Biobank networking for dissemination of data and resources: an overview

Abstract: In response to the increasing global demand for high quality biospecimens and data for biomedical research, biobanking is rapidly gaining popularity as an efficient and user-friendly platform for translational research. The advent of increasingly sophisticated technologies for specimen and data analysis, in the face of growing economic pressures, are converging to encourage consolidation, centralization, and harmonization of biobanks into networks. Several types of biobank networks exist worldwide. Individuals involved in network establishment and day-to-day function hail from varying disciplines, including health care, academia, information technology, and the pharmaceutical industry. However, they may work together within and between networks to enhance the rapid progression of patient/donor-centered research through standardization of procedures and robust quality management systems. Regularly updated standards, policies, and guidelines are published by large biobanking organizations and made available to biobankers and those interested in biospecimen science. A biobank network’s ability to reliably disseminate specimens and data depends on a variety of factors, including a well stocked inventory, a robust information technology platform, and adequate support, including the goodwill of collectors who supply specimens, and of end-users who return experimental data to the network. High quality experimental data may be recycled, thus accelerating biomarker discovery. Access to large amounts of personal data, however, carries risk, and ethical issues surrounding data protection are of paramount importance. All biobank networks require data security measures in keeping with local ethical and legal requirements. Return of results to individual donors is another emerging ethical and administrative challenge for biobank networks as technology steadily increases the overlap between research and patient care. Finally, as the bioresource impact factor concept is further developed, and as more scientific journals require biospecimen and source details in submitted manuscripts, biobank networks will be securely established as an essential platform for biomedical research.

Keywords: biobank, networking, data, resources, dissemination

Introduction

Biobanking is revolutionizing the way we look at science, medicine, and health. A recent issue of the popular press periodical *Time Magazine* highlighted biobanking as one of the top ten ideas changing our world.¹ Loosely defined, a biobank (or biorepository) is a collection of biological samples and associated data, systematically organized for use by stakeholders, such as researchers and health care providers.² Biobanks and their networks differ in scale, scope, purpose, structure, and level of government support, and as such, a clear definition of the term “biobank” has been the subject of some debate within the biobanking community.³–⁵

Biobanking as an activity is not new, but it is emerging as a young scientific discipline in and of itself, supporting other disciplines in the quest to improve science and...
ultimately public health. In this paper, some of the salient points regarding activities and debate surrounding specimen and data dissemination in biobank networks are discussed. Issues surrounding access, privacy and security, and return of results are also reviewed.

What types of biobanks exist?
The traditional practice of collecting and storing human biological samples (biosamples) and data for research has been fragmented, without communication between investigators and without standardization of collection, processing, or storage methods for the samples and data in question. This approach increases the likelihood of irreproducible results and the waste of scarce biological and financial resources, and leads to much slower progress in biomedical research and development. According to the broadest definition, a biobank may collect specimens from plants, animals, and the environment, and a variety of types of biobanks are in operation all over the world. However, biobanks that collect human specimens are the most common.

Informed consent – the first step in biobanking
Informed consent in biobanking has been extensively studied. While the principles of informed consent for biobanking are similar to those of consent for specific research studies, key differences exist. First, in biobanking, samples given at one point in time may be used for several future studies, each generating new data. Moreover, the potential exists for studies involving analysis of previously generated data alone, without any need for the original samples. Most importantly, at the time of consent, the future studies in which the donors’ samples and data will be used are unknown. As such, consent in the context of biobanking cannot truly be considered informed, given that as technology becomes more robust and more information can be gleaned from each donated sample, the potential risks and benefits to the donors cannot be fully known. To address these differences, a shift away from specific informed consent towards broad consent and/or tiered consent is gaining support. In broad consent, a donor consents to the use of their samples and data for future unspecified research with no need for recontact. Tiered consent offers donors a range of consent options, from broad consent through consent on their behalf by an institutional review board to recontact. Consent models (Table 1) based on a number of regulations, guidelines, and laws worldwide are elegantly reviewed by Salvaterra et al. Recent studies of informed consent form contents within and across countries in Europe show that tremendous heterogeneity exists, and suggest that improved harmonization may lead to a better balance between donors’ rights, protection, and autonomy, and research potential, through facilitation of biobank networking.

Table 1 Consent models for biospecimen donation and future use of samples and data

<table>
<thead>
<tr>
<th>Consent model</th>
<th>Definition</th>
<th>Language</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prospective opt-in</td>
<td>Requires active consent to donate</td>
<td>&quot;My samples and data may be used in any future ethics board-approved studies.&quot;</td>
<td>• No need to recontact.</td>
</tr>
<tr>
<td>Prospective opt-out</td>
<td>Requires active refusal to donate</td>
<td>&quot;My samples and data may be used in any future ethics board-approved studies in the following fields: cancer/infectious disease/metabolic conditions/degenerative diseases/etc.&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;I wish to be contacted for consent to release my samples and data for any future ethics board-approved studies outside this/these field(s).&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;My samples and data may be used in the research study only.&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;My samples and data may be used in any future ethics board-approved research studies.&quot;</td>
<td></td>
</tr>
<tr>
<td>Broad/Blanket</td>
<td>Allows any future use</td>
<td>&quot;My samples and data may be used in any future ethics board-approved studies.&quot;</td>
<td>• Considered most pro-research.</td>
</tr>
<tr>
<td>Categorical/Tiered/</td>
<td>Use options chosen from a &quot;menu&quot;</td>
<td>• No need to recontact.</td>
<td>• Considered &quot;uninformed&quot; by some.</td>
</tr>
<tr>
<td>Multi-layered</td>
<td></td>
<td>• Administrative burden: low.</td>
<td>• Administrative burden: low.</td>
</tr>
<tr>
<td>Specific</td>
<td>Consent for specific project only</td>
<td>&quot;My samples and data may be used in the current research study only.&quot;</td>
<td>• Donor control level: intermediate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Requires reconsent for additional use.</td>
<td>• Administrative burden: high.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Donor control level: high.</td>
<td>• Donor control level: high.</td>
</tr>
</tbody>
</table>
The varying models of informed consent answer some of the challenges to the legal and ethical validity of consent to participate in biobanking. In these models, participation is framed as a “moral act of a responsible citizen”, with the consent process engendering trust that the scientific establishment will use biobank materials and data for the public good. It may be argued that changes to the consent process should involve participant groups, thus encouraging the public to take on a more active role as citizen scientists, beyond passive donation, in the realm of biobanking. It is recognized that this role may be limited to biobanking in industrialized countries where higher literacy rates, better funding, and more developed infrastructures for dissemination of information exist. In less developed nations, such as those across Sub-Saharan Africa, there is a growing trend toward tiered consent, but guidelines and legislation surrounding consent vary from country to country.

What are the benefits of biobanking?

Human biosamples and clinical data may be collected routinely in the health care setting, during the course of clinical trials, or as part of long-term large-scale population studies of healthy volunteer donors. The widespread availability of large amounts of samples and data may be beneficial to a wide array of stakeholders, including patients, health care providers, academics, policymakers, and pharmaceutical and other health care businesses. The advantages of coordinated standardized human biobanking include stronger clinical correlation because specimens are well annotated with clinical data, implementation of best practices to increase sample quality and reduce redundancy and waste, job creation to support collection, processing, and annotation, infrastructure leverage for smaller research organizations, research efficiency, and acceleration of progress to personalized medicine. From a global economic perspective, it is estimated that the market value of demand for human biospecimens and related services is expected to reach close to $2.3 billion by 2015. Moreover, it is estimated that if in the USA alone, standardized specimen collection were to reduce annual clinical trial costs by just 2%, the research community would save $115 million per year.

Functions and advantages of biobank networks

For biomarker discovery and development of new targeted treatments, research is required in large cohorts of well annotated biosamples. While individual centers may enroll small numbers of patients in any one research study, a biobank network of several centers using the same standard operating procedures for consent, sample procurement, processing, storage, and annotation, may more rapidly, efficiently, and more cost-effectively reach the required cohort size to achieve statistically significant results. A biobank network links biobanks by ethos, policies, procedures, and attention to quality. A common database is used and there is a consensus regarding release of samples and data for research. Biobank networks achieve the large sample numbers required for collaborative translational projects quickly and efficiently. They maximize the use of resources, follow international best practices, and because standard operating procedures are monitored and quality management is a keynote, samples are of optimal quality. A common database allows the construction of a web portal and sample catalog, with restricted sample and clinical data access limited to authorized end-users. Biobank networks are ideally suited to high volume translational research, national and international projects, and academia-industry partnerships, and can provide numerous samples if requested by biopharma companies.

Today there are national or regional biobank networks in 23 different countries and some European Union networks (Table 2). In an effort to highlight some of the challenges faced by biobanks, Shickle et al introduced a classification of biobank networks by category, including storage, bring-and-share storage, catalog, partnership, contribution, and expertise, recognizing that many networks may have features overlapping more than one category, due to the scope of their activities. Regardless of their type, all biobank networks require adequate information technology platforms to support biobanking activities. Such platforms include not only the minimum data and metadata sets related to the donors (eg, epidemiology, pathology, and follow-up information) and the donated samples (processing, inventory specifications, distribution parameters), the capability of building a web portal, but also subserve additional functions, such as those related to investigator project management, billing, quality control, and absorption of study results.

Biobank networks and industry

There is a gradually evolving rapprochement of biobank networks and industry. This is hardly surprising, given that all new biomarkers and targeted therapies after validation are manufactured and sold by biopharma companies. Moreover, the raw materials for what may become vital new patient interventions are stored in our biobank networks. The European, African, and Middle Eastern Society for
<table>
<thead>
<tr>
<th>#</th>
<th>Network name</th>
<th>Website</th>
<th>Details</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Australasian Biospecimen Network</td>
<td><a href="http://www.abrn.net/">http://www.abrn.net/</a></td>
<td>8 centers</td>
<td>Australia</td>
</tr>
<tr>
<td>2</td>
<td>Biobank Graz</td>
<td><a href="http://www.meduni-graz.at/en/biobank">http://www.meduni-graz.at/en/biobank</a></td>
<td>15 centers</td>
<td>Austria</td>
</tr>
<tr>
<td>4</td>
<td>Latin America and Caribbean Tumor Bank</td>
<td>NA</td>
<td>11 countries</td>
<td>Brazil</td>
</tr>
<tr>
<td>5</td>
<td>The Canadian Tumour Repository Network (CTRnet)</td>
<td><a href="http://www.ctrcanada.ca/">http://www.ctrcanada.ca/</a></td>
<td>6 sites</td>
<td>Canada</td>
</tr>
<tr>
<td>6</td>
<td>Canadian Virtual Brain Tumour Bank</td>
<td><a href="http://www.braintumourbank.ca/">http://www.braintumourbank.ca/</a></td>
<td>4 sites</td>
<td>Canada</td>
</tr>
<tr>
<td>7</td>
<td>China National Genebank</td>
<td><a href="http://www.nationalgenecentre.org/en/new2.html">http://www.nationalgenecentre.org/en/new2.html</a></td>
<td>7 hospitals</td>
<td>People’s Republic of China</td>
</tr>
<tr>
<td>8</td>
<td>Danish National Biobank</td>
<td><a href="http://www.biobankdenmark.dk/">http://www.biobankdenmark.dk/</a></td>
<td>National</td>
<td>Denmark</td>
</tr>
<tr>
<td>9</td>
<td>The European Human Frozen Tissue Bank Network – Tubafrost</td>
<td><a href="http://www.tubafrost.org">http://www.tubafrost.org</a></td>
<td>Eurocan platform</td>
<td>European Union</td>
</tr>
<tr>
<td>10</td>
<td>EuroBioBank</td>
<td><a href="http://www.eurobiobank.org/">http://www.eurobiobank.org/</a></td>
<td>10 countries, 21 centers</td>
<td>European Union</td>
</tr>
<tr>
<td>11</td>
<td>The Danubian Biobank Consortium</td>
<td><a href="http://www.danubianbiobank.de/">http://www.danubianbiobank.de/</a></td>
<td>20 centers</td>
<td>European Union</td>
</tr>
<tr>
<td>12</td>
<td>National Biobank Network</td>
<td><a href="http://www.healthbio.fr/healthbio.asp?viewID=331">http://www.healthbio.fr/healthbio.asp?viewID=331</a></td>
<td>For industry</td>
<td>Finland</td>
</tr>
<tr>
<td>13</td>
<td>CEPH Biobank</td>
<td><a href="http://www.cephb.fr/index.php">http://www.cephb.fr/index.php</a></td>
<td>National</td>
<td>France</td>
</tr>
<tr>
<td>14</td>
<td>French Biobanques</td>
<td><a href="http://www.biobanques.eu/">http://www.biobanques.eu/</a></td>
<td>77 members</td>
<td>France</td>
</tr>
<tr>
<td>15</td>
<td>The German Biobank Registry (TMF)</td>
<td><a href="http://www.biobanken.de/en-gb/home.aspx">http://www.biobanken.de/en-gb/home.aspx</a></td>
<td>111 elements</td>
<td>Germany</td>
</tr>
<tr>
<td>16</td>
<td>Biobank Ireland Trust</td>
<td><a href="http://biobankireland.com/">http://biobankireland.com/</a></td>
<td>4 centers</td>
<td>Ireland</td>
</tr>
<tr>
<td>17</td>
<td>Israel Collaborative Bioresource for Research &amp; Development</td>
<td>NA</td>
<td>4 centers</td>
<td>Israel</td>
</tr>
<tr>
<td>18</td>
<td>The Organization of Italian Regional Networks</td>
<td>NA</td>
<td>6 regions</td>
<td>Italy</td>
</tr>
<tr>
<td>19</td>
<td>Telethon Network of Genetic Biobanks (TNGB)</td>
<td><a href="http://biobanknetwork.org/">http://biobanknetwork.org/</a></td>
<td>10 centers</td>
<td>Italy</td>
</tr>
<tr>
<td>20</td>
<td>The Biobank Japan Project</td>
<td><a href="http://www.src.jrc.it/english/project/person/index.html">http://www.src.jrc.it/english/project/person/index.html</a></td>
<td>National</td>
<td>Japan</td>
</tr>
<tr>
<td>21</td>
<td>HUNT Biobank, Norway</td>
<td><a href="http://www.ntnu.edu/hunt">http://www.ntnu.edu/hunt</a></td>
<td>National</td>
<td>Norway</td>
</tr>
<tr>
<td>22</td>
<td>National Depository of biological samples</td>
<td>NA</td>
<td>National</td>
<td>Russia</td>
</tr>
<tr>
<td>23</td>
<td>Spanish National Biobank Network</td>
<td><a href="http://www.bancoadn.org/en/home.htm">http://www.bancoadn.org/en/home.htm</a></td>
<td>63 institutions</td>
<td>Spain</td>
</tr>
<tr>
<td>25</td>
<td>The National Biobank Project</td>
<td><a href="http://www.biobanks.se/">http://www.biobanks.se/</a></td>
<td>10 nodes</td>
<td>Sweden</td>
</tr>
<tr>
<td>26</td>
<td>Biobank Suisse – A Collaborative network</td>
<td><a href="http://www.biobank-suisse.ch/">http://www.biobank-suisse.ch/</a></td>
<td>8 centers</td>
<td>Switzerland</td>
</tr>
<tr>
<td>27</td>
<td>The Netherlands Dutch Biobank Network</td>
<td>NA</td>
<td>National</td>
<td>The Netherlands</td>
</tr>
<tr>
<td>29</td>
<td>UK DNA Biobank Network</td>
<td><a href="https://www.ukbiobank.ac.uk/">https://www.ukbiobank.ac.uk/</a></td>
<td>National</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>30</td>
<td>The Cooperative Human Tissue Network (CHTN)</td>
<td><a href="http://www.chtn.ncl.ac.uk/">http://www.chtn.ncl.ac.uk/</a></td>
<td>6 divisions</td>
<td>USA</td>
</tr>
<tr>
<td>31</td>
<td>Wales Cancer Bank</td>
<td><a href="http://www.walescancerbank.com/">http://www.walescancerbank.com/</a></td>
<td>9 centers</td>
<td>Wales</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not available.
Biobanking Translate Working Group found that 72% of biobanks already had or would like to develop industry partnerships and provide samples for industry. Samples can be acquired directly by industry from the hospital biobank network or indirectly through a company which acts as an intermediary between the biobanks and industry.

There are different types of partnerships and collaborations, but most are project-based. An intermediary company may either pay an administration fee or provide funding for personnel and equipment in return for obtaining samples requested by biopharma companies. It is very important that industry agreements provide a net benefit for the biobank network, and not function merely as a quid pro quo provision of services. Otherwise, key biobank staff might be distracted from their primary function of real-time biobanking. Also, industry and honest broker biobankers have very different perspectives on the ease of collecting, storing, and retrieving biosamples. Hence, practical business advice is very helpful for negotiations, and agreements should have a trial period to test assumptions of productivity.

Industry is often cited as a potential funder for biobank networks. This is true, but specification of the amount and duration of the funding is outside the control of the biobank network, unless the project is very large. Much industry funding can be regarded as a series of small grants. An intermediary company that sources samples for pharma may produce a more lasting partnership, because such companies are likely to have continuous requests from pharma. An industry partnership might therefore contribute to the sustainability of a biobank network. The collaboration may exist simply for a biobank network to provide samples and data, with reimbursement by an administrative fee, expenses, or part-time personnel. A true partnership to complete a project together maximizes resources and expertise, and usually involves secondment or hiring of personnel to work part-time for the biobank network (or the translational laboratory) as well as on the project at hand. The biobank network must evaluate very carefully the value of the work done on the project, so as to be reimbursed appropriately.

Science of biobanking

Biobanking as a scientific discipline is growing at a rapid pace. A cursory PubMed search for English language publications using the terms “biobank” and “bio bank” yielded a steady increase from the early 2000s (Figure 1). In addition to supporting studies, the results of which may lead to discovery of new diagnostic and prognostic biomarkers as well as treatments, biobank networks also support a new research discipline examining the science of biobanking itself, ie, biospecimen science. Indeed, in an effort to determine the best methods for procuring and providing the highest quality specimens and data for research, and ultimately for patient care, studies related to important issues in biospecimen science, such as methods for cryopreservation, and the

![A PubMed search for biobank publications (2000s)](image)

Figure 1 Steady increase in biobanking papers since 2000.
effect of freeze-thaw cycles and other preanalytical variables on specimen quality,\textsuperscript{20–23} are emerging. In parallel, biobanking is developing a variety of specialized subdisciplines related to all aspects of the field, including biobanking cryobiology, biospecimen science for biomarker validation, bioinformatics and data sharing, and specific ethical issues as they relate to specimen and data collection and dissemination in the post-genomic era. Thus, the main objective of all biobanking activity is to increase the quality of samples and data such that research results will be of optimum validity and reliability. Organizations such as the National Cancer Institute, the Organization for Economic Co-operation and Development, and the International Society for Biological and Environmental Repositories create and regularly update best practices for all aspects of biobanking (Table 3).\textsuperscript{24–29} Recently, the College of American Pathologists has instituted a voluntary accreditation program that biobanks may adhere to.\textsuperscript{30}

### How do biobanks disseminate samples and information?

The ability of a biobank network to disseminate biosamples and donor data depends on cooperation among its stakeholders. Willingness to cooperate is in turn dependent on an individual’s own motivation or the provision of incentives. For example, in a network of hospital-integrated biobanks, collectors, who are often surgeons, must be provided with incentives strong enough to convince them that making the extra effort involved in biobanking (requesting informed consent, calling the biobank when a surgical specimen is ready to be taken from the operating room to pathology, all of which take time and are not directly related to their daily routine) is worth their while.\textsuperscript{31} Otherwise, they will continue to consent patients only for their own specific projects, and the potential that might have been gained from sharing their patients’ material and data with the entire research community will have been lost. Occasionally, a breakdown of trust between collectors in different centers in the same network may occur, leading to withdrawal of participation by a collector or a group of collectors. For example, a group of thoracic surgeons collecting lung tumors may end their participation in the network if they feel threatened by a group of competing surgeon investigators who may wish to use their collected specimens. Indeed, such issues of perceived threats to “ownership” of specimens may severely damage a network and require preventive incentives. Such incentives might include reduced cost recovery fees for samples or a higher level of access to samples and data for research for more active collectors who consent more donors to the network.

To provide samples and data for the research community, well-constructed rules are required. As such, most biobank networks construct and institute sample access policies for dissemination of biosamples and associated data.\textsuperscript{32,33} Projects must be ethically approved and scientifically sound and feasible. Researchers should be proficient in the techniques they propose to use and this may merit a pilot project. Derivatives or sections can be released, rather than frozen or standard pathology formalin-fixed paraffin-embedded blocks. Applications for formalin-fixed paraffin-embedded tissue can be reviewed by the Sample and Data Access Committee and a pathology subspecialist. This harmonizes the process of obtaining tissue for research in an institution. Although all biobankers wish to see tissue samples used, that usage must not plunder the entire resource, compromising future research. Importantly, adequate residual formalin-fixed paraffin-embedded tissue must be available for molecular pathology examination to compare with tissue from a patient’s recurrent tumor. The committees deciding sample and data access take several factors into consideration when deciding which investigators will receive which biosamples, particularly for highly valued specimens. Prioritization may be on a first-come, first-served basis, but other factors, such as collaboration between investigators, funding, which may in turn help the biobank recover its operating costs, ethical approval

### Table 3 Available guidelines for biobanks and biobank networks

<table>
<thead>
<tr>
<th>Organization</th>
<th>Guideline/best practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISBER (International Society for Biological and Environmental Repositories)</td>
<td>Best Practices for Biorepositories: collection, storage, retrieval, and distribution of biological materials for research\textsuperscript{24}</td>
</tr>
<tr>
<td>NCI (National Cancer Institute)</td>
<td>National Cancer Institute Best Practices for Biospecimen Resources\textsuperscript{25}</td>
</tr>
<tr>
<td>OECD (Organization for Economic Co-operation and Development)</td>
<td>OECD Guidelines on Human Biobanks and Genetic Research Databases\textsuperscript{26,27}</td>
</tr>
<tr>
<td>OECI (Organization of European Cancer Institutes)</td>
<td>Ethical and legal recommendations. From the Biobank to the Research Biorepository, with an emphasis on building public trust and support\textsuperscript{28}</td>
</tr>
<tr>
<td>WHO (World Health Organization)</td>
<td>Guideline for obtaining informed consent for the procurement and use of human tissues, cells, and fluids in research\textsuperscript{29}</td>
</tr>
</tbody>
</table>
parameters, and previous interaction with the network, may also be important (Figure 2).

Many biobank networks provide samples internationally, eg, the HUNT Biobank and the Australasian Biobank Network. Recently, in an attempt to overcome international legal and ethical access barriers to genomic studies, the Public Population Project in Genomics and Society (P3G) consortium has formulated a generic access agreement, drawing on a variety of source agreements from large-scale population studies underway.\(^34\) In these cases, not only do legal and ethical issues surrounding the treatment of research samples in the sender and recipient countries require consideration, but logistical variables concerning shipping are also important. Although costly, rather than risk damage to or loss of precious biosamples, a biobank may opt to ship a test sample to the recipient prior to shipping actual study specimens. The payment arrangement as to who incurs the shipping cost (investigator versus biobank network) may vary between networks and between studies within a given network, depending on circumstances.

**Figure 2** Factors and elements affecting the dissemination of resources in a biobank network.

**Abbreviations:** QC, quality control; IRB, institutional review board; IT, information technology.
Return of research data and/or materials to the biobank

Dissemination of biosamples and data leads to experimental results which may then be returned to the biobank. Such raw experimental data can greatly increase the value of stored biosamples, because they provide a source of information for further investigators who may recycle those data to make new discoveries (so-called “secondary use”). Thus, investigators are often encouraged to return raw data gleaned from use of acquired samples. For many biobanks and biobank networks, however, enforcing the return of raw data can be difficult.

Thus, as part of a material transfer agreement, many biobank networks include a clause which requires investigators to return raw experimental data to the biobank, and/or any information related to publications resulting from biosample use. Material transfer agreements may also include specific provisions regarding specific studies. For example, if a study is likely to use the last aliquot of a particular tissue sample, the biobank may require that a sample of DNA extracted from the aliquot be remitted to the biobank by the recipient investigator.

Sharing between investigators within and between biobank networks

For individual investigators, one of the short-term goals is production of high quality manuscripts. This may be achieved through generation of data derived from successful experiments conducted on biosamples supplied by the biobank (primary use of samples and data). However, it may also be achieved through studies focusing on secondary data analysis of returned research results. For effective secondary use of data to take place, primary investigators must abide by the material transfer agreement, and the data provided must be of high quality. A network’s protocol for data sharing should be carefully designed and implemented, and should be transparent, both to the public and to the research community. Such protocols may evolve with changing technologies. A recently published example is that of the data sharing policy and protocol of the National Heart Lung and Blood Institute which manages individual level data from over 560,000 study participants from over 100 trials. The policy outlines conditions regarding timing of data set release, data submission, and data access requirements.35

Risk consideration in data dissemination

In 2001, amidst increasing concerns over privacy and confidentiality in medicine, the oversight committee of the Ethical Force Program compiled a consensus series of 34 measurable expectations for the protection of privacy and confidentiality in health care in the USA. Concerns included release of identifiable health information without patient consent, use of health information by employers, large-scale breaches of confidentiality in electronic systems, and inconsistency in existing health information privacy norms. The expectations for use of identifiable health information outlined over a decade ago fell into content areas which all biobank networks deal with today on a daily basis, ie, transparency, consent, collection limitation, security, individual access, data quality, information use limitation (or waiver of consent for uses beyond the consented limitation), and accountability.36

The large-scale use of biosamples and data, and the potential secondary uses of data long after the primary experiment has been carried out, has fuelled discussion of one of the most important issues in biobanking governance. Biobank networks relying on strong information technology platforms must take into account the delicate balance between the scientific value of the data provided and submitted on the one hand, and donor protection on the other. Data sets provided for secondary use must be redacted to a degree sufficient to reduce the risk of reidentification, but not so much as to reduce scientific utility.

In the USA, a consortium of stakeholders headed by the American Medical Informatics Association, including health care companies, academia, public health organizations, and private sector research organizations, initiated a framework for the exchange of electronically formatted health-related data.37 This framework was then reviewed and refined by American and European experts,38 and recommended by the US National Committee for Vital and Health Statistics.39 Following this lead, a panel of experts in the European Union convened the European Summit on Trustworthy Re-use of Health Data, led by the International Medical Informatics Association (Brussels, 2012). At this summit, the experts were presented with three scenarios in which large-scale data sharing was proposed: integration of multiple databases; collection and reselling of patient electronic health record by a vendor; and public health surveillance on a global level. Results of discussion revealed a high level of consensus on the public health benefit of data sharing in scenarios pertaining to integration of databases and global public health surveillance, on the need to fully inform participant donors, and on the need for tight government regulation. Interestingly, however, no consensus was reached regarding individual benefit of data sharing in any of the three scenarios. Trust,
however, was seen as an integral part of any proposed data sharing framework, requiring clarification of the benefits of data sharing, and issues of privacy and data protection being taken into account.40

Many large-scale biobank networks use the open-access translational research information systems model for their information technology infrastructure to establish a data warehousing infrastructure and facilitate rapid dissemination of research findings.41 These systems integrate a wide variety of data types. While the use of open-access translational research information systems may be contingent upon participant anonymity, the security measures taken may not be enough to protect patient privacy. Thus, data warehouses must be subject to tight regulatory controls, such as the Health Insurance Portability and Accountability Act (HIPAA) and its accompanying privacy rule or the National Institutes of Health (NIH) deidentification policy for genome-wide association studies.42–44

Deidentified data can be reidentified if they are unique or “distinguishing” (publicly available deidentified hospital discharge records, now unavailable due to the HIPAA and Health Insurance Portability and Accountability Act safe harbor policy), a naming resource exists (such as voter registration lists), and there is a mechanism for linking the deidentified data with the naming resource. It has been argued that reidentification is relatively simple for anyone with a computer, data, and an understanding of database software.45 For all biobank networks using open-access translational research information systems, person-specific features must be classified into high-risk identifiers (eg, demographic data) and low-risk identifiers (diagnosis-related and treatment-related data). In the biobanking world, as in other fields of health care, participant reidentification may be possible through genotype-phenotype correlation, familial information, trails and location-based patterns, genome sequence data (even when aggregated), and laboratory reports and expression (functional genomics) data. However, the likelihood of an “attack” is probably quite low. Nonetheless, given that complete privacy protection cannot be guaranteed, data managers should be able to estimate the risk of reidentification,46 and all biobank networks should adopt a list of technical and policy-based mechanisms such as those suggested by Malin et al to help improve data privacy protections. These include k-anonymity, assessing replicability of molecular data types, establishing formal access policies, implementing data use agreements and transparent informed consent procedures which specifically address future use of data, putting in place procedures for redress in the unlikely event of a data security breach, audits, and varying levels of access for personnel.41,47

Although there is a public fear of individual reidentification through deidentified research data sets, it should be borne in mind that an attacker wishing to identify an individual still requires an identified DNA sample. Research data sets from biobanks may be a potential source of data for an attacker; however, the question has been raised as to why, other than to prove that it is possible, an attacker would use a sample to determine whether or not an individual’s DNA was in a research dataset. Other than “proof of concept” attacks, which are reported in the media and in the literature, the realistic risk of an attack is believed to be very low, particularly when risk mitigation steps are taken.48 Parenthetically, while not addressed in all informed consent forms for biobanking, biobankers and biobank donors, and indeed anyone who has provided biological material in the clinical setting, should be aware that privacy protection is not absolute. In Italy, for example, clinical and/or research samples provided through informed consent may be used (without reconsent) for identification in the course of a criminal investigation.49 In the USA, biological material, banked or otherwise, may be subpoenaed in the course of a criminal investigation. Clinical material cannot be protected from subpoena. Whether or not material biobanked for research is protected by a confidentiality certificate will depend on the informed consent practice at the institution in which the material was banked.50

The public, understanding biobank networks, biosamples, and data-sharing

Patient care

Biobanking is patient-friendly. Among the many stakeholders in networked biobanking, the patient/donor is central. Specimens donated to the network are taken for biobanking on the understanding that their removal does not compromise pathological evaluation or future medical care. For example, when a specimen is brought from the operating suite to the pathology department, a pathologist must decide whether there is enough residual tissue for banking, and must oversee sampling of that tissue and determine its quality (eg, percentage of viable cells, percentage of tumor cells, percentage of necrosis) for future use. This is particularly relevant for hospital-integrated cancer biobanking, where the pathologist’s role is central in safeguarding the immediate, and sometimes long-term, interests of patients.51 On occasion, a banked specimen is required subsequently for decisions regarding patient care. In cancer biobanking, patients who develop a recurrence may require molecular comparison of the biopsied
reccurrent tumor with the originally donated specimen in order to modify treatment. Part of the original specimen may be in frozen aliquots in a research laboratory or in the biobank. If so, the tissue is provided to pathology for additional diagnostic molecular investigations. Alternatively, the patient’s medical oncologist may request part of the biobanked tissue for pretreatment tissue testing. Such a once-off request might be for the purpose of testing the effectiveness of an approved targeted treatment or an experimental new drug. It is conceivable that the original donated sample might be needed in its entirety. In this situation, it is likely that the patient/donor has advanced disease, and every possibility is being explored. It must be remembered that the patient is more important than the biobank network. Biobankers are neutral custodians of donated biosamples. Donors are free to change their minds about research use of their specimens and data, and patient care has priority over research. If the biobanked tissue leads to a successful therapeutic outcome, that is a win-win situation for everyone, not just the patient.

It is highly desirable that patient groups and the public in general understand the role of biobanks and the scientific and societal benefits of the biobank network. The biobank network, which has multiple stakeholders, is in a good position to coordinate information provision, which could feature a pharma partner and the media. People should be given opportunities to understand that research on cancer tissue will lead to better and more specific treatments for cancer and other serious diseases. They should see that taxpayers’ money is being spent to good advantage. Understanding broad concepts, they will be better equipped to interpret research results presented in the media. Eventually, it may happen that consent for biobanking blood or tissue, from which there is a “soft opt out”, is considered standard health care, and authorized on entry to hospital or other health care settings. This system applies in countries such as Belgium and the Netherlands where a code of conduct stipulates that secondary use of human residual biosamples for research is allowed unless the patient has objected to this use. An alternative version using deidentified specimens, and relying on advanced privacy protection measures to the extent that such biobanking is not considered human subjects research, has also been recently introduced in the USA.

**Return of results**

What about research findings that patient donors want to be made aware of? Although protection of participant donor privacy is critical to biobank networks, public expectations with respect to privacy may be changing. The use of social media in the health care sphere is increasing, particularly with the advent of the Internet-savvy e-patient. While direct interaction with a patient for the purpose of patient care is discouraged, under strict guidelines, social media may be used for other activities, such as clinical trial recruitment. A cursory Google search reveals that many biobanks worldwide have Facebook pages, and some receive hundreds of “likes”. However, it is unclear as to whether or not biobanks successfully recruit donors through social media. It is conceivable that biobank advertising through social media may potentially lead to donor recruitment, enhance public awareness and debate, and may also lead to increased financial support for the biobank. Literature is scarce, but one study has shown that biobank advertising through social media has the potential to reach the public and to result in positive interactions (“likes”), at low cost.

Many Internet-literate patients are actively taking part in research activities whereby they directly place their own health care data on websites such as “Patients Like Me”, which has recently launched an open research exchange. Similarly, many biobanks and biobank networks worldwide not only serve the research community, but also have interactive websites where participant donors may receive information regarding the biobank, fill in health-related questionnaires, and see results of studies using banked samples and data. In general, however, individual research results are not reported back to the participant, a phenomenon which has become an issue of active debate among biobankers, since the boundaries between biobank research and health care are becoming increasingly blurred.

With the advent of whole genome sequencing and whole exome sequencing, the ability to generate large quantities of data which may or may not be clinically actionable, is fuelling the debate on sharing data (particularly of unintended research findings unrelated to the research or clinical question) with the individual donor. There is evidence that research participants wish to receive genetic information regarding disease-causing mutations identified incidentally for conditions other than that under direct study. In a recent survey of researchers’ views on return of incidental genomic research, the overwhelming majority believed that highly penetrant, clinically actionable findings should be returned. However, it was recognized that returning incidental findings may be burdensome to the research community, as many researchers lack expertise in identifying findings to be returned, and there is no generally accepted policy on return of incidental findings.
As pertains to the clinical setting, in March, 2013, the American College of Medical Genetics and Genomics issued recommendations for reporting incidental findings in whole genome sequencing or whole exome sequencing. The list includes 56 genes in 24 inherited conditions which should be sought and reported in addition to the analytic process, regardless of the reason the test was ordered, and without regard for patient preference as to whether or not they wish to receive the added information. The recommendations were called a “valuable starting point for deliberation” of issues surrounding optimizing and standardizing genomic testing in the clinical setting. However, since the 56 listed genes are recommended as add-ons to the test ordered clinically, their results do not truly constitute incidental findings. Ethical concerns over patient autonomy and right to know have been raised, as have concerns regarding return of results of testing in children for the possible purpose of parental benefit.

Following a year of discussion, the American College of Medical Genetics and Genomics recently changed the recommendation to include an opt out clause for patients to choose prior to testing, in order to avoid generation of results they would not wish to receive. With the debate unresolved in the clinical setting, one must approach the issue of sharing individual research data, particularly incidental genomic research results, with biobank donors, with extra caution. Participant donors who consent to dissemination of their samples and deidentified data for future research endeavors are generally not consented to a specific project, but rather to a variety of future investigative options, which may include genetic studies. A donor’s specimen may yield one genetically abnormal result at one point in time, and another genetically abnormal result at another point in time. Is it logical and/or ethical to recontact a donor repeatedly? While some biobanks, such as the Estonian biobank, allow access to individual results, the traditional attitude toward return of individual results is that it may be difficult to administer; the cost of validation of research results may be prohibitive; the donor may have died in the time elapsed between specimen donation and study completion; and the result obtained, even if valid, may not be actionable. Thus, many biobanks opt to disseminate results in a collective form, by listing published studies resulting from the use of banked samples and data in a public forum such as a website.

**Special considerations in ethnic and cultural diversity**

Little is known regarding the attitudes of minority groups towards biobanking and issues surrounding biobanking, such as return of research results. “Deliberative engagement”, in which participants are informed and discuss policy questions, was used to examine African Americans’ views on biobanking and return of results. In this study, attitudes and beliefs were surveyed before and after engagement using four educational slide presentations followed by discussions. Results for the clientele of a university-based practice were compared with those for the clientele of a federally qualified health center serving underserved populations, and showed positive attitudes for both groups toward biobanking, and general support for broad consent. However, participants in both groups expressed a need for education and understanding of biobanks. Similar attitudes were expressed, from an appreciation for respect shown them by the recruiter, a desire to receive future research results, and concerns regarding confidentiality breaches, including the possibility that their biological material may be handed over to law enforcement authorities. Given that trust is an important factor in research participation, knowledge of and respect for participants in special ethnic and cultural groups is paramount in recruiting. As such, in the underserved population, a preference was expressed for an African American recruiter. Similarly, in one large teaching hospital in Jerusalem, Israel, the recruiting process is carried out by orthodox recruiters, and takes into account the extra time required for rabbinical consultation for orthodox Jewish potential donors.

**Evaluating the impact of biobanks and biobank networks**

Biobank networks are in an ideal position to encourage individual participants to become “citizen scientists”. By allowing future investigators to use their data, particularly longitudinal follow-up clinical data, as well as samples, donor contributions to scientific advancement are of infinite value.

Adequate supply and effective implementation of sample and data sharing policies by a biobank network benefit the scientific research community, and may also benefit the biobank network itself, as it strives to build a reputation as a reliable source of biosamples and data. The work around providing biological samples with or without associated data is time-consuming and may be quite labor-intensive. Since many networks function in an academic setting, the question as to how to credit biobank stakeholders for their efforts often arises, with no definitive answers. Major health research funders seeking to promote bioresource sharing have not provided adequate incentives or tools to achieve the goal of maximizing the translational research potential in
The authors report no conflicts of interest in this work.

References


