Nature Reviews Cancer | AOP, published online 12 August 2010; doi:10.1038/nrc2913

PERSPECTIVES

SCIENCE AND SOCIETY

Integrating biobanks: addressing the practical and ethical issues to deliver a valuable tool for cancer research

R. William G. Watson, Elaine W. Kay and David Smith

Abstract | Cancer is caused by complex interactions between genes, environment and lifestyles. Biobanks of well-annotated human tissues are an important resource for studying the underlying mechanisms of cancer. Although such biobanks exist, their integration to form larger biobanks is now required to provide the diversity of samples that are needed to study the complexity and heterogeneity of cancer. Clear guidelines and policies are also required to address the challenges of integrating individual institutional or national biobanks and build public trust. This Science and Society article highlights some of the main practical and ethical issues that are undergoing discussion in the integration of tissue biobanks for cancer.

It is widely accepted that although basic scientific studies carried out using cell lines and animal models can be informative about the cellular and molecular aspects of cancer there is a clear requirement to confirm this in human samples. The concept of patient-specific and disease-specific ('targeted') therapy has expanded rapidly in recent years. Many researchers believe that this concept of personalized medicine will provide the solution to the considerable challenges posed to the clinical treatment of cancer. The move from the traditional 'one size fits all' approach for the treatment of cancer to targeted approaches seems to offer genuine hope for improved patient outcomes. There are a few examples for which the concept of a highly effective drug treatment targeted towards a specific limited patient population has become reality. The use of imatinib (Gleevec; Norvartis) in chronic myeloid leukaemia[1], the use of monoclonal antibodies that target the epidermal growth factor receptor (EGFR) in patients with EGFR-expressing metastatic colon cancer[2] and the use of the ERBB2 (also known as HER2)-specific monoclonal antibody trastuzumab (Herceptin; Genentech) in ERBB2-positive breast cancers[3] are such examples.

The potential benefits of this personalized approach to the treatment of disease are considerable. They include the identification of improved biological targets using validated biomarker studies, the capacity to increase the likely success of clinical trials by preselecting the patient population and the fact that this will in turn reduce the time, cost and the likelihood of failure of clinical trials[4]. Information from validated biomarker studies also allows the re-introduction of drugs that have failed in a clinical trial setting or that have been withdrawn from the market to be re-applied in a more targeted way. Similarly, biomarker studies might also offer the potential to avoid adverse side effects, and this would, in turn, lead to higher compliance with various treatment regimes.

The need for large integrated biobanks

One of the biggest limiting factors to the successful translation of basic scientific cellular and molecular studies into improved patient outcome has been the lack of access to large, appropriate and well-annotated cohorts of human tissue[5,6]. Focused disease-specific institutional biobanks have had some success in translational and personalized medicine (as described above). However, owing to the complex and heterogeneous nature of cancer, it is now clear that much larger biobanks are required[7,8]. For genetic main-effect studies 2,000–5,000 samples are needed, for lifestyle main-effect studies 2,000–20,000 samples are required and for gene–lifestyle interaction studies 20,000–50,000 samples are required[8]. Only when these larger resources are available can we truly understand the interactions between gene, environment, lifestyle and disease and translate this knowledge into the clinic through innovative diagnostics, therapeutics and preventive strategies for cancer. These larger resources can only be achieved by the integration of existing biobanks that already have a wealth of information and samples. However, there are many obstacles and challenges associated with such integration, including technical, logistical, ethical and legal ones.

Groups across both North America and Europe have started to address these obstacles and challenges to move this process forwards. Initially, national programmes were established that linked previously collected biobank samples. These included the Canadian Tumour Repository Network (CTRNet; see the CTRNet website; Further information), which was established to link cancer researchers with provincial tumour banks, thus creating new opportunities for translational cancer research to improve cancer outcomes in Canada and beyond. This network gave researchers access to tissue and clinical data. The Organisation of European Cancer Institutes (OECI) TuBaFrost database (see the TuBaFrost website; Further information) was established in 2003 to network European frozen tissue pathology banks for cancer research. In this initiative, the tissue collection process was standardized, a code of conduct for the exchange of residual human material for research was developed based on European legislation and a web-based sample request process was developed. Other examples include the Spanish tumour bank, EuroBoNet (see the EuroBoNet website; Further information) and the Office of Biorepositories and Biospecimen Research (OBBR; see the OBBR website; Further information) in the USA. These initial biobanks led to further expansions and networking to create resources across Europe and the United States. The establishment of the

European Biobanking and Biomolecular Resources Research Infrastructure (BBMRI) programme (see the BBMRI website; Further information) illustrates moves to coordinate existing biobanking activities across Europe[9,10] (BOX 1). The development of the Cancer Bioinformatics Grid (caBIG®) infrastructure (see the Cancer Bioinformatics Grid website; Further information) is connecting research organizations across the United States[11] (BOX 2). The National Comprehensive Cancer Network (NCCN; see the NCCN website; Further information) has collected patient data outcomes for breast cancer, non-small-cell lung cancer, colorectal cancer, non-Hodgkin's lymphoma and ovarian cancer. The data, along with records of patient treatments and patient outcomes, have allowed retrospective comparative studies to be conducted. The analysis of subpopulations in these databases has already led to changes in clinical practice[12,13].

There are many reports outlining the complexity of biobanking that provide strong recommendations and the identification of best practice for all aspects of the process[14] (see the National Cancer Institute Best Practices for Biospecimen Resources website; Further information). FIGURE 1 outlines the steps involved in biobanks. As the requirements for complex multi-institutional and international collections to study cancer processes have been established, this article focuses on some of the important practical and ethical issues related to the integration of biobanks.

Box 1 | **The Biobanking and Biomolecular Resources Research Infrastructure**

The European project Biobanking and Biomolecular Resources Research Infrastructure (BBMRI) was established in 2008 to network European biobanks with the aim of improving resources for biomedical research and therefore contributing to the improved prevention, diagnosis and treatment of disease. The resource includes 261 biobanks across 23 countries with a total of more than 16 million samples. It is only possible to achieve this using a federated network of centres in European countries, which is best described as a distributed hub structure of existing biobanks. This will provide the flexibility to facilitate expansion and multiple uses (see the BBMRI website; Further information).

The mission of the BBMRI is:

- To benefit European health care, medical research and, ultimately, the health of the citizens of the European Union
- To have a sustainable legal and financial conceptual framework for a pan-European Biobank infrastructure
- To increase scientific excellence and efficacy of European research in the life sciences, especially in biomedical research
- To expand and secure the competitiveness of European research and industry in a global context, especially in the fields of medicine and biology

Practical implementations to integration

Definition and use of biobanks. There are many types (BOX 3) and definitions of biobanks, which are informed by the goals, objectives or statements of purpose, and so differ from institution to institution. Defined in the most general terms, a biobank is a collection of biological specimens and corresponding participant data. The goals of the biobank will therefore determine the types of material collected, the scope of the research and the consent used to collect the samples. A clear definition of the scope of a biobank is essential to define informed consent for its participants, determine standard operating procedures for sample collection and storage, define the data management system and determine the use of the samples. All of these characteristics are not only required for a high quality resource but also for building public trust, avoiding the misuse of the resource and setting clear expectations for the contributors to and users of the biobank.

However, providing a clear definition of the scope of a biobank is a problem when integrating biobanks that have been collected across different institutions or countries for different reasons. There is a clear need for such flexibility, owing to the ever-expanding development of technology and understanding of the complexity of cancer, but this flexibility has effects on patient consent, standard operating procedures, information technology management systems, and the use and distribution of samples.

Standard operating procedures. The careful and well-documented processing and annotation of samples is essential to provide a useful resource for scientific interrogation. Valid data are defined as those that are reproducible, but variation is an enemy to such reproducibility. The importance of standard operating procedures has been well established to prevent variation in the patient consent, collection, processing, storage and distribution of samples. Several important documents have been produced that clearly define the best practices for biobanks[14,15] (see the National Cancer Institute Best Practices for Biospecimen Resources website; Further information). However, balancing the standardization and flexibility of biobanks is of considerable importance. The main challenge related to biobank integration is the consistency of standard operating procedures across multiple sites, so that differences observed across patient populations are not affected by where and how the samples were processed. There are ongoing international collaborations addressing these issues; for example, the Data Schema and Harmonisation Platform for Epidemiological Research (DataSHaPER; see the DataSHaPER website; Further information) has been developed through the Public Population Project in Genomics (P3G) and Promoting Harmonisation of Epidemiological Biobanks in Europe (PHOEBE) and involves 25 international biobanks. These international collaborations have defined several core high priorities for improved study integration, such as sample collection, preliminary processing and storage, as well as the importance of documenting key aspects of the standard operating procedure, such as the time of collection and processing so that this variable is recorded and linked with the sample. These networks have all identified the need for quality assurance systems that operate across the sites and that can address compliance with, and the implementation of, standard operating procedures, but their implementation has not yet been agreed.

Information technology management systems. A central component of any biobank is the bioinformatics and data management system. It is clear that to harness the full potential of any biological sample, information needs to be collected not only about the standard operating procedures and pathological status, but also the demographic, diagnostic, medical and family history and clinical outcome data of the participant. For such data, the patient must be identifiable to a member of the biobank so that the information can be updated over time. Issues associated with such information extend beyond the storage of data and linking it to a specific participant and also include data reporting, data searching and mining, data accessibility and network

security, and require personnel to enter, manage and maintain the data in the future. There are a large number of commercially and institutionally available data management systems but there is no common consensus on which system delivers an overall solution. Biobank information management systems (BIMS) are emerging as a proposed solution that allow the storing, tracking and most importantly the searching of several sources of data at the same time. The Karolinska Institute has developed such a system, and it has been proposed that this BIMS model could be used to manage the data from the planned LifeGene project, which will include 500,000 Swedish participants and will follow them for decades using both questionnaires and testing of individuals. Another such project is the UK Biobank, which will gather samples from 500,000 participants with extensive medical and family histories and follow them for 30 years. This project is expected to collect 15 million aliquots of blood and DNA. However, it is clear that a strong and close interaction between the information technology development team, data managers, biobank managers, clinicians and researchers is required to build a system that is fit for purpose and that also contains the appropriate security features to protect participant confidentiality and so foster public trust. Individual biobanks need to agree to use standard identification schemata, data formats, data quality assurance and control processes, database architectures and common security processes when building their systems.

The integration of databases is one of the most challenging issues for establishing biobank networks. Ideally, data should be exchangeable across multiple national and international sites using a web-based system, and facilitated by agreed consent, ontologies, naming conventions, data formats and definitions that are underpinned by common standard operating procedures for data collection, uploading, management and audit. Unfortunately, there is no such agreement; however, the World Wide Biobank Summit has indicated that biobanks should agree on minimum data sets that are interchangeable between biobanks and identify the complete ontology and multi-lingual definitions of the data set. In addition, access policies and levels of security to protect the identity and confidentiality of patients must be agreed. This can be achieved using network security and access control that limit different users to different aspects of the data. For example, data managers would be able to see patient-identifiable information that is applicable to their institution only, so that they can update the appropriate information, and researchers or external collaborators would see the de-identified codes of participants as well as the relevant clinical and scientific data that are applicable only to their level of access and requirements. This approach would help to address issues of confidentiality and reassure participants of privacy, which is central to building public trust. However, it is clear that an important requirement for these information technology systems is that they are built in a flexible way to accommodate the ever-evolving requirements of participants and users of biobanks.

Ethical issues associated with integration

Patient consent. Informed consent is a fundamental and key mechanism put in place to protect the interests, welfare and rights of research participants and it is generally accepted as an absolute imperative for the collection of human samples[16,17]. Central to the consent process is information that educates the participant about the research and how samples will be collected and used, along with details of potential benefits and risks. This information allows participants to make an informed decision regarding their involvement in a study. A clearly defined biobank makes it easier to inform participants about the process and the use of their samples; however, with increasing pressure for flexibility in the use and sharing of samples, the consent process has become more complicated. In addition, biobanks are typically established for long-term use, therefore the details of future research projects cannot be known. Obtaining new informed consent for each new project is in most cases difficult, if not practically impossible. However, it has been argued that general consent to future research projects is not sufficient. To what extent can such consent about the unknown future be accurately described as informed consent?[18–20] One proposed solution to this problem is to accept general consent for future research if the participants have the option to withdraw their data and if future research projects are approved by an Institution Review Board (IRB) or Research Ethics Committee (REC). It can be argued that adopting a somewhat lower standard than that of classic informed consent could be justified because the risks to individuals resulting from research on their human biological material do not involve physical harm, as is the case with clinical trials[21] (BOX 4). In addition, it has been proposed that owing to the benefits that are associated with the knowledge generated through these potential future studies, and the fact that this information could accrue to all individuals

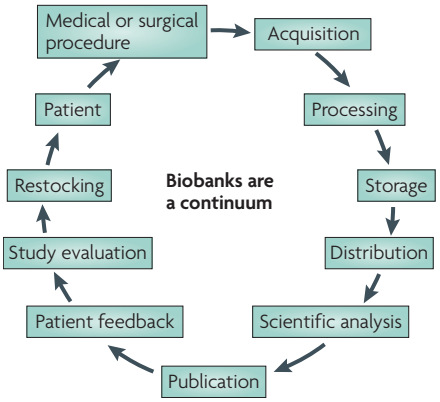

Figure 1 | **The steps involved in establishing a biobank.** The steps involved in biobanks are a continuum, starting with participant consent, leading to the collection, processing, storage and distribution of samples. The process finishes with the publishing of data and informing participants of how their samples have been used, which then supports the continuation of recruitment to biobanks.

Box 2 | **The Cancer Bioinformatics Grid (caBIG®)**

The Cancer Bioinformatics Grid (caBIG®) was launched in 2004 by the US National Cancer Institute and is already enabling many research organizations to move their clinical and basic research programmes forwards. The caBIG® programme aims to create a virtual web of interconnected data, individuals and organizations, with the ultimate goal of improving the interaction of those involved in treatment-focused research, leading to improved patient outcome. It is essentially an infrastructure for connecting research organizations that allows for a systems approach to biomedical research. It relies on the widespread use of sophisticated and powerful information technology, which offers improved data management and data exchange and sharing. This will facilitate a move from the traditional self-contained or isolated model of research organizations to the concept of large networks that allow faster, larger and standardized collaborative research. The CaTissue suite developed by caBIG® is a biobanking managing system designed to collect, manage, process, annotate, request and distribute biospecimens and associated information (see the Cancer Bioinformatics Grid website; Further information).

as well as future generations, it may be possible to justify an expansive use of participants' samples for future studies[22].

In recent research conducted in the United States on stored biological materials, it was found that biobanks used consent forms that offered individuals the option of designating how their biological materials could be used with varying degrees, including whether the materials could be stored; whether participants could be contacted again in the future; whether the materials could be used for genetic testing; and whether the materials could be shared with researchers who were not part of the original research team. In addition, a few biobanks offered options for participants to designate how their biological materials could be used in future research — for example, by specifying whether identifiers could be retained, or specifying the research topics for which the materials could be used[23].

In support of Elger and Caplan's[17] suggestion that general consent should be the international standard, the European Council has taken the view that broad consent to future research use is acceptable[24]. This implies that individual donors renounce any ongoing rights to exercise control over the uses of their donated materials and the source itself[25]. This point has been emphasized by Shickle[26] who states that "providing that there is proper disclosure and so on, then the choice for the individual is to participate on the terms offered or not. There is a 'negative right' not to be included in the research without consent. There is no 'positive right' for a biobank to be run in such a way just because an individual would like it to be so."

Investigator access. The scientific potential of biobank samples and data can often be fully exploited only if their use is not confined to individual research projects specified in advance. Donors should be able to give generalized consent to the use of their samples and data for the purposes of medical, including genetic, research. For this reason, donors should also be able to consent to the transfer of samples and data from biobanks to third parties for the purposes of medical research. Most biobanks provide samples and data to various researchers who are often not directly affiliated with the biobank holders. The transfer of samples and sharing of data raises several ethical questions. If samples and data are physically transferred outside the biobank, the recipient might use the material for purposes other than those to which the sample donors have consented or the samples might be transferred to third parties, thus putting the privacy of donors at risk. If recipients use samples for their own purposes, the investments and efforts of the original biobank holder could be threatened. Biobank material is a valuable resource and the ethical issues of allocating such material emerge if the samples and data are misused. However, if the transfer of samples is too restrictive because of transfer agreements that include complex and costly control mechanisms, the resulting inaccessibility of biobank material does not allow a large number of researchers to use this resource in the interests of science and for the good of the public[21].

However, except in circumstances prescribed by law, it has been strongly argued that the transfer of samples or information must take place in a simple coded or double-coded way, with the recipient of double-coded samples having no access to the code. Should the recipients' research require an association with personalized data, this may be provided only by an official of the biobank to which the donors originally entrusted their samples and data. The rationale for this is that an alternative anonymization strategy, which permanently destroys the information that links a sample and donor identity, risks the loss of information of scientific value, as there is no way to prevent multiple inclusions of the same participant, and this approach does not allow retroactive validation and demonstration of reproducibility. To facilitate clear autonomy, transfers of samples and data to third parties must be fully documented for future reference according to standard operating procedures[27].

Subject to the consent of donors, the transfer of samples should be permissible provided that the recipient is subject to standards of donor protection and quality assurance equivalent to those that are applicable to the original institution in charge of the biobank. Transfers of existing biobanks to third parties with the inclusion of personalized donor data should be possible only with the approval of a REC or IRB[28].

Patient access. An individual's right to access their genetic data is linked to control over their own identity. This varies across biobanks, as do the ethical justifications for the differences. In Estonia, personal genetic data is available to participating individuals on request or the participants can decide that the information should be withheld from them. Likewise, in Iceland, patients are entitled to information about their health (for example, condition, prognosis, risks and benefits of treatment) on request. The UK Biobank takes a different and less open approach. Although the Biobank Protocol acknowledges that individuals will have the legal right to access their personal data if required, it also makes it clear that they will not routinely receive any individual information that relates to their blood samples (including biochemistry and genetic findings). The assumption that participants should only be provided with health information about themselves in a clinical situation in which a trained professional can provide appropriate interpretation and guidance lies behind this prohibition. The justification for this position is that it would not be constructive and might even be harmful to provide health information but no interpretation, counselling and support. The UK Biobank cannot provide such counselling and support because it is purely a research initiative[29]. The German Ethics Council recommends that if individual communication of research results to the donor is agreed, then they must also be told as part of the information to be given that they must divulge these details in certain circumstances — for example, when

Box 3 | **Types of cancer biobanks**

A population-based biobank is defined as a large repository of donated human DNA and/or its information, collected from volunteers with and without cancer, which is used to identify the genes that contribute to human disease. This is an essential resource if we are to understand the genetic risk factors that are associated with cancer development and the genetic profile of the patient that is associated with the development of cancer. These resources also need to include high-quality lifestyle data.

Disease-based biobanks are defined as a collection of biological material from patients with cancer, and are essential to understand the molecular and cellular development of cancer at a specific stage.

These types of biobanks can be collected for clinical purposes; for example, they could be established for medical purposes, such as a blood bank. Specimens can also be collected as a by-product of diagnostic or curative procedures; for example, from hospital pathology departments. Biobanks can also be collected for specific research investigations; for example, clinical trials.

Box 4 | **Donor information required for consent**

In light of ethical controversies, informing participants about the details of storage and research involving their samples is crucial. It is widely agreed that sample donors must be informed about:

- The voluntary nature of participation
- The type of consent used (informed consent for the primary clinical or research purpose, or a general consent for future research)
- The circumstances of sampling, including the risks and benefits of the procedure
- The aims of storage (clinical purpose or research purpose) and the nature, extent and duration of the proposed use, including the possibility of genetic analysis
- The extent of and conditions for the possible future transfer of samples and data
- The measures taken to protect confidentiality (who will have access to the samples and information and the risks for individual donors and groups)
- The form of data storage and combination
- The anonymization or pseudonymization of samples and data and other ancillary donor protection measures and any provision for state access to samples and data
- The right to withdraw consent
- The limits of withdrawal and the fate of samples and data in this instance and if the biobank closes down
- Whether individual or aggregated research results will be disclosed to sample donors
- The possible consequences of the communication of results of genetic analyses for the donor and their relatives, including possible obligations to divulge (for example, to insurance institutions)
- The use of biobank samples for commercial purposes (patients planned, benefit sharing planned, and if so, in which form)
- Issues of payment of expenses, remuneration or benefit sharing

concluding new employment or insurance contracts in the future. In addition, when such individual communication to the donor has been agreed, the findings must be imparted by a person with the appropriate counselling skills, especially when communicating the results of genetic diagnosis[28]. However, most biobanks take a different position regarding the right of donors' access to data. Participants are kept informed about research that is currently in progress or completed. These findings are usually made available through regular information sheets, newsletters or a dedicated website. It is clear that continuing public engagement and the feedback of findings is necessary for an acceptable public profile of biobanks[30]. It is recommended that such an approach should be taken following the integration of biobanks, as the ability to identify individual samples shared across resources would not be possible owing to anonyomization. However, if participants are informed at the time of donating samples of their ability to access research data, then new mechanisms need to be identified to relay information from future studies to participants from the collaborating users of the biobank.

Withdrawing samples or data from biobanks. The right of participants to withdraw their personal data reflects the basic principles regulating medical research according to the Nuremberg Code and the Declaration of Helsinki. However, exercising this right conflicts with the interest of science and industry to maintain the statistical integrity of population-based databases, particularly given the aims of using these databases for longitudinal studies. There are two important problems to withdrawing: namely, what can be withdrawn and when it can be withdrawn[31]. Although all studies discuss the right to withdraw from research generally, almost one-third failed to discuss this right specifically with respect to research involving biological materials. Therefore, specific discussion of the right to withdraw biological materials may be necessary to preserve this basic right regarding participation in research[23].

The UK Biobank offers a set of graded options for withdrawal (complete withdrawal, discontinued participation and no further contact requested), which attempts to balance the interests of the participant for data removal with the interests of the scientific community. In Estonia, there seems little room for negotiation, as donors have the right to have their data deleted from the database on request, and any violation of this right, including coercion to participate, is punishable as a criminal offence[29]. The German Ethics Council argues that donors must have the right to withdraw their consent to the use of their samples and data at any time. It should not be possible to waive this right. However, there should be a provision for donors to allow samples and data to continue to be used in the case of withdrawal if they are anonymized — that is, if the link to the participant's identity has been eliminated[28]. This is particularly important if the biobank has been integrated into a larger network and samples are shared across that network.

Participants need to be made aware of the fact that it will not be possible for samples and data that are included in completed research to be extricated from the research results and destroyed. The right to withdraw from a biobank does not include the right to withdraw research results that have already been accumulated. Rather, it means a prohibition of obtaining new data and analysis from the samples. Even if biobanks decide to retain existing data in an impersonalized form, questions arise as to whether complete and irreversible anonymization can be considered as a solution to the problem. When can withdrawal take place? One obvious point in time when withdrawal will be impossible is when samples and data have been analysed and are part of a publicised research result. Also, at this point, the participant's data will have been merged with the other data of that particular study and integrated into the result to create new data sets. However, participants could potentially withdraw at any point before this occurs[31].

The comparative analysis of practices in Europe and the United States has shown that, although the exact details may be different, all biobanks examined deem the provision of a withdrawal option to be important. The fact that informed consent is requested and that information is provided about the mechanisms of withdrawal also plays a decisive part in ensuring that participant autonomy is maintained[23]. However, the withdrawal of samples that have been shared with network partners has not yet been resolved and so is a challenge to the integration of biobanks.

The roles of IRBs and RECs. As can be observed, the role of IRBs and RECs seems to be crucial for the management of biobanks. Therefore, before conducting a research project involving the use of biobank samples and data, the consent of an IRB or REC should be necessary for studies in which bodily substances are to be collected from a donor's body for research purposes; the project calls for identified samples (that are linked to the individual in a way that makes them immediately identifiable) to be transferred to external researchers, and

entire existing biobanks are to be transferred to third parties with identified samples included.

The involvement of an IRB or REC, and the need for its favourable opinion, is intended to ensure that a narrowly worded consent is not exceeded, that a consent in broad terms is not inappropriately given an even wider interpretation and that exceptional situations in which consent may be waived are not illegitimately invoked[15]. To ensure the integration of biobanks, IRBs and RECs need to be proactive in cooperating with one another. To some extent, this process is already underway. The European Forum for Good Clinical Practice (EFGCP) (see the European Forum for Good Clinical Practice website; Further information) is an example of the development of common policies and standardized procedures for European RECs. Such initiatives need to be encouraged and supported for biobanking integration so that IRBs and RECs can work from a common set of standards, procedures and documentation that will streamline ethical applications without compromising the underlying ethics. IRBs and RECs are becoming the guardians of biobanks as they move towards more integration.

Perspectives

Biobanks have the potential to be resources that will facilitate an understanding of the interactions between genes, environment and lifestyle in cancer. It is now clear that the integration of existing biobanks into large resources is required to translate this knowledge into clinical use through the development of innovative preventive strategies, diagnostics and therapeutics. There are many obstacles and challenges to such integration, but the BBMRI and caBIG® are addressing these issues to implement a federated network of biobanks. Many of the technical and logistical issues have been addressed; however, IRBs, RECs and scientific management groups that are responsible for these biobanks are still grappling with the social and ethical issues. These issues include consent, withdrawal of samples and data, patient access to information, and research access to information and samples. The roles of IRBs and RECs are becoming central to the management of biobanks, and their integration and the establishment of harmonized international policies are still required to facilitate the smooth sharing and integration of biobanks. If such issues are not dealt with, the public will withdraw its trust and individuals will start to revoke their consent, which would fundamentally affect the quality and completeness of such resources, preventing the ability to draw scientifically valid conclusions from them.

R. William G. Watson is at the UCD School of Medicine and Medical Science, University College Dublin, Ireland.

Elaine W. Kay is at the Department of Pathology, Beaumont Hospital, Ireland, and the Royal College of Surgeons in Ireland.

David Smith is at the Department of General Practice, Royal College of Surgeons in Ireland.

Correspondence to R.W.G.W
e-mail: william.watson@ucd.ie

doi:10.1038/nrc2913
Published online12 August 2010

1. Cohen, M. H. *et al.* Approval summary for imatinib mesylate capsules in the treatment of chronic myelogenous leukaemia. *Clin. Cancer Res.* **8**, 935–942 (2002).
2. Siena, S., Sartore-Bianchi. A., di Nicolantonio, F., Balfour, J. & Bardelli, A. Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. *J. Natl Cancer Inst.* **101**, 1308–1324 (2009).
3. Lemieux, J., Provencher, L., Latreilly, J., Rayson, D. & Wang, B. The role of neoadjuvant Her2-targeted therapies in HER2-overexpressing breast cancers. *Curr. Oncol.* **16**, 48–57 (2009).
4. Tan, D. S. W. *et al.* Biomarker-driven early clinical trials in oncology. *Cancer J.* **15**, 406–420 (2009).
5. Bonavida, B. Preface. Antibody therapies for cancer. *Oncogene* **26**, 3592–3593 (2007).
6. Taube, S. E. *et al.* A perspective on challenges and issues in biomarker development and drug and biomarker co development. *J. Natl Cancer Inst.* **101**, 1453–1463 (2009).
7. Caporaso, N. E. Why have we failed to find the low penetrance genetic constituents of common cancers? *Cancer Epidemiol. Biomarkers Prev.* **11**, 1544–1549 (2002).
8. Burton, P. R. *et al.* Size matters: just how big is BIG? Quantifying realistic sample size requirements for human geneome epidemiology. *Int. J. Epidemiol.* **38**, 263–273 (2009).
9. Bevilacqua, G. *et al.* The role of the pathologist in tissue banking: European consensus expert group report. *Virchows Arch.* **456**, 449–542010.
10. Riegman, P. H. J., de Jong, B. W. D. & Llombart-Bosch, A. The organization of European cancer institute pathobiology working group and its support for European biobanking infrastructure for translational cancer research. *Cancer Epidemiol. Biomarkers Prev.* **19**, 923–926 (2010).
11. Buetow, K. H. An infrastructure for interconnecting research institutions. *Drug Disc. Today* **14**, 605–610 (2009).
12. Goosner, M. National Comprehensive Cancer Networks (NCCN) — big data basis outcomes research begins to yield results. *J. Natl. Cancer Inst.* **102**, 215–217 (2010).
13. Cambon-Thomsen, A. Accessing the impact of biobanks. *Nature Genet.* **34**, 25–26 (2003).
14. International Society of Biological and Environmental Repositories. Best practices for repositories: collection, storage, retrieval and distribution of biological materials for research. *Cell Preserv. Technol.* **6**, 3–56 (2008).
15. The German Ethics Council Biobanks for Research, Opinion. Deutscher Ethikrat [online] (2004).
16. Sheikh, A. A. Genetic research and human biological samples: some legal and ethical considerations. *Med. Law* **23**, 897–912 (2004).
17. Elger, B. S. & Caplan, A. L. Consent and anonymization in research involving biobanks. *EMBO J.* **7**, 661–666 (2006).
18. Annas, G. Rules for research on human genetic variation – lessons from Iceland. *N. Engl. J. Med.* **342**, 1830–1833 (2000).
19. Chadwick, R. & Berg, K. Solidarity and equity: new ethical frameworks for genetic databases. *Nature Rev. Genet.* **2**, 318–321 (2001).
20. O'Neill, O. Informed consent and genetic information. *Stud. Hist. Philos. Biol. Biomed. Sci.* **32**, 689–704 (2001).
21. Ravitsky, V., Fiester, A. & Caplan, A. (eds) *The Penn Center Guide to Bioethics* (Springer Publishing, New York, USA, 2009).
22. Hakimian, R. & Korn, D. Ownership and use of tissue specimens for reseach. *JAMA* **292**, 2500–2505 (2004).
23. Wolf, L. E., Bouley, T. A. & McCulloch, C. E. Genetic research and stored biological materials: ethics and practice. *IRB* **32**, 7–18 (2010).
24. Council of Europe Draft explanatory memorandum to the draft recommendation on research on biological materials of human origin, Strasbourg. Council of Europe Steeting committee on Bioethics [online] (2006).
25. Campbell, A. V. The ethical challenges for genetic databases: safeguarding altruism and trust. *Kings Law J.* **18**, 373–382 (2007).
26. Shickle, D. The consent problem with DNA biobanks. *Stud. Hist. Philos. Biol. Biomed. Sci.* **37**, 503–519 (2006).
27. Helgesson, G., Dillner, J., Carlson, J., Bartram, C. R. & Hansson, M. G. Ethical framework for previously collected biobank samples. *Nature Biotechnol.* **25**, 973–976 (2007).
28. Slater, B. & Jones, M. Biobanks and bioethics: politics of legitimation. *J. Eur. Pub. Pol.* **12**, 710–732 (2005).
29. Putnina, A. Exploring the articulation of agency: Population genome project in Latvia. International Summer Academy on Technology Studies [online] (2002).
30. Gertz, R. Withdrawing from participating in a biobank — a comparative study. *Eur. J. Health Law* **15**, 381–389 (2008).
31. Hansson, M. G. *et al.* Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol.* **3**, 266–269 (2006).

Competing interests statement
The authors declare no competing financial interests.

DATABASES

National Cancer Institute Drug Dictionary: http://www.cancer.gov/drugdictionary
imatinib | trastuzumab

FURTHER INFORMATION

Canadian Tumour Repository Network (CTRNet): http://www.ctrnet.ca
Cancer Bioinformatics Grid: http://cabig.cancer.gov
DataSHaPER: http://www.datashaper.org
EuroBoNet: http://www.eurobonet.eu
European Biobanking and Biomolecular Resources Research Infrastructure (BBMRI): http://www.bbmri.eu
European Forum for Good Clinical Practice: http://www.efgcp.be
National Cancer Institute Best Practices for Biospecimen Resources: http://biospecimens.cancer.gov/bestpractices
National Comprehensive Cancer Network (NCCN): http://www.nccn.org
Office of Biorepositories and Biospecimen Research (OBBR): http://biospecimens.cancer.gov
TuBaFrost: http://www.tubafrost.org

ALL LINKS ARE ACTIVE IN THE ONLINE PDF