Biobanks and Biobank Networks

Eva Ortega-Paíno* and Aaro Tupasela**

*Coordinator for the BBMRI.se Service Centre for Southern Sweden. Lund University (Sweden).
**Section for Health Services Research, Department of Public Health, University of Copenhagen.

Introduction

Even after the first decade of the 21st century, one of the remaining and major scientific challenges to overcome is the lack of reproducibility in medical research. In mid 2015, Freedman and others, published an article in PLOS Biology where low reproducibility rates were shown within life science studies. The authors highlighted that half of the annual research investment in the USA is lost due to irreproducibility. This irreproducibility caused mainly by pre-clinical issues, such as the study design, data analysis, laboratory protocols, inappropriate biological reagents and reference materials, point to the lack of a standards and best practice frameworks. Furthermore, this continuous lack of standardization, for example in biological material, not only in the collection, but also in the storage, could lead to the discovery of false biomarkers. Poste, in his comment published in Nature stressed the huge gap existing between literature, where more than 150 000 articles have been published claiming thousands of potential biomarkers, and the reality, where only around a hundred are used in the clinic. Therefore, there is an increased need for more stringently characterized biospecimens in biomedical research.

In the same article, Poste noted the need of demanding from funding agencies, such as the NIH, the European Framework Programme among others, better returns for the money invested in biomarker research. To achieve these goals, it is important to establish common standards, not only for storage and for processing of biospecimens, but also for discovery and validation, as well as harmonized data from those samples. As if this were not enough, it should also be compulsory to possess the whole range of cross-disciplinary capabilities needed for translational research, that is to say, from laboratory to the clinics. Therefore, partnerships between academic laboratories, industry and health care should be encouraged to form larger research networks and the need of high quality samples made a priority.

In the era of high throughput technologies and the search for new potential biomarkers for therapy, prognosis or even diagnosis, it is of great importance to gather samples from donors where the pre-analytical variables have been minimized as much as possible. These pre-analytical variables are classified in a first group called in-vitro (those that happens in the process from “needle to freezer” such as hemolysis, delay in centrifugation, transport, etc.) and those variables such as use of drugs, alcohol or smoking as well as other factors like sex, weight, age, habitat and lifestyle, among others, known as in-vivo variables. All these variables, in-vivo and in-vitro, should be controlled, annotated and taken into account while selecting individuals for a study.

---

with the aim of matching as many of them as possible for reducing interferences in the study and therefore, the ulterior analysis.

So, where can these very high quality specimens be found? The answer is simple: from harmonized and standardized **biobanks**. Time Magazine listed biobanks as “one of the ten ideas that will change the world”.³ According to Hewitt and Watson, biobanks are collections of human, animal, plant and microbial samples, and the data associated to them that must be managed according to professional standards.⁴ The most common biobanks contain human material such as whole blood, plasma, serum, urine and sperm among others. According to Asslaber and Zatloukal there are many different types of biobanks the most important ones for research, however, are the population-based biobanks and the disease-oriented biobanks.⁵

The fast development of the “-omics” platforms such as genomics, transcriptomics, metabolomics, proteomics, etc… has produced an enormous amount of genotypic and phenotypic data stored in several databases around the world as well as the samples associated to them are spread in different biorepositories. Without a proper understanding of the conditions under which the samples from which this data has been derived, it is unclear to what extent this data is meaningful in a clinical sense. The process of biobanking, therefore, needs better quality control from the very beginning of the biobanking process to the very end, when data is being analysed.

As we mentioned, there are many different types of biobanks. Without any doubt, the longitudinal population-based biobanks are the most common format. Frank has noted that the population biobanks in the Nordic countries is a type of collection where the entire population is a cohort.⁶ It is well known, given the organization of the Nordic countries where each citizen gets a personal number at the moment of birth, that this makes the Nordic countries a ‘goldmine’ for biobanking. This personal number or identification number is used to link information from each individual to all National Quality registers as well as healthcare records. This provides researchers with unparalleled information about the state of health of an entire country and allows them to develop predictive disease models. These predictive approaches can be used as prevention in medical programs, since volunteers with or without any disease are donating samples at different time points with the aim of finding biomarkers/genes that can lead and explain different diseases caused, mainly, by environmental factors as well as, for example, the diet. The typical samples stored for this purpose are isolated DNA and whole blood, as well as the data associated to them (environmental factors, family history, lifestyle, etc.).

Some notable examples of this are, the VIP (Västerbotten Intervention Program), which is a long-term project in the Northern part of Sweden where individuals with age of 40, 50 and 60 are invited for screening. This study gathers in total around 254,000 inhabitants. The participants are asked to complete a questionnaire concerning various lifestyle factors and to donate a blood samples to be frozen for further research.

---

purposes. In Denmark, the Danish Blood Donor Study (DBDS), for example, has collected samples from over 60,000 Danes and it seeks to be the largest collection in Denmark. Another example in Skåne, the southern part of Sweden, is the MDC (Malmö Diet and Cancer study). This study started in the early 1990s as a screening survey in the middle-aged population of Malmö. The study gathered 28,000 participants, all from Malmö. As in the VIP study the individuals had to answer a survey as well as donating blood. Among all the participants, 62% were women. Cardiovascular risk factors were measured in a random subsample (n=6000). Sixteen years after the first recruiting process, a new clinical examination and blood sampling was performed (n=3700). 

Population based-biobanks are also a good complement to disease-biobanks, in which the samples from the individuals have been stored after diagnosis and/or even treatment and will be used to answer some specific disease-driven questions. Unlike population-based biobanks, disease-oriented biobanks can be used soon after the collection of samples has concluded, with a lower cost associated. An example of this type of study, and the biobank associated to it, is the SCAN-B (Swedish Cancerome Analysis Network-Breast) project in the region of Scania (Sweden). This multi-center initiative started in 2010, enrolling hospitals from the southern healthcare region (Malmö, Lund, Helsingborg, Växjö, Halmstad, Kristianstad and Karlskrona). Three years after the region of Uppsala joined SCAN-B as well. The goals of the project are focused on finding new ways for diagnosis, prognosis and treatment.

Another example of a disease-based biobank is the Auria Biobank in Finland (www.auria.fi). Established in 2012 the biobank collects samples from patients coming into the regions hospitals after they have consented. In 2014, the biobank also transferred a large amount of samples from the University of Turku Hospital pathology collection in accordance with the Finnish Biobank legislation, which was established in 2013. What makes these types of biobanks unique and important is that they are able to capture a large amount of samples from specific hospital regions and track patient outcomes over long periods. It is hoped that these types of biobank configurations will also provide industry with new resources for conducting research.

These examples of the different formats of biobanks show not only the need of collecting biospecimens to find an answer to why some people develop diseases and others don’t, but it also points to the need to develop common standards between the

biobanks to ensure that samples are of similar quality, as well as to make sure that the results from one study are comparable to results from other studies. Furthermore, in relation to more rare conditions it is difficult to find the necessary samples in just one biobank. This means that researchers may need to request samples from numerous biobanks in order to gain sufficient statistical significance. Ensuring that samples are of similar quality and standards means that their comparison and use generates fewer errors.

The rise and development of biobank networks

During the past few decades the emergence of biobank networks have become an important framework and activity through which the hundreds of biobanks around the world coordinate and develop their activities. According the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI), in Europe alone, there are over 500 biobanks, which contain more than 60,000,000 samples.\(^\text{13}\) (Holub et al., 2016: 560). With such a large number of collections and samples, the need for common standards and procedures is of utmost importance in order to improve the quality and reliability of research results.

There is a clear need for collaboration in biobanking and the development of biobank networks serves as an important medium through which collaboration is facilitated. It is easy to understand that not only cancer, but also other diseases are present worldwide and therefore this implies the need of comparing between different ethnic groups and sub-populations, patterns of disease, as well as prevalence in different areas in the world. Furthermore, some commentators have noted that it is difficult to publish findings of genetic association studies unless several populations have ben combined, preferably from different countries.\(^\text{14}\) This increased pressure for generating replicable results across populations has also contributed for an increased demand for standards among biobanks. Considering all this, it is important to join forces and gather samples from as many biobanks in the region, in the country and all over the world as possible.

During the last two decades, different international networking initiatives in Europe, North America, Africa and Australia such as P3G, CTRNet, TuBaFrost, BBMRI-ERIC, Australasian Biospecimen Network among others, have contributed to the proliferation of biobanking networks. These actors are central in the collection, processing, storing and dissemination of biological samples and the clinical data associated to them and are, therefore the pillars and the key of success of well-organized and managed biobanks.\(^\text{15}\) Well-functioning network-activities will be reflected in subsequent high quality research.

Biobank networks are not only a way to provide samples and data to researchers. They also serve important social and normative functions. As Mitchell and Waldby have noted the collection and exchange of human tissue samples is not just a technical, but

\(^{13}\) P. Holub et al., ‘BBMRI-ERIC Directory: 515 Biobanks with Over 60 Million Biological Samples’ (2016) Biopreservation and Biobanking 14, 559–562.

\(^{14}\) J. Kere, ‘Miten Suomessa kerättyjä DNA- ja kudosnäytteitä voidaan hyödyntää? - [How can DNA and tissue samples collected in Finland be utilized]’ 123 Duodecim 864.

also a social act. The processes of collecting, storing and sharing samples, constitutes, according to them a 'tissue economy', which serve as a conduit through which values and norms are also established and shared. Within biobanking networks there are a plethora of examples through which social norms and standards are established. These include forms of informed consent, establishment of norms regarding data sharing practices, discussions surrounding appropriate pricing, as well as quality standards for the collection and storage of samples.

Biobank networks also play an important role in policy setting, both nationally and internationally. The need to organize biobanking activities nationally is in part related to the international standards and requirements set forth within international networks. In Finland, for example, the national biobanks have sought to organize themselves using the hub and spokes model (http://www.bbmri.fi/). The same organization principle can also be seen in other countries, such as Sweden and Norway, for example. The coordinating bodies of biobank networks also play an important role in organizing and developing common resources for network members. This can include training and education, as well as serve in facilitating the setting of common standards. Biobank networks also provide an opportunity for biobanks to gain access to new resources, such as financial opportunities gained through entering into research collaborations with new partners and collaborators. In this sense, biobank networks can serve as accelerators of research and provide important new avenues through which scarce and expensive resources are better used and maintained in the end.

**Biobanks: a matter of quantity and quality**

It is important to stress that quality in biospecimens (samples gathered in a harmonized, standardized manner as well as the data associated to it) is a compulsory requirement, but this is not enough for developing high quality projects. For this purpose, it is needed to have enough samples that can provide statistical power to the analysis. Therefore, quantity is also a key factor to consider. For many of the non-communicable diseases (cardiovascular diseases, cancer, chronic respiratory diseases and diabetes), enough individuals are enrolled in research projects within the country and/or even the region. However, while talking about rare diseases the issue of quantity is relevant, since the number of cases is very low in each country and the need of statistical power is a must in research. We cannot forget either that biopharmaceutical companies operate in a global world, whereas the fact of diversity of ethnicity needs to be also addressed. This need opens up the issue of international collaborations for research, which makes biobank networks of great significance. Even for common diseases, there is a need for larger samples sizes. Despite Large Prospective Cohorts (LPC) having thousands of

---

samples, one cohort may not have enough samples of a specific type to reach statistical significance. Therefore, biobank networks serve an important function of providing researchers the ability to search for specific sample types of specific patients in a broad range of biobanks from around the world.

These valuable collections would need to be accessible, managed in a proper manner as well as standardized, as it has been highlighted before in this chapter. The information about these samples should present a common ontology so researchers could exchange not only the samples but also the data associated to them. For this purpose, there have been different approaches. One of these, MIABIS (Minimum Information About Biobank Data Sharing) launched within BBMRI network had as their goal the harmonization of information from biobank samples throughout Europe.\(^{22}\) The goal of MIABIS was to develop a minimum data set, consisting of 52 attributes, for biobanks and studies in human biospecimens. This data, at the aggregate level would allow the reutilization of samples and furthermore will save money and resources in the long run, increasing at the same time the quality of research by number of specimens and their data associated.

Although collaboration and exchange are key concepts for the success of any biobank network, as well as data sharing, it is also important to remark that this data belongs to a number of individuals who must have their integrity protected. For this reason, some questions have raised concerning legal, ethical and societal issues (ELSI) in Biobanking. It will important to discuss the individual's' rights and protect their integrity.

In the next section, we will briefly describe the BBMRI-ERIC infrastructure in Europe, which has become one of the largest biobanking networks in the world. The network highlights the possibilities that formalized biobanking networks offer to its members in relation to coordination of activities, as well the development and harmonization of common standards.

**BBMRI-ERIC as a platform for biobank networks**

BBMRI-ERIC is a European research infrastructure, which was first initiated in 2008 when it entered the European Research Infrastructure Preparatory Phase of the ESFRI roadmap. Between 2008 and 2011 their main funding came through the European Framework Programme 7. Five years later, in 2013 BBMRI-ERIC was awarded legal status and since then has grown to comprise 16 full member states (Austria, Belgium, Czech Republic, Estonia, Finland, France, Germany, Greece, Italy, Latvia, Malta, the Netherlands, Norway, Poland, Sweden, and the United Kingdom), 3 observers (Cyprus, Switzerland, Turkey) and one international organization (IARC/WHO). However, countries such as The Netherlands and Sweden were pioneers in the organization and already initiated to link their BBMRI activities in 2009. The research infrastructure offered by BBMRI-ERIC is important in that it is designed as a research infrastructure whose goal is to support interaction between members, as well as coordinate activities. The ERIC status facilitates the collaboration between biobanks and biomolecular resources not only for giving access to their collections, but also by providing expertise

---

and services free of charge. As such, it is not a funding organization, which funds research, but rather serves to facilitate activities, which utilize samples and data from biobanks around Europe and elsewhere as well.

From a historical perspective, the BBMRI network sought to improve the interoperability of biobanks in Europe. As they note in one of their early publications:

The move towards a universal information infrastructure for biobanking in Europe is directly connected to the issues of semantic interoperability through standardized message formats and controlled terminologies. The information infrastructure has become a critical component in life sciences research.\(^{23}\)

This perspective highlights one of the major visions driving biobank networks, namely the possibility of combining both samples and data from biobanks across Europe, and ultimately the world. The concept of interoperability of biobanks and data derived from biobanks had its roots in early attempts at analyzing data from different population studies in Europe. The GenomEUtwin project was one of the first such attempts at developing a platform in which samples from over 600,000 twins could be analysed. In order to make the analysis possible, the groups had to develop common standards and terms to describe the different variables that they had used, such as BMI.\(^{24}\) From the experiences gathered in the GenomeEUtwin project and others, it became apparent that in order to further facilitate the interoperability of other types of biobanks, there needed to be a common ‘language’ that could be used to facilitate the comparison and use of samples and data. This “comparability” has been one of the main tasks tackled in the Global Biobank Week held in September 2017 in Stockholm. The need of harmonization within the different biobanks around the world was extensively discussed. This harmonization should be taken in consideration in the different steps from the needle to freezer and back to the patient and engulf not only the sample but also the data associated to it.\(^{25}\)

The development of the BBMRI research infrastructure took place within the context of the European research infrastructure (ERA) which has been developed, among other things, with the Innovation Union concept in mind.\(^{26}\) The interest to develop biobanks as a common resource for research, however, has not been the mission of the European Union alone. The OECD has also had an interest in promoting the development and use of biobanks to bolster research activities, as well as innovation.\(^{27}\) In addition to the

bolstering of biomedical research in universities, the hope with these initiatives and efforts has been to encourage the use of biobanks in industry.\textsuperscript{28}

Within the BBMRI network, each member state is expected to sign a memorandum of understanding (MoU) in which they agree to fund the development of biobanking activities and facilities within their own countries. This commitment is important in that it is focused on the support for structures that maintain biobanking facilities and activities. At the same time, member states are expected to fund research, which utilizes biobanking samples. Given the vast differences in the number and scope of biobanks in various countries, the ways in which each member state organizes its activities is left up to them. A common example, however, regarding this system has been called the “hub and spokes” model where within a country there will be a central coordinator, which supports the various biobanks around the country.\textsuperscript{29}

BBMRI-ERIC developed also the idea of the Expert Centre (EC). An EC, following international standardized conditions, should be the link between public and private sectors carrying out the analysis of the biological specimens. These EC according to van Ommen, GJB and others, will join the expertise of both academia and industry and this alliance will lead to a win-win scenario where 1) collaborations in research will increase; 2) the flow of information, knowledge, data and technologies will be shared between partners; 3) ethical and legal issues will be ensure under the frame of the adequate code of conduct 4) resources will be managed and used in a more efficient manner; 5) questions regarding ownership (of samples and data) will be minimized and finally 6) an innovation will be the outcome of a competitive research and development within the partners.\textsuperscript{30}

There are several EC that could be used as examples. Among them BARCdb, platform developed in Sweden, links biobanks to molecular technology platforms in the worldwide -omics scenario.\textsuperscript{31} (www.barcdb.org) Other examples could be the EGC (Estonian Genome Centre) and SciLifeLab (Science for Life Laboratory), a state of the art national initiative of large-scale molecular analyses. SciLifeLab offers services to both academia and industry and provide also, apart for new developed technologies, bioinformatics support.

In the previous section, we discussed the MIABIS system that has been designed as a system for the minimum amount of information that should be available regarding biobank samples. From a broader perspective MIABIS plays a crucial role in the functioning and role of biobank networks. From a political and social perspective, the MIABIS system, its development and implementation can be seen as a technical translation and implementation of European research policy in which human tissue samples and their related information become a system through which European

\textsuperscript{30} Supra note 26.
integration can be achieved on one level.\textsuperscript{32} In this sense biobank networks serve a broader political role in unifying European research practices. The MIABIS system serves as the basis for the smoother and more efficient functioning of the European tissue economy in relation to the biobanks, which agree to meet these standards. It helps to create and facilitate a system of exchange through the establishment of common standards.

The network as a metaphor for activity across borders is also emblematic of the emergent governance structure relating to biobanking where the nation state is no longer the central actor. As Andrew Barry has noted, biobanking can be considered to form a “technological zone” in which governance of activities, as well as the establishment of standards operates beyond the traditional nation-state borders.\textsuperscript{33} Technological solutions, such as MIABIS, are seen as systems, which address the problems, which since states are unable to adequately answer and solve. At the same time, however, the networks, which operate across borders form new locations of governance and power through which information and data are channeled.

The BBMRI network also works towards establishing common services which its members can make use of. One example of this is the Common Service ELSI, which seeks to help discuss, establish and foster ethical, legal and social norms surrounding biobanking in the member states. Given the broad range of practices, which have developed around biobanking during the many decades that it has been practiced, it has become an important component to develop common ethical and legal guidelines for operating and managing biobanks. Although many of the legal issues which biobanks face are also dictated by ‘hard law’, such as the EU’s data Directive, there are also a large amount of practices and procedures that can be guided by ‘soft law’ or commonly agreed upon good practices within the biobanking community.

Apart from the joint efforts in biobank ontology, summarized in MIABIS, BBMRI-ERIC has also organised working groups for the comparison and establishment of Quality Standards based on the CEN/TS. BBMRI-ERIC quality services provide the required tools, knowledge and experience on quality management for biobanks and research on biomolecular resources.\textsuperscript{34} These working groups have produced different self-assessment surveys or pre-examination processes for biobanks with the aim of evaluating quality of DNA, RNA and proteins in different tissues such as snap frozen and FFPE tissues, as well as for venous whole blood. They also evaluated the pre-requisites in metabolomics for urine, serum and plasma, covering thus all possible pre-analytical variables existing in the biobanked samples.

All these efforts together with the development of IT tools such as sample/data locator and negotiator carried out by the BBMRI-ERIC common service IT will pave the way to assist and advice in the collection, distribution and use of biobank high quality samples. This will facilitate the good practice for high quality research, which finally will revert to the patient in Precision Diagnostics and Personalized Medicine. It is hoped that this will lead to better health and therefore, quality of life.

Discussion

In this chapter, we have tried to elucidate the importance that biobanks and biobank networks play in biomedical research. Biobanks are central to the enactment and implementation of collection practices in which good quality samples and related information are collected from the population. This activity is central to the production of good research where the results are both replicable and valid. Maintaining and striving for good quality is also a valuable social goal since most of the large biobank collections, whether they be cohort studies, disease-based collections or diagnostic collections in hospitals have been collected and are maintained using public funding.

Biobank networks also serve an important function of giving rise to larger collections when they are combined. In both the study of complex diseases, as well as rare conditions, it is necessary to have access to enough samples to be able to derive meaningful conclusions. In many cases a single biobank is not enough. As such, biobank networks facilitate the process of identifying which biobanks have the relevant samples that a researcher may need to conduct their study.

Both quality and quantity in biobanking should provide a more reliable avenue towards identifying and developing more accurate and clinically valid diagnostics and treatments for patients. As we have discussed above, both quality and quantity is needed if we are to trust the biomarkers that are developed to identify disease causing genes, as well as gene-environment interactions. Without robust quality systems, as well as sufficient sample sizes, this may prove to be increasingly difficult and unreliable. Many of the purported benefits of Personalized Medicine are based on the assumption and goal that we will have reliable and valid biomarkers for diagnostics and treatment. In order to facilitate research into personalized medicine it is necessary to provide the necessary research material, which can be used to study and validate possible biomarkers for personalized medicine.

In a broader sense, biobanks and biobank networks can be seen as technological implementations of political goals where, at least in Europe, the goal is to produce a more cohesive research market that will facilitate and accelerate the discovery of new biomedical innovations. Biobanks are also seen to serve as the basis of future innovations, as well as drivers for the knowledge-based bio-economy. While facilitating these political aspirations, biobank networks also give rise to new governance structures in which new ethical and legal (soft law) norms are established and exercised. These norms have considerable implications in relation to how we perceive the acceptability of new practices in relation to the collections, distribution, and use to biobanking samples. As such biobanks and biobank networks play a crucial role in maintaining the social and technical norms, which allow for tissue economies to emerge and function.