CHAPTER 3

Introduction: The Cardinal Role of Biobanks and Human Biospecimen Collections in Biomarker Validation: Issues Impeding Impact of Biomarker Research Outcomes

PASCAL PUCHOIS,*[a] LISA B MIRANDA*[b] AND ALAIN VAN GOOL*[c,d,e]

[a] Trans-Hit Biomarkers Inc., Montreal, Canada; [b] Biobusiness Consulting Inc., Greater Boston Area, USA; [c] TNO Quality of Health, Metabolic Health Research, Leiden, the Netherlands; [d] Laboratory of Genetic Endocrine and Metabolic Diseases, Department of Laboratory Medicine, Radboud University Nijmegen Medical Center, The Netherlands; [e] Faculty of Physics, Mathematics and Informatics, Radboud University Nijmegen, The Netherlands
*Email: pascal.puchois@trans-hit.com; lisabmiranda@biobusinessconsulting.com; alain.vangool@tno.nl

3.1 Introduction

Biobanking is a burgeoning market sector. Hundreds of biobanks exist around the world, with many others in prospective creation regularly. It is estimated

RSC Drug Discovery Series No. 33
Comprehensive Biomarker Discovery and Validation for Clinical Application
Edited by Péter Horvatovich and Rainer Bischoff

Published by the Royal Society of Chemistry, www.rsc.org

that more than 1500 biobanks are currently in operation globally; the greatest prominence of activity aggregated across North America and Europe. The majority of biobanks appear to be cancer oriented, comprising as high as 70% of active biobanks. Other biobanks include a research focus that is disease based, *i.e.* CNS, cardiovascular, metabolic disorders, infectious, immune or rare diseases based. Biobanks collect, manage and store human or animal biological specimens such as blood or tissue from diseased and healthy participants. Historically, collective biological samples or biosamples have proven invaluable in establishing the natural evolution of disease and its transmission. In addition to increasing understanding of disease in general, biospecimens offer a precious resource to support advancement of translational research related to biomarker development and validation.

In the last decade, biospecimen research has also contributed to development and validation of novel diagnostic tests. Diagnostic tests can be utilized to diagnose and monitor disease progression as well as predict drug efficacy and risk for adverse events. Such tests can also facilitate stratification of patients based on response to therapy, thereby optimizing clinical treatment options and patient outcomes alike. One case study is the human epidermal growth factor receptor-2 (HER2) gene. Predictive testing for HER2 protein led to development of trastuzumab drug. Predictive tests for other genetic biomarkers (*e.g.*, BCR/ABL tyrosine kinase, EGFR/KRAS) have been similarly codeveloped for other targeted anticancer therapies. Utilization of biospecimen collections has also bolstered development of prognostic tests that stratify patients by disease outcome to guide clinical decision making. Recent examples include Oncotype DX 21-gene breast cancer and MammaPrint assays, both of which predict the risk of disease recurrence in certain women with early-stage, node-negative breast cancer.[1] Based on these and other successful proof-of-concepts, one may surmise that biospecimen research if harnessed efficaciously may considerably contribute to the discovery and validation of novel biomarkers and related clinical therapies.

To date, utilization of biomarker-disease based and molecular-medicine-related research potential has been at best somewhat diminutive. Only a minority of candidate biomarkers progress beyond initial discovery. The few biomarkers that do proceed beyond the initial phase of development tend to have a high degree of unreliability. To further complicate matters, the ability to progress candidate biomarkers into large-scale validation studies poses further challenges. Quality and relevance of biospecimen and data collections remain key limiting factors for biomarker research in general. In the present chapter, we commence with an overview of the biobanking–biomarker collaborative landscape in an effort to elucidate, fundamental planning considerations for optimizing biomarker research and validation. We will next present prevailing challenges in biospecimen cultivation and utilization that compound biomarker research outcomes, and then follow with recommendations as to how these challenges may be ameliorated. In tandem, we offer guidance as to how biospecimen and biomedical research practices may be augmented to advance biomarker discovery and amplify development outcomes towards maximal

impact, emphasizing that cardinal collaborations are imperative to achieve such impact.

3.2 Navigating the Biobanking–Biomarker Collaborative Landscape

3.2.1 Crucial Considerations in Planning Biobank–Biomarker Research Collaborations

Historically, biomarker research practices and biobanking collection priorities have been developed prior to systematic surveying of the therapeutic and/or drug discovery market landscape. The majority of specimens collected have been interdependent and largely driven by internal stakeholder research biospecimen supply and demand. The past decade's surge in biobanking formalization sheds light on acute issues that require early incorporation into technical and scientific planning processes if the short- and long-term impact of biomarker research outcomes is to improve. Biomarker scientists may not be cognizant of the requisite due diligence affiliated requirements of the collaborative value chain, that is how to strategically plan for and successfully navigate biomarker–biobanking research pathways. Extramural and interoperational research activities tend to be frequently impeded by limited transboundary education, feasibility review and prospectively proactive logistical planning. To ensure "fit for purpose" and impactful research and clinical outcomes alike, it is especially important that biomarker scientists acquire education on quality issues biobanks face in operations. Scrutiny is not restricted to biospecimen quality review; it permeates the continuous research pathway. It is crucial that biomarker scientists optimize approach and evaluation of biobanks during preinvestigative feasibility planning. It is a false assumption that biobanks and their respective collections are equivalent to each other. Biobanking collection practices differ dramatically; the relevance and quality of collections vary in tandem. Historically, the majority of biospecimen and data collections are not prospectively designed with biomarker research requirements in mind. Clarity of understanding is a prerequisite to more aptly understand how best to define and cultivate relevant biomarker-ready biospecimen and data collections.

Biobanks range in infrastructure, operational models and extent of expertise. Biobank infrastructures permutate from informal collections in a researcher's freezer to dedicated facilities.[2] Historically informal collections were referred to as "tissue banks" or "biobanks" when inventory included biofluids and other source materials, in additional to tissue specimens. The term "biobank" is often the vernacular. The following terms are also relatively synonymous but tend to encompass a significantly higher degree of formal infrastructure: biospecimen resources, biorepositories, tissue banks, core bioresource facility, sample management laboratory, biolibraries and biological resource centers. Some bioresources identify with the term biological resource centers (BRCs). BRCs

tend to be broader in scope of operations and services, and are designed as an applied research complex that processes source biological materials into cells and cellular components, and function in manufacturing environment conditions.

Trans-Hit Biomarkers (internal database) has identified over 1500 biobanks worldwide; it is unknown how many human biospecimen banks exist in total. It is known that biobanks are prominent in both nonprofit and commercial sectors across the market. In the nonprofit sector, biobanks function in government research laboratories, nonprofit research institutes, academia, and community-based hospitals, cooperative entities (United States based examples *e.g.* Eastern Cooperative Oncology Group (ECOG), Southwest Oncology Group (SWOG), and Gyneacologic Oncology Group (GOG)) and nonprofit charitable/community-based organizations (US-based examples *e.g.* Accelerated Cure Project for Multiple Sclerosis, National Cancer Society (NCS), Juvenile Diabetes Research Foundation (JDRF)). In the profit sector biobanks operate as services providers, as commercial laboratories and sometimes as Clinical Research Organizations (CROs), but may also function as pharmaceutical and biotech-based sample management facilities. As the precompetitive domain amplifies, market expansion of biobanks appears to be increasing. The resulting variance in market activity has a causative effect on further evolving biobanking model form and function.

The most widely known models in human biobanking are biosample-based bench collections, clinical trial repositories and pathology core resources. Recent efforts have focused on centralization of biobanking resources in concert with development and harmonization of networks. These activities as well as network to network linkages will continue to burgeon; it is not possible to discern if such efforts will be sustainable long term without timely integration of formal business models and cost recovery. Aggregation of sample and data collections is also an increasing trend. The shift towards aggregation appears largely due to the urgency to increase sample collection size aimed at ameliorating statistical significance of downstream research results. As a result, some scientists have begun to search for retrospective compatible specimens for their research efforts from other sources, both internal and external. Others have begun to amalagamate prospective banking of particular samples to acquire a statistically significant relevant sample size. Biobanking research objectives also differ. While some biobanks focus solely on collecting specimens for a key research purpose, other biobanks may collect for a wide range of research needs. Disease-specific banks are becoming more popular; it is speculated that disease-specific banks may have a greater propensity of producing quality collections due to their concentrated focus. Moreover, population-based banking which ranges in prominence geographically, appears to be a renewed consideration for those with access to disease-specific registries. Population-based banking entails procuring samples from both diseased and healthy individuals in communities towards alleviating a specific clinical outcome. Population-based banking has occurred for decades in the context of public and private epidemiologic research but has rarely been harmonized with central biobanking efforts. Historic collection of specimens has also occurred in vast

quantities *via* drug clinical trials in which many biosample and data collections have been collected and stored for private research efforts. All of these collections may hold potential value for biomarker research providing evidence-based quality parameters can be defined and achieved.

There is often confusion in the nomenclature in regards to what denotes a "commercial" biobank *vs*. a noncommercial biobank and how operationally these biobanks may contrast. The domain specific term refers to what are in reality, commercial tissue procurement organizations. Commercial tissue procurement organizations appear to be a growing trend with at least fifty commercial tissue procurement organizational biobanks operating in the for-profit sector alone. While these resources may bear physical resemblance to academic bioresources in general, what differs is that these organizations typically procure specimens and data secondhand from primary sources, *e.g*. academic medical centers and community-based hospital partners. These partners tend to be located in a variety of geographic areas ranging from North America, Western and Eastern Europe, as well as Asia and South America. It is meaningful to understand the downstream issues resulting from access and availability of primary versus secondary source biological material and data. Issues tend to pertain to quality, *i.e*. the extent and authentication of annotation. It is difficult to conclude which type of bioresource offers the highest quality of biospecimens and data collections. However, it does appear that commercial vendors tend to be more aware of their inventory and have shorter lead times related to feasibility studies and to provision of retrospective samples. The type of biospecimens available may also differ, *e.g*. primary materials (blood, urine, tissue) or bioproduced materials, *e.g*. derivatives (extracted RNA, DNA, tissue array). One may therefore presume that biospecimen and data collections are likely to vary in quality and content based on the source origin and type of organization collecting them.

Harmonization of standardized sample management methodologies and research practice is thought to be a logical solution to defining and improving quality of biosample collections and related biobanking activities. However, the wide geographic span and vast diversity of biobanking activities creates challenges to expeditious standardization. In addition to disparities in infrastructure, research expertise and type of collections, biobanks differ in terms of types of sample management protocols, bioprocessing techniques, and biobanking workflows. Variation also exists in volume of biosample throughput, along with access to and utilization of manual and automated and benchtop sample management technology and instrumentation. There are some who intimate that greater implementation of automated technologies may alleviate prevailing issues related to standardization that impact quality. Degree of cohesion of oversight may also vary, which affects the extent of quality biobanking and sample management practices. Tissue banks are typically governed by internal parties, *i.e*. coordinating Directors, who provide primary oversight aided by key adjunct advisors and/or principal investigators, Secondary oversight is typically offered by founding stakeholders, *i.e*. department heads, and administrators. Core facilities tend to have scientific

advisory committees composed of interdisciplinary parties both internal and external to the direct bioresource. Oversight may also be offered *via* biospecimen "utilization" committees. A biospecimen utilization committee's degree of contribution, expertise and effort provided tends to vary from informal gestalt and "weigh in" on organizational planning and operations to formal review of upcoming projects. Current oversight practices, *i.e.* scientific and technical advisory tend to range in frequency (1–4 times a year) and mode (*e.g.* face to face meetings, conference calls, and/or emails). Further information on bioresource oversight may be referenced in Section B.1.2.2 of the 2011 NIH/NCI BBRB best practices focusing on guidance related to bioresource operation and management.[2]

As briefly referred to above, biobanking experts have contributed to the development of experiental best-practice-based technical and operational guidelines with the aim of improving quality and harmonizing standardization of specimen and data collections. Currently, there are four established primary best-practice reference documents:

- "National Institutes of Health/National Cancer Institute's Biorepositories and Biospecimen Research Branch (BBRB) Best Practices for Biospecimen Resources."[2]
- "International Society of Biological and Environmental Repositories (ISBER) Best Practices for Repositories: Collection, Retrieval, and Distribution of Biological Materials for Research."[3]
- "Organization for Economic Co-operation and Development (OECD). Best Practice Guidelines for Biological Resource Centers".[4]
- "World Health Organization International Agency for Research on Cancer (IARC). Common Minimum Standards and Protocols for Biological Resource Centers Dedicated to Cancer Research").[5]

Scientific working groups have published reporting guidelines that dispense advice as how to report research methods and findings. While originally developed to increase the quality of data reported and/or presented in publication, guidelines *e.g.* REMARK (prognostic tumor markers)[6] (STARD (diagnostic accuracy),[7] STROBE (observational studies),[8] and BRISQ (biospecimen reporting)[9] can be effective tools when incorporated prior to execution of research, particularly in enhancing prospective biomarker research protocols and biobank collection procedure design. Further information including a comprehensive list of reporting guidelines may be obtained *via* the Equator Network website (http://www.equator-network.org). If used efficaciously best practice and reporting guidelines may then serve as tools to enhance research parameters, define benchmarks and increase statistical relevance. The long-term goal for both best practice and reporting guidelines is to incorporate empirical data and evidence-based lessons learned from biospecimen science research over time.[10] However, prevalence, rate of adoption as well as frequency and manner of utilization of best practice and reporting guidance remains unconfirmed. One could envision that statistics on these variables may function

as a predictor of the level of a biospecimen and data collection's inherent quality. However, evidence to this effect is still required to demonstrate this hypothesis. The BBMRI/Gen2Phen Bioresource Research Impact Factor (BRIF) working group is currently collaborating to evaluate the impact of related biobanking supported research outcomes. Over time these efforts could include validating this idea and yield much information as to the actual impact of biospecimen based research, including biomarker-based studies.[11]

As awareness of the inherent scientific and market value of biospecimen and data collections has increased, stakeholders have realized that high-quality, dedicated, formalized bioresources have the potential to serve as hubs for biotechnological research and development. In an effort to aggregate and optimize resources both public (*i.e.* government network models) and private institutions, have evolved towards assembling networks of biobanks. One well-known network is the Centro Nacional de Investigaciones Oncologicas (CNIO)'s National Biobank Network which was created in September 2000 and incorporates over 25 institutions across Spain and Latin America.[12] More recently, networks have engaged in research-driven business-based collaborations in public–private consortia partnerships. There has been an increase in consortia activity cross-sector as well as crosscontinent. Examples include the National Foundation for Cancer Research's Tissue Bank Consortium in Asia (http://www.nfcr.org/index.php/research/tissue-bank-consortium-in-asia), the United States Foundation for the National Institutes of Health (FNIH) Biomarker's Consortium (http://www.biomarkersconsortium.org/), and European networks like the Biobanking and Biomolecular Research Infrastructure (BBMRI) (http://www.bbmri.nl/en-gb/home) and the String of Pearls Institute (http://www.string-of-pearls.org/). Efforts in formalizing networks and assembling consortia in general have been focused on linking bioresources and their collections demographically (country, continent, inter/intracontinental), by disease and/or research focus *via* common operating standards, unified adoption of best practice, quality-control procedures, informatics platforms (databases and specimen locators); such initiatives are still in development. Biobanking societies can also play a role in harmonization biobanking education and awareness. Societies *e.g.* the International Society of Biological and Environmental Repositories (www.isber.org), Public Population Project in Genomics (P3G) Consortia (www.p3g.org), European, Middle Eastern and African Society for Biopreservation and Biobanking (ESBB) (www.esbb.org/) have all played a role in moving the field forward.

3.2.2 Biobanking Challenges that Impede Impact of Biomarker Research Outcomes

Biobankers and biospecimen researchers in tandem with scientists who perform biomarker research have the capability to advance biomarker research and development aimed at improvement of outcomes related to incidence and prevalence of disease, diagnosis, treatment, as well as disease prevention and quality of life specific to both diseased and healthy populations. Currently, these

capabilities and any related impact they may offer are impeded by significant common challenges. Some of the key challenges have been recently described.[13–15] These include, but may not be limited to: long-term sustainability of biobanking operations and infrastructure, formal execution of cost recovery, real-time implementation of best practice and quality sample management review and bioresource evaluation, expedition of "fit for purpose" and evidence-based biobanking practice, availability of tools to enhance interoperability of specimen sharing and proactive governance and utilization review.

Trans-Hit Biomarkers Inc performed a survey (internal data) to pinpoint the key issues that scientists in academia and industry face when searching for biospecimen collections.

- Biomarker scientists reported difficulties in locating sources of biospecimens due to the lack of a dedicated comprehensive global registry, in particular biospecimens for rare disease states.
- During their searching process most scientists noted that biobanks were unable to perform rapid multicriteria search based on associated medical data. They also found that the majority of biobanks they interacted with did not have facile access to medical data. Experience demonstrated a high degree of variance of information technology and bioinformatics tools in practice (electronic sample inventory management systems) which created problems with achieving interoperability between the researcher and the biobank.
- Downstream analysis was further complicated by the fact that operating procedures for key steps in the biospecimen management process were frequently missing *e.g.* time for warm and cold ischemia for tissue.
- Requisite clinical and scientific input in the design and management of collections was inconsistent.
- For a given disease, few control groups were consistently available within the same biobank. Control groups that were available were not appropriate for assessing biomarker specificity (*e.g.* prostate cancer, benign prostatic hyperplasia, prostatitis, and other cancers).
- Additionally, the entire cycle of disease progression was not covered by collection. (For example, it is common to find collections in the cancer field for late-stage patients, but be very difficult to find specimens from cancer patient with early-stage disease.)
- Biomarker researchers noted that available specimen sets were inconsistent on overall search. (For example, some biobanks had tissues, while others had serum, DNA, *etc.*); incidence was rare to find biobanks with inventory that included a wide range of biospecimen types for the same patient.
- In addition to issues with samples, scientists repeatedly had problems with finding the complete associated clinical data sets and medical follow-up data.
- Quality management due diligence was arduous; a review of audit trails for individual biospecimen sets was not always possible.
- The most crippling issue reported was restricted access to industry partners by numerous academic biobanks.

– Lastly, the overall consensus amongst biomarker scientists was that timelines to conduct feasibility studies and the administrative process were highly unsustainable.

Understanding the challenges biobanks face helps clarify the high prevalence of reported impediments. There appears to be four key areas of dysfunction: 1) unsustainable biobanking operational models, 2) difficulties in locating and accessing biospecimen and data collections, 3) inefficient research business practices and resulting inability to engage in timely collaborations and 4) low prevalence of satisfactory, relevant biospecimen and data collections. Further insight on these issues is offered below.

– Issue #1: Unsustainable biobanking operational models

The most prevalent impediment to both quality biospecimen and biomarker research today is the disparity in biobanking business infrastructure and formal business models. Both historically and from one standpoint, biobanks have not and continue still to refrain from establishment of an econometrically data-driven business plan, outlining a short- and long-term strategy regarding the size of the biobank and return of investment. In the middle, an unconfirmed percentage of biobanks do practice cost recovery and have business practices, but business models tend to be limited in scope and/or planning. At the other end of the spectrum, some biobanks have extremely formal business and cost-recovery models with limited flexibility or that are resistant to market adaptation. Even still, many biobanks have no formal business infrastructure. The end result of lack of business planning tends to frequently result in significant lags in engagement and conduct of research, quality review and curation of specimen and data collections as well as transfer of specimens and data. Such issues can also compound risk of loss in return on investment and contribute to increased costs of scientific and clinical collaboration.

Sustainability is further retarded by insufficient strategic planning, specifically foresight regarding budgeting of funding and costs to support research projects for most biobanks, even those that generate revenue *via* cost recovery. In recent years, many governments and funding bodies have encouraged the establishment of large-population biobanks and networks. Despite continued funding, most are either underfunded or funded with unsecured revenue streams. Stakeholders are imprecise as to the significant resource investment, time and cost of operations that are required for a quality biobanking effort. Development of such a facility may require anywhere from $3 to 6 million USD depending on infrastructure. Operating costs and initial resource investment vary per bioresource depending on the size of the biobank, its business model, related activities and collections. Labor, equipment and materials tend to be the largest costs.[14,16]

– Issue #2: Difficulties in locating and accessing biospecimen and data collections

One of the key challenges to obtaining samples for biomarker research is success in locating and identifying relevant sample collections. Locating and targeting the correct set of samples and data is a pivotal requirement to support biomarker research. The primary mechanisms of acquiring specimen collections have been: 1) direct procurement from local sources (*e.g.* clinicians' colleagues within the same hospital) or 2) establishment of new relationships for individual interrogation of collections of identified sample and data sets (*via* phone, email, and web-based queries). Some repertories exist but most of these registries are national or managed by biobanks associations and access for searching is restricted. As of today, there is no exhaustive full accessible international repertory. Manual investigation of available samples is time consuming and costly and occurs prior to quality testing or downstream analysis of the sample. New trends to improve specimen searching, cultivation and access include creation of web-based technologies *i.e.* online catalogs and biolibraries, web portals and specimen locators. The NIH Specimen Resource Locator (http://pluto3.nci.nih.gov/tissue/default.htm) is one example. However specimen locator-based databases typically include access only to the resources within their biospecimen research network. Another example is the gratis Trans-Hit Biomarkers portal, B4B-Hub (Biobanks for Biomarkers: www.trans-hit.com) that allows biomarker scientists to post their biospecimen requests that are automatically sent to all biobanks members. Trans-Hit's database includes more than 1500 biobanks worldwide. In the future, greater availability of informatics tools would allow biomarker scientists rapid real-time capability to search multiple data fields and more aptly evaluate criteria for donor-specific sample sets and associated clinical data.

– Issue #3: Inefficient research business practices and resulting inability to engage in timely collaborations

Access to biobanks continues to be a prevalent barrier to research. Frequently biobanks restrict access to intramural scientists, preventing access to external scientists, particularly those from industry. Many biobanks refuse to allow access to their biospecimen collections citing institutional ethical constraints or questionable rationale. This practice limits many biomarkers to be effectively used in biomarker validation and in translation from bench to bedside. Thus, it decreases the potential for the biobank to add value to their basic research programs by unifying applied research to facilitate new biomarkers. As a result, biobank and biomarker scientists are limited in capability to increase their knowledge base requirements to perform quality, "fit for purpose" and benchmark-level biobanking and sample management. In general, "fit for purpose" is a term used to clarify that the methodology of the research design should match the intended downstream purpose of the research application. Benchmark recommendations are still being explored across the industry overall, but the concept involves the idea that performance metrics emulate the highest level of defined practice. This lack of extramural collaboration and shared learning ultimately hurts the ultimate stakeholder in

biomarker research, the donors that support such research as well as prospective patients who may benefit from related discoveries. When access is available to external scientists, some biobanks implement policies with unreasonable access fees that appear to be applied solely to scientists from industry.

Another major obstacle is lag in delivery of biospecimens, even from collections already in stock and timelines to deliver biospecimen collections often due to labor required to investigate requests and delayed execution of administrative overhead. As a result, too frequently biomarker development projects are critically delayed.

In some countries, *e.g.* China, export of some types of biospecimens is not allowed. This seriously hampers validation studies and commercialization of *In Vitro* Diagnostic (IVD) tests in China, unless partnering strategies with local companies are being followed.

Overall, such constraints in biobank access, donor recruitment, and project management related timelines create extensive lags, often putting biomarker research programs at dramatic risk.

In addition to issues related to access, there are significant inefficiencies with the lags in the research administrative process. While requisite due diligence is understandable (review and approvals by the biobank's biospecimen use committee and/or institutional ethical board (IRB) approval and/or financial committee) this process tends to be overcomplicated. Global divergence in biobanking ethical, legal and regulatory frameworks induces significant delays, *e.g.* some institutions refuse to implement standardized language templates. Inefficiencies executing contractual agreements for extramural research limit industrial biomarker collaborations due to strict filing timelines that the large majority of biobanks in operation are not usually aware of.

Often, terms for collaboration tend not to be well defined or justified. The terms do not necessarily appear reasonable to the contributions garnered. The most frequent area of difficulty is the negotiation of intellectual property (IP).[17] Some biobanks insist on sharing of downstream IP in accordance with provision of biospecimens. More often biobanks desire coauthorship in future publications, which is justified mainly when there is a scientific input from the biobank. Terms for collaboration, including IP should be well discussed early on. On the other hand, when papers are published by biomarker scientists, they often fail to provide sufficient and crucial information about the source and biopsecimen.

Another prominent issue tends to be agreement on laboratory location of conduct of downstream analysis. Some biobanks will only collaborate if laboratory work can be conducted in their laboratory. This request is particularly problematic for industrial organizations to comply with due to regulatory requirements aimed at ensuring environmental quality control throughout the process chain related prefiling and precommercial activities.

Once terms are rendered and agreed upon, additional delays may occur due to lags in approval of material transfer agreement, where simple agreements may take months or up to a year. Mayol-Heath[18] and colleagues recently published an article in which they offered insight into navigating the regulatory process for biospecimen research in the United States as depicted. It is required

that each biobank follow specific steps to obtain and maintain institutional regulatory board approval. This process can take an average of 1–4 months, depending on the related requirements (whether informed consent is needed), complexity of the research project, types of samples being collected and/or requested and frequency of IRB review board meetings.

– Issue #4: Low prevalence of satisfactory, relevant biospecimen and data collections

Many biobanks do not have complete or well-documented Standard Operating Procedures (SOPs). This deficiency increases difficulty in qualifying relevance and quality of biospecimen and data collections. SOPs for annotating, collecting, processing, storing biospecimen are usually different from site to site and can result in significant variation in studied biomarkers that has nothing to do with disease. Care should be taken when selecting biospecimen collections and biobanks. This is even more crucial for validation studies, when collections from different sources may be required.

Few cohort-based biobanking projects have prospectively stored an extended variety of specimens for a single subject (plasma, serum, urine, tumor, DNA, RNA, *etc.*). Most centralized tissue based biobanks are initiated in pathology departments and historically have focused on collection of fresh tissue in excess, which is usually stored in FFPE and frozen formats. Many biobanks do not collect other fluids that can be very useful for translational research and biomarker validation.

When developing and/or validating a biomarker it is important to look at biomarker profiles in control cohorts. Controls could be matched healthy subjects but may also be other types of patients who might express same profile of biomarkers or the same kind of clinical symptoms. While many biobanks may have very substantial collections for a given medical field they typically do not collect control groups.

Another key issue is that clinical and/or downstream annotation not readily available if not planned in advance – many biobanks do not have an easy and full access to clinical data. The biobanks that do have access to clinical annotation often obtain this data through retrospective medical file review, which is more difficult to validate and not necessarily "fit for purpose". As a result, data sets vary dramatically in terms of content, range of annotation, quality and completeness.

Biomarker research is often conducted utilizing aggregate samples obtained from either one or a variety of retrospective collections. Retrospective collections can certainly be used for biomarker discovery and in many cases for biomarker validation. Depending upon the final intended use, application of a biomarker, which may be submitted to different regulatory requirements (please refer to Chapter 2 of this book on regulatory aspects) retrospective samples could potentially save significant time and money compared to prospective studies. Harnessing historical collections from retrospective studies may offer the benefit of expediting such valuable research and reduce costly delays

and limit the need to execute lengthy, expensive validation studies, therefore dramatically accelerating clinical validation of future diagnostic products. However, for eventual future regulatory submission to the US Food and Drug Administration (FDA), prospective collections can be mandatory.

Clinical trials often include dedicated biomarker studies and having access to archived patient specimens from prospective clinical trials could greatly contribute to the validation of biomarkers, particularly new companion or prognostic biomarkers. Pharmaceutical sponsors have dedicated high caution and budgets for well designing and conducting their clinical trials, collecting high-quality specimens and clinical data. As mentioned by Lindpainter,[19] even if only a fraction of the more than four million subjects enrolled worldwide in interventional trials were to be captured, progress would be enormous. However, regulatory and ethical issues (*i.e.* patient inform consent) considerably limit the use of biospecimen that have been collected during past clinical trials. The ability to acquire approval from sponsors of these clinical trials (which were successful or failed), to have access to these collections is also critical and uncertain.

3.2.3 Effect of Process-Chain Impediments on Impact of Biospecimen Collection Quality

The health-care industry spends billions of dollars annually on biomarker research for personalized medicine. Despite this investment, research into biomarker-disease-associated molecular changes in body tissues and fluids hasn't yet delivered on its promise. Technologies such as proteomics and genomics have contributed voluminous literature documenting thousands of claimed biomarkers;[20,21] Rare candidate biomarkers failed to replicate when tested on larger-sample cohorts and therefore have not been fully validated and/or approved for routine clinical practice. Several papers reported alarming findings concerning the reliability and reproducibility of published scientific data, which could not be replicated by a third party.[22,23] The poor reproducibility of results is alarming but may possibly be explained by differences in laboratory equipment, protocols, and personal skills and related to biomarkers, the most relevant factor is likely to be the variation in starting material that may differ between laboratories.

Besides the lack of substantial validation clinical studies, one of the primary impediments to progress the full validation for biomarkers is the lack of standardization in how specimens are collected, handled and stored. It is well known that the quality of the starting material determines the accuracy and reliability of diagnostic assay results. Therefore, it is crucial that the appropriate samples are managed in a robust manner to ensure the conclusions made based upon the derived data are reproducible and reliable. Ensuring the highest quality sample-management processes is of vital importance to ensure the robustness of the data, decision making and conclusions from studies.[24,25] Inaccurate results owing to compromised tissue quality can lead to false-positive or false-negative results with therapeutic consequences that can harm patients and affect their eventual outcome. As pointed out by Foot,[26] "if

there are no samples, there is no science''. Many studies have demonstrated the adverse impact of the inferior quality of samples on the generation of bias and therefore the reliability of the data derived from their analysis.[27,28]

The major impediment to progress in the validation of new biomarkers is the lack of standardization in how specimens are collected, handled and stored. It is known that at least 60% of errors occurring in routine clinical chemistry laboratories are due to mistakes made in the preanalytical phase. Most of them are attributable to mishandling procedures during collection, handling, preparing or storing the specimens. Sample integrity is of paramount importance in order to make robust conclusions from the data, from the point of collection, through handling and shipment, storage and sample management and to ensure efficient retrieval of the correct sample and ability to link with relevant clinical data. In nearly one fifth of the cases they can produce inappropriate investigations and clinical decisions.[29] If this observation is true for routine assays (already validated and commercialized), even though the laboratory itself is accredited and/or has an outstanding performance, one can imagine what happens when new biomarkers are studied in a noncontrolled quality laboratory environment. This is particularly important for biological fluids but also for tissues especially in the cancer field.[30–38] It is essential that the levels of expression of biomarkers that are used to make treatment decisions be accurately attributable to the underlying biology and not due to artifacts. Several authors have identified a variety of issues and barriers that can affect the transfer of clinical tests from research to clinical practice: differences in sample collection, handling or storage.

The effects of tissue degradation and decay have been reported for many molecules[34–36,39] Hicks and Boyce[40] recently reported how the quality of tumor tissue, determines the accuracy and reliability of companion diagnostic assay results and therefore optimal therapeutic strategies. Preanalytical variables are the ''weakest link'' of immunohistochemistry, as the properties of the tissues analyzed and, consequently, the results, may be affected by several factors, including the time to collection (length of time that tissues are subjected to warm ischemia between removal of the tissue at surgery and fixation), details of fixation (type of fixative agent used and length and conditions of fixation), dehydration steps, and conditions for paraffin embedding. These preanalytical parameters are beyond the control of investigators, are most often unrecorded, and constitute a major potential source of bias.[41]

Variables involved with tissue samples begin with taking the sample from the patient through medical or surgical procedures, and continue from acquisition through handling and processing, storage, distribution, analysis, and restocking unused samples.

A succession of variables can affect tissues samples:

– The type of sampling procedure: for instance, Pinhel *et al.*[42] comparing needle core and excision biopsy breast cancer specimens have shown the immunohistochemical reactivity for phospho-Akt and phospho-Erk1/2 was markedly reduced in the latter specimen type.

- The warm ischemia time (the time the surgeon has clamped the blood flow to the tissue and it is cut off from oxygen and nutrients before removing the sample from the patient): during this warm ischemia time the tissue remains at body temperature. Ligation of the blood supply to living tissues being excised during surgery leads to hypoxia, ischemia, metabolic stress, and the progressive degradation of macromolecules that are of potential clinical interest. During the period between artery ligation, tissue removal from the patient, and the start of fixation, the tissue remains alive and reactive, and the level of expression of gene transcripts and proteins can change significantly during this ischemic interval.[43–45]
- The cold ischemia time (the time between removal of the tissue from the body until it is fixed or frozen): immediately after removal of a tissue sample from a patient, cells in the tissue adapt to the absence of blood flow, ischemia, hypoxia, metabolic acidosis, accumulation of cellular waste, absence of electrolytes, and temperature changes.[43] Within a few minutes major changes start to occur in the protein signaling pathways of the tissue as it remains at ambient temperature. While awaiting fixation or freezing, the tissue is alive and "wounded". Post-translational-modified proteins, such as phosphoproteins are among the most labile molecular features. Espina *et al*.[43] reported fluctuations of up to more than 7-fold in some phosphoproteins when samples were incubated at room temperature for various times after excision. Studies have shown that degradation of a number of proteins and peptides starts immediately after disconnecting a tissue sample from its oxygen and nutrient supply.[46] The first changes in gene and protein expression are observed within minutes from excision of a tumor, and after a half-hour 20% of all detectable genes and proteins significantly different expression,[47] reflecting the shutdown and activation of different pathways.

Recent reports suggest that delays from tissue collection to the initiation of formalin fixation also may have an adverse impact on the analysis of hormone receptor assays.[48,49] Khoury *et al*.[49] also showed that the level of expression of estrogen receptors (ER) by immunohistochemistry starts to fall after a two-hour delay from collection to the start of fixation and similar changes are seen for progesterone receptor (PR) after only 1 h. Most surprisingly, these investigators demonstrated that HER2 fluorescence *in situ* hybridization intensity begins to show compromised interpretation after only one hour, and this reduction becomes statistically significant after two hours. The implications for these findings are that some tumors with excessive ischemic times may be classified falsely as negative for the expression of these important therapeutic targets because of protein degradation during this period. Invalid testing of breast predictive factors could have serious consequences for patients.

Time intervals from blood-vessel ligation or tissue removal to examination in the pathology laboratory may vary from minutes to hours depending upon the time of day, the type of procedure, and the proximity of the operating room while the stability of studied biomarkers might vary between sites. In order to be able to calculate and monitor cold ischemic times, the American Society of

Clinical Oncology ASCO/College of American Pathologists CAP task force has developed guidelines for ER and PR testing requiring that breast biopsy and excised breast tissue samples be placed in formalin within one hour of excision[50] This requires that the collection time for each sample be recorded in the operating room, the transport of the tissue samples to the laboratory be expedited, followed by a record of the fixation start times.

– Other variables include the type of fixative, the fixation time, the rate of freezing, and the size of the sample. Formalin fixation and paraffin embedding (FFPE) is still the gold standard for tissue-sample preservation in the clinical setting. However, pathologists consider formalin to create "standard artifacts", but still little is known on the effects of formalin fixation on many (potential) biomarkers. Length and conditions of fixation have been branded as the "Achilles heel" of phosphoprotein assessment in clinical specimens.[42] Dehydration steps and conditions for paraffin embedding (temperature of the paraffin) may also have an influence. To complicate matters even more, different constituents of the sample will undergo decay at a different rate, phosphoproteins being among the most sensitive.

In the future, the use of reliable decay indicators of quality of biospecimen should allow simplifying and testifying as to whether a biospecimen collected and stored according to different protocols may be appropriate for biomarker research. With respect to fluids, several biomarkers have been studied but none have demonstrated so far their true usefulness. Cytokines CD40L has been proposed to be a potential marker to assess the quality of storage for blood samples.[51] For tissues, developing a list of tissue protein stability surrogate markers also remains an important goal for molecular profiling research.[52] Indeed there is an urgent need to document metrics and a cutoff at which the change is so great that it renders a specimen inadequate for diagnostic or research purposes.[52] For any of these quality-assessment assays to provide meaningful information, they must be able to guide the selection of tissue for a given assay or measurement; ideally by providing a numerical cutoff above which that assay can be used reliably.

3.3 Recommendations: Reducing Disparity of Impact and Lag in Outcomes of Biomarker Research

3.3.1 Technical and Scientific Recommendations for Biomarker Scientists

For the context of this discussion, we regard biomarker scientists as those focusing on the discovery, validation, development and implementation of biomarkers.

– Scientific and technological advancements in genomics, transcriptomics, proteomics, metabolomics, and novel approaches in bioinformatics and

systems biology have spawned new insights into the molecular processes underlying complex diseases, which are important drivers for the discovery of novel biomarker candidates. Before 2001, biomarkers appeared less than 1000 times in patents and scientific papers. In recent years more and more data have been produced about biomarkers as putative future diagnostic tests. By January 2012 PubMed referenced more than 553 000 hits, while a Google search produced 8 500 000 results for the term ''biomarkers''. Every week several articles are issued announcing an academic research group has identified a new protein, gene, or other types of marker to better diagnose patients with a specific medical condition. Although the vast majority of these scientific discoveries are pertinent inventions, most of these discoveries are not translated into commercial products or R&D pipelines. Results obtained with many of these biomarkers remain controversial and not conclusive. In many cases the clinical proof-of-concepts are weak, and are often based on small, poorly designed studies with limited numbers of patients and/or the use of suboptimal quality sample collections (poor quality, incomplete clinical information, combination of different sample cohorts with distinct characteristics that create bias, *etc*.). Actually, the success rate in biomarker discovery and validation versus the level of biospecimen and data collection quality has a high degree of association. Ioannidis and Panagiotou[53] described that the initial effect size of a biomarker in selected clinical populations can in the majority of cases not be confirmed in meta-analyses of subsequent studies. Surprisingly, even publications in prestigious peer-reviewed journals or from several independent groups did not ensure reproducibility. Too many scientific papers suffer from various pre- and postanalytical bias and confounders relevant to sample collection, processing and downstream testing, including assays. These biases fail to be recognized in time and often lead to invalid discoveries, *i.e*. putative biomarkers that can't withstand further testing. There is an obvious and urgent need for scientists to ensure the reliability and reproducibility of their data before publishing any new biomarker candidates. As a key factor in the value chain, biobanks can certainly perform an important role in generating much needed ancillary data to validate the robustness of the scientific findings. Unfortunately, most of the literature published by biomarker scientists contained no or little information about methods and materials and the types of biospecimens used.[21,54] As poor-quality samples can only serve to jeopardize research results, it is of the utmost importance to critically select high-quality collections before starting a biomarker project. By carefully selecting biospecimen collections and biobanks, scientists are more likely to increase their chances to adequately address scientific questions, and fulfill regulatory requirements in anticipation of bringing a biomarker to market.

– It is crucial for biomarker scientists to confirm the objectives and research hypothesis of interest upfront, and to define the future intended use for the final assay be made in order to better define the development plan and

therefore inclusion criteria for patients. Consultation with biobanking colleagues will also ensure that specimen and data collections are timely, relevant and can be procured prospectively.

– It is also important to ascertain what type of collection, retrospective or prospective biobanking collections is most ideal to support biomarker research. When using biospecimen collection, *a fortiori* for eventual future regulatory submission to the FDA, great care must be taken to utilize the most appropriate samples. Both prospective and retrospective samples can be used, but each has its own set of drawbacks and regulatory concerns (Table 3.1). The path for clinical and regulatory plans must be previously defined. The use of retrospective samples is not the right solution in every case. However, depending upon the classification of product within the regulatory guidelines and intended use application, retrospective samples could potentially save significant time and money. When prospective studies are needed for the clinical validation of new biomarkers, the setup and monitoring investigator sites, respecting Good Clinical Practices guidelines and maintaining high-quality standards for biospecimen sampling, shipping and storage are crucial (Table 3.1).
– If biospecimen and data collections are needed for a biomarker project, project plans should aim to factor in time and resources required to investigate numerous biobanks before finding the right source and collaboration. Currently, too many researchers rely on whatever specimens they can obtain conveniently from local institutions. Such specimens may not be appropriate for the biomarker research project of interest. Often, biomarkers scientists may need to consider procuring collections from alternate sources. In an effort to expedite examination of

Table 3.1 Pros and cons of utilizing prospective and retrospective biobanking collections.

	Pros	*Cons*
Prospective Sample Collection	Designed to target specific clinical/therapeutical needs	Long time to collect
	Collects all required clinical data	Time-consuming legal/regulatory hurdles
	Addresses intended use	Costly
Retrospective Sample Collection	Collection readily available	May or may not have all required patient data available
	Samples targeted at the therapeutic area in general	Informed consent may be restrictive, or if no informed consent is available: must need the FDA/legal requirements for use as left-over specimens
	Less expensive	
	Addresses the intended use	

sources it is recommended to utilize existing local as well as international scientific connections. Scientists should aim to decrease searching lags *via* utilization of biobank hub portals and locator tools (*e.g.* the B4B hub: www.trans-hit.com/).

- In unison with biobanks, biomarker scientists should aim to define which type of donors, samples and data are needed. This involves definition and provision of inclusion and exclusion criteria, demography (*e.g.* ethnicity, gender, smoking habits, alcohol consumption, *etc.*), clinical requirements (*e.g.* disease of interest, specific patients), as well as specificities regarding type of control groups required (healthy, symptomatic or asymptomatic populations, other diseases groups of interest, *etc.*). It is important to carefully determine the number of patients and their inclusion criteria; well-designed studies are crucial to improve reliability and efficiency of research about biomarkers.[55,56]
- It is also important to define quality requirements for biospecimens and data collected relative to the biomarker of interest. Too many sources of bias in specimen impact results and interpretation of biomarker studies.[56] This includes not only consideration of clinical sampling procedures (*i.e.* application of tourniquet, resting time, posture, level of activity for blood samples) but also biobanking methodologies (mode of cryopreservation per type of sample (plasma *vs.* serum), ischemic window, length of storage, *etc.*). An example of bias is given in Table 3.2.
- Once the project requirements have been defined, due diligence steps should include formal evaluation of each biobank. While a portion of this process may be possible virtually, biobank evaluation may best be conducted *via* onsite audits. On site audits should include confirmation and review of existing SOPs (from sampling to storage and beyond) as well as accessible medical annotation per case. During the audit process, it is important to confirm the range and level of scientific expertise engaged. Efforts should be made to clarify how and in what fashion such expertise has been applied. For example, one biobank may have a satisfactory biospecimen collection in one area of expertise, but not in another one). It is essential to confirm if pathologists and/or clinicians are directly involved and the time points of direct involvement. Such details are of critical importance to ensure quality research and accurate interpretation of data. Once a biobank has been properly evaluated, feasibility planning should involve early education of ethical and regulatory requirements *e.g.* informed consent, ethical committee approval processes and related timelines, confidentiality, specimen and data-acknowledgement requirements, and any other institutional guidelines. One should understand that these requirements may vary per country, per sector, per research domain and in some cases per donor/source population, *e.g.* varying requirements for pediatric populations, postmortem samples, leftover samples, *etc.* Contractual agreements should include unified templates for material-transfer agreements. Intellectual property issues and coauthorship in case of publication should be clarified upfront. These

Table 3.2 Sources and "locations" of bias in marker research. Extracted from Ransohoff and Gourlay.[56]

Source of Bias	*Location of Bias: Before or After Specimens Are Received in the Laboratory*		*Example*
	Before	*After*	
Features of subjects, determined in selection: Age Sex Comorbid conditions Medications	X		Cancer subjects are male, whereas control subjects are mainly female. Bias: Assay results may depend on sex
Specimen collection	X		Cancer specimen come from one clinic, whereas controls come from a different clinic. Bias: Assay results may depend on conditions that differ between clinics
Specimen storage and handling	X	X	Cancer specimens are stored for 10 years because it takes longer to collect them, whereas control specimens are collected and stored over 1 year. Bias: Assay results may vary with duration of storage, or with different number of thaw–freeze cycles.
Specimen analysis		X	Cancer specimens are run on one day, whereas control specimens are run on a different day. Bias: Assay results may depend on day of analysis in a machine that "drifts" over time.

Note: The table shows examples of different sources of bias and the location of the bias before or after specimens are received in the laboratory. The list is not exhaustive: other biases may be important, and the biases listed may or may not be important in any given research study, depending on details of biology and technology (*i.e.* what is being mesured and how it might be influenced).

processes can be quite lengthy. Lags in collaboration tend to be intradependent on the biobank/institution of interest, but in some cases can be 6–18 months. Therefore, these lags should be factored in early on as part of the project planning.

– Akin to pilot testing of collections in progress, it is recommended that feasibility planning involves pilot shipping to confirm that the material-transfer process does not introduce additional variability. Not only will this process allow confirmation of quality requirements for shipping, it will also confirm reliability of timelines and preferred shipping carriers. Biomarker scientists should be aware of international best practices for shipping, which may include IATA guidelines.

– Lastly, it is recommended that whenever data obtained from materials from biobanks are published the study analysis should contain a detailed description of all the confounding parameters that may have influenced the overall results.

3.3.2 Technical and Scientific Recommendations for Biobankers

For the context of this discussion, we regard biobankers as those developing, managing and providing access to biospecimens in a biobanking environment.

Biomarker research requires timely availability of both retrospective and/or prospective collections. An unfortunate reality is that biobanking projects tend to rarely be designed with the planning considerations required to adequately support biomarker research on a broad scale. It would be highly beneficial if biobanking projects were prospectively coordinated internationally with this in mind. Such foresight may reduce the risk of monetary loss and increase return on investment.

It is recommended that biobanks evaluate the potential end products and "fit for purpose" requirements prior to biobank design and setup. Planning considerations, should aim to elucidate the basic and applied research objectives for collections. This would dramatically assist future sustainability of biomarker collaborations. Currently, the practice is quite the opposite. A large majority of biobanks appear to function as "libraries storing books that will never be read". It is estimated that biobanks are only utilizing a relatively small percentage of their biostores. Efficacy would be increased dramatically if biobanks were aware of biomarker-specific research requirements and bilateral planning considerations of biomarker scientists in the academic and industrial sectors. It is necessary that equilibrium exist to support supply and demand. Therefore, it is recommended that biobanks factor in biomarker-planning considerations early on to ensure return on investment of such collaborations and related work products.

Outlined below are primary recommendations, prerequisites and procedures that biobanks should aim to employ in order to optimize the collaborative process and related biomarker research outcomes.

3.3.2.1 Biospecimen Research and Biobank Operational Planning Efforts

– For biomarker research outcomes to have the most relevant impact the following planning efforts should be exercised regularly. Biobank planning efforts to support biomarker research should survey prospective end users to confirm project specific requirements. To remain educated and ensure relevant research projects are being selected, biobanks should also aim to keep abreast of trends in the biomarker research field. Planning should also include areas of market need; this should aim to be a key driver in prospective design of biospecimen and data collections as well as cultivation of historical archives and/or retrospective samples. Therefore,

research budgets should factor in funds for labor involved in investigative R&D efforts.

- Biobanks interested in supporting biomarker research projects should aim to define an efficient pathway to aide timely biomarker collaboration. To decrease lags in research and assure quality and relevance of research outcomes, biobanking regulatory frameworks should factor in issues specific to support high-quality biomarker research. Biobanks should work to educate bioethics committees of the need to work with collaborators in industry as well as academia to ensure samples are being utilized appropriately and in a manner that offers the fullest opportunity to respect the patient's wishes. Such efforts will aim to ensure that the intent to perform research is fulfilled. Informed consent guidelines should include statements that acknowledge that third parties may have access to their information and samples, keeping this aim in mind. Regulatory advisors should be educated to understand that fulfilling market need that is commensurate with and driven by patient and/or disease incidence and prevalence-related needs.
- Biobanks can only function effectively if requisite expertise is available at the appropriate level of commitment. It is possible that the most effective biobanks are those with several dedicated managing scientists and technical teams. Experience has demonstrated that the optimal biobank collections are those tended by a pathologist or clinical chemist with high incidence of well-defined interdisciplinary activities. Career paths for Technical Directors, Biobank Managers as well as professional technical staff should also be established to ensure benchmark-level cultivation becomes a reality. When possible, it is recommended that biobanks engage biomarker experts from industry and academic on both the Scientific Advisory Board and Biospecimen Utilization Committee.

3.3.2.2 Scope of Collection

- It is recommended to keep scope of collection targeted. Biomarker collections may be less sample-size-dependent if they are well planned. Therefore, rather than aiming for broad, large, wide ranging recruitment targets in multiple disease research areas of interest, it is recommended to focus on available expertise.

3.3.2.3 Control-Sample Considerations

- Biobanks should increase the prevalence of control group biosample collection.[57] When developing and/or validating a biomarker it is important that biobanks collaborate prospectively with biomarker scientists to factor in biomarker profiles in control cohorts. Control cohorts may consist of matched healthy subjects but also other patient populations who have demonstrated expression of common profiles of biomarker or the clinical symptomology (*e.g.* pancreatic cancer versus

pancreatitis). Studying other populations of patients may also prove helpful to demonstrate the specificity of a biomarker (*e.g.* pancreatic cancer *versus* colon cancer). While biobanks may plan their collections appropriate to a designated domain of interest, many fail to include necessary control groups.

3.3.2.4 Sample-Set Methodological and Annotative Considerations

- When building a biobank collection for future research it is preferable that a wide range of sample types and formats as feasible per individual/case be procured. Formats vary and can refer to procurement of primary materials, *i.e.* blood, urine and tissue with and without cryopreservative agents, but can also refer to processing formats for primary components (serum, plasma, Buffy coat) as well as secondary derivatives, *e.g.* (DNA, RNA, *etc.*). For blood, the minimum sample set suggested is a plasma and serum sample. Sample-set requirements should be carefully considered for each type of sample and its familial sample set. Source-based requirements should be assessed and factored in methods and materials, then documented in research protocols.
- Definition of biosample collections should be performed in concert with well-established best-practice-based research methodology, procedures and evidence-based protocols. To reduce preanalytical variability for a given biomarker, strict adherence to SOPs and standardized protocols is crucial. Lack of adhesion to simple guidelines and protocols regarding the preanalytical phase will compromise the quality of biospecimen collection. SOP development should mirror recommendations provided by best and evidence-based practices. Several initiatives are attempting to improve the practice of biospecimen management or biobanking. For example, the US National Cancer Institute's Cancer Human Biobank (caHUB) has historically established stringent guidelines to ensure that samples from healthy individuals and cancer patients are collected, annotated, stored and analyzed under standardized conditions and accompanied by appropriate donor medical information. NIH/NCI BBRB recommendations for minimal data sets can be found at: http://biospecimens.cancer.gov/bestpractices/Appendix1.pdf. As these guidelines were created as a broad template, it is important for biobankers to work with biomarker scientists to confirm how such data sets may need to be customized to support biomarker research specifically.
- Baseline procedures should aim to detail methodologies factoring in reporting guidelines early on and be applied during each step along the specimen lifecycle and process chain. For example, this may include instructions for biospecimen sampling. *e.g.* (identification of used collecting tube types, sample volume, *etc.*) along with annotation on sample handling (*e.g.* ''place on ice'' or ''transport at room temperature'')

as well as instructions regarding processing times including specifics (*e.g.* centrifugation specifications, aliquoting process, type of storage vials, *etc.*). For tissues, it is important to collect detailed information on ischemic times: warm ischemia time, cold ischemia time, lag between the collection of a biological sample and processing for testing in the laboratory, type of fixative, fixation protocol, fixation time, size of the tissue specimen, storage temperature and length of storage. Biobanks should ensure that their quality-management system addresses cold chain logistics along the process chain prior to time of collection through time of distribution and beyond. Any deviations should be documented to assist interpretation of downstream analysis.

– Biospecimen data collection should be accompanied by well-annotated clinical data sets. For biomarker research projects to be successful it might be necessary to collect significant clinical outcomes. Collected data should not be limited to solely pathology information. The US NIH has provided initial guidance on factors that may be beneficial to collect. This includes physiologic factors, *e.g.* gender, age, menopausal status, stage in the menstrual cycle, concomitant infection, pregnancy, fasting, exercise, smoking, consumption of alcohol or caffeine-containing beverages, time of day and year, use of a tourniquet, time of rest before sampling, posture (sitting *versus* supine), stress, medication, *etc.* Medical annotation should be complete prior to provision of samples and ideally accessible on a real-time basis when needed. The biobank must have the capability to perform expeditious multicriteria search queries to enable real-time efficient feasibility studies.

3.3.2.5 *Inventory Management, Workflow and Sample Tracking and Temperature Monitoring Practices*

– Informatics tools and software solutions for biospecimen inventory, clinical annotation, sample logistics and downstream analysis management should be coordinated to factor in workflows and annotation relative to biomarker research. A sample laboratory/biospecimen inventory management system (LIMS or BIMS) should be capable of recording and reporting on the following details on a sample-by-sample basis: biospecimen management life-cycle history, length and mode of storage, shipping and sample temperature, sample integrity and stability at time of sampling and handling intervals and long-term stability. In order to facilitate tracking of temperature and variation of individual sample stability, biobank stakeholders should aim to consider proactive implementation of temperature-sensing and -tracking techniques *e.g.* microelectromechancial systems or "MEMs"-based technologies *i.e.* those being in development by Bluechiip™.[58]
– The ideal biobank should offer an electronic sample inventory management system that eliminates paper-based tracking and reporting processes through all stages of a sample life cycle. Ideally, the system

should have a reasonable level of interoperability with both internal and external collaborators, either through a common biospecimen inventory module as well as facile methods of exchanging and examining data. Best practices for biobank inventory management have been initially offered by the (NIH/NCI BBRB Best Practices (http://biospecimens.cancer.gov/bestpractices/to/bri.asp)). Biobanks should keep abreast of updates in best practices relevant to inventory management as well as guidance on annotation specific to biomarker research. Biobanks and biobanking product vendors should plan accordingly to ensure inventory management systems factor in workflows specific to biomarker research.

3.3.2.6 Sample-Quality Management, Assurance and Auditing

Quality practices should include established data collection procedures to promote standardized data collection. Data-collection practices should document the mode and procedures involved with data collection. Clarifications should differentiate data collection characteristics, *i.e.* electronic *versus* manual, single *versus* double capture, data cleaning and level of completeness). Logistical details should be captured, *i.e.* capability and time frame required to retrieve the latest data from the health-care system, anonymization procedure, and ability and steps needed to recontact donors for additional information and/or samples. All of this information should be presented to guide the feasibility review of biomarker projects.

Ideally, biobanks should aim to establish quality standards, *e.g.* quality decay indicators. Quality decay indicators, if available, may assess molecular integrity of the biosample and ascertain if biosample still represents the *in vivo* situation confirming that the sample is still valid for assays and reliable interpretation of results (Refer to guidelines in Section 3.3.3).

- Biobanks should aim to perform regular internal audits to assess the quality of biospecimens and data as it is relevant to biomarker research. Yearly external audits are also recommended. Quality assessments should include quality analysis of specimens as well as a review of the data collected. This may include but not be limited to details, *e.g.* adherence to tracking of sampling, sample management and storage procedures (''5W rule'': to know when, who, what, why, where), the quality of related medical annotations (kind of data being retrieved, data consistency per case). Throughout this process the question of interest should confirm that both specimens and data are sufficient to support successful biomarker research outcomes.
- Biobanks interested in supporting biomarker research for long-term projects should aim to track efficiency measures. Such measures should be provided at key iterative points in project development towards guiding biomarker scientists in evaluating candidate biobanks. Outcome measures might include review of sample utilization, *e.g.* number of biospecimen used over a definedperiod: ratio used/stored, internal versus external use, types of projects performed and related outcomes (number of publications

and relevant citations, number of biomarker studies resulting in successful proofs of concept, commercial end-products, *etc.*). Hofman *et al.*[59] recently proposed four major categories of indicators performance indicators to be used for the evaluation of a biobank: quality, activity, scientific production and visibility.

3.3.3 Joint Recommendations for Biomarker Scientists and Biobankers

To progress the maturation of biomarkers from bench to bedside, it is imperative that those with expertise related to biomarker discovery, validation, development and implementation work tightly together to build a successful pipeline. The biobankers, on the one hand, need to interact with a variety of biomarker scientists, both on a strategic level and on a practical level, to implement the recommendations outlined above. A few recommendations are presented below.

3.3.3.1 Content of Biospecimen Collection

The quality, relevance and content of a biobank determine its utility. Importantly, a tight interaction between the collectors and the users of biospecimen (respectively, the biobankers and the biomarker scientists), will ensure that the content of a biobank is used to maximal extent for biomarker research and development. There could be considerable differences between biomarker scientists in their interests in clinical biospecimen. Some may just need clinical samples to confirm their preclinical findings in cellular or animal studies; others may focus on clinical impact and need samples to perform large-scale epidemiology through genetic analysis of selected genes in defined population groups. Through strategic alignment, the biomarker scientists can enhance definition of what biospecimen collections are collected, including the necessary matched control samples, thus tailoring the biobank collection to more aptly support biomarker research requirements.

3.3.3.2 Collection of Biospecimens

In various ways, the biomarker scientists mutually complement biobankers in collection and management of biospecimens. Through their functional and often global network of scientific interactions, a wide array of clinical expert centers may be approached to share clinical samples and thus enrich the collection of biospecimen from different origins and increase the intrinsic value of the biobank. The biobankers should indicate the gaps in their collections, thereby prompting the scientists to look for specific clinical samples. Interesting collection of samples can be obtained in clinical trials when the effect of pharmaceutical drugs is being tested. Particularly when serial samples are obtained from the same subject, drug effects before and after administration

can be determined. Sampling from dosed healthy volunteers will potentially enable a study of mechanistic or toxic effects of a drug in a nondiseased state, or the validation of biomarkers representing those effects. Sampling in patients, however, will enable more clinically relevant studies on the mechanism of action of the drug.

Through coordinated interaction, biobankers can advise scientists how to design the clinical trials and ensure the storage of high-quality samples that can be used for predefined use in a later stage. It is important that biomarker scientists and biobankers collaborate closely on evaluation and cultivation of biomarker research utilizing biospecimen collections. In unison, biobanks and biomarker scientist should aim to work together to prospectively identify, define, elucidate and benchmark, quantify discrete outcomes and parameters for biomarker biospecimen analysis and quality review. Each stakeholder can assist the other in elucidating each other's needs and requirements, as well as success factors for increased impact of biomarker research outcomes.

3.3.3.3 Quality Control

A recurring issue with biospecimens stored in biobanks is how to ensure the quality of the sample. Through their technology platforms, biomarker scientists can support such analysis by performing a thorough analysis of control samples at regular intervals. For instance, an LC/LC-MS/MS proteomics scan of a body fluid will typically reveal 1000–5000 peptide peaks, depending on the complexity of the sample and analysis method used.[60,61] By recurrent scanning of the control samples with a fixed standard protocol at predefined times, the scientist can determine whether the abundance of some components of the proteome has changed. The subsequent identification of such components will learn whether there has been a loss of sample integrity, a protein aggregation or other events that has changed the sample over time. Recently, we have applied this method to maintain the stability of cerebrospinal fluid, demonstrating the use of this body fluid for biomarker research.[62,63]

3.3.3.4 Biomarker Research

Any biomarker research project that aims at clinical relevance and impact has to have access to the best-suited clinical samples to generate the proper results. In a powerful combination, biomarker scientists can work together with biobankers to ensure that such samples are available at the start of a joint research project. Our experience has shown that often clinicians are enthusiastic to participate in clinical biomarker projects, but despite this, may not be able to collect the required number of samples by their own. A biobank can coordinate the engagement of several of such clinicians, ensure the quality of the samples in a similar manner and thus make sure the required number for the study is met.

Both biobanks and biomarker scientists should aim to evaluate unique funding opportunities to ensure sustainability of biomarker research. Interesting opportunities arise by national and international funding of biomarker

projects, such as NIH or EU-FP7 grants. Such funding often has a more long-term objective to innovate disease management in a particular area, and aims to bring together multiple scientists, clinicians and biobankers together under a functional umbrella. Creative models for collaboration can be beneficial if implemented. Moreover, increasingly such funding is through public–private partnerships, strengthening the synergy between pharmaceutical, biotech and diagnostic industry on one hand, and academic scientists and biobankers on the other hand. Research funded through such grants almost always has the objective of application in real life, and given the nature of many biomarker projects, application in clinical care is greatly stimulated through this principle. A variety of collaboration business models including the risk–benefit model should be considered on promoting sustainability.

3.3.3.5 Crossfertilization of Best Practice

Finally, and perhaps as important as the findings made by biomarker scientists using good samples from biobanks, there is a mutual growth in experience among the participants and stakeholders in the importance of high-quality clinical samples. From biomarker projects and sample handling, there will be a growing sense of awareness of how to collect biospecimen and to use them in biomarker research and development studies. This will result in definition of best practice that, upon publication and communication, will steer the biomedical field towards the optimal positioning of synergy between biobanks and biomarker research. Potentially, such best practice can also drive the development of reporting standards for evidence-based-practice and clinical care, *e.g.* Biospecimen Reporting for Improved Study Quality (BRISQ) and coding systems, *i.e.* SPREC or referred to as "Standard Preanalytical;[64] these variables should be jointly tracked when relevant). Recommendations outlined in this chapter may function as early best practice for biomarker–biobanking collaborations.

3.4 Utilization of Biobank Samples for Biomarker Discovery and Development

Over the past ten years there has been an accelerated growth in the biobanking field as well as in the biomarker domain. We next review aspects of recent technical developments that utilize clinical biosamples to support the discovery and development of biomarkers.

3.4.1 Biomarker Discovery

There are in general two ways in which clinical biomarkers are being discovered. First, molecular studies of cellular or animal model systems can yield candidate biomarkers that are subsequently validated in human samples of clinical relevance. Such an approach is strongly driven by the fact that

preclinical model systems are more accessible and readily treatable with pharmaceutical interventions. However, in many cases there has been limited translation between preclinical and clinical behavior of the biomarker, including its level of abundance, modification, kinetics and interactions with other macromolecules[65] Alternatively, human material is used directly to identify biomarkers which circumvent the cross-species translational step. This optimal approach has led to several clinically applicable discoveries, including the increasing list of driver mutations underlying aggressive tumor development that have become the basis of personalized cancer medicine[66] Several discovery technologies have had requirements for minimal sample quality or sample amounts with the result that a study cannot be designed properly or cannot be performed at all but recent developments have improved this situation.

3.4.1.1 Sample Quality

Many biomarker studies did not include analysis of the quality of the samples and may have resulted in differential biomarkers that at least partly can be related to the sample isolation and storage process (our unpublished observations). As outlined above, the integrity of the stored samples themselves can already be improved by standardizing sample handling procedures and by including wide molecular profiling of control samples. In addition, the multiplex nature of targeted biomarker analysis methods such as proteomic/metabolomic mass spectrometry using Multiple Reaction Monitoring can now include decay indicators to determine the quality of the biosamples within the same study, which is a great advantage. Also, method developments have made formalin-fixed-paraformaldehyde-embedded (FFPE) tissue samples accessible to biomarker discovery and validation. FFPE is generally regarde as a stable preservation method, but for long the crosslinking during sample preparation has disabled the identification of novel biomarkers through molecular profiling. However, recently methods were developed to reverse the crosslinks and make fixed tissue samples amenable to hybridization-based transcriptomics profiling[67] and even to the more challenging mass-spectrometry-based proteomics.[68] In particular, in combination with isolation of histologically defined areas through laser-capture microdissection,[69] this enables enormous possibilities for retrospective molecular assessment of stored fixed samples.

3.4.1.2 Sample Amounts

Several high-end technology platforms increased their sensitivity and robustness, thereby lowering the minimally required amounts for biomarker discovery. For instance, the next-generation sequencing platforms have developed with increasing robustness and reduced cost, such that a whole genome scan is feasible for many biomarker laboratories. Depending on the research question, the entire population or subfractions of cellular DNA and RNA can be analyzed to identify variants in sequence, abundance and/or

modification, including those in genomic DNA, methylated DNA, transcribed mRNA and miRNA.[70] Whereas complete genome sequencing requires substantial material, the sequencing of subpopulations can now be done with minimal amounts that is first amplified under well-controlled conditions. As such, biosamples from patients can be subjected to whole exome sequencing in a routine clinical setting to identify Mendelian genetic predisposition to disease.[71] Alternatively, novel hybridization-based microarray platforms have been developed that use only minimal amounts of material to profile DNA/RNA variants with a large variety of probes, with a similar performance as amplification-based expression profiling methods.[72,73] Recent improvements in mass spectrometry, the key technology to focus on protein, peptide and metabolite biomarkers, have increased the sensitivity of biomarker detection and consequently the minimal amounts needed for such studies.[74] For example, only 20 μl of a cerebrospinal fluid sample was sufficient to yield robust data (coefficient of variation $<7\%$) to identify potential peptide biomarkers for multiple sclerosis.[75] It can be expected that further technological developments, particularly in combination with microfluidics[76] can further decrease the sample amounts needed for biomarker analysis.

3.4.2 Biomarker Development

3.4.2.1 Biomarker Validation

To progress a biomarker from discovery to development, it is imperative to do a thorough biomarker validation in relevant clinical samples to confirm the results obtained in the initial study. Three important components strongly depend on each other in biomarker validation: the assay, the samples and the test design. The biomarker assay needs to be developed and applied following a fit-for-purpose principle, whereby limited robustness is needed in this stage of biomarker development. Clinical samples need to be selected to support the objective of the biomarker validation, need to be available in sufficient numbers and quality, and preferably be derived from different clinical sites to verify that the biomarker is specific to the disease or mechanism and not to a certain source. The biomarker test design should ensure that a sufficient number of biosamples is tested, whereby information on the technical and biological variation of the biomarker is to be obtained. Optimally, biomarker validation is to be done by several independent laboratories using shared protocols and aliquots of high-quality biosamples, ensuring that the obtained data can be compared in a meta-analysis across the participating laboratories.[77] Recently, the key biomarkers for Alzheimer's disease Aβ42, T-tau, P-Tau have been validated in such a multicenter testing, defining a best-practice example.[78]

Emerging technologies for biomarker validation tend to follow a multiplex approach, in which the biomarker of interest is compared with other biomarkers within the same assay, thus demonstrating the added value of the novel biomarker, and at the same time determine the quality of the sample through analysis of decay markers. The use of next-generation sequencing and micro-arrays to study DNA/RNA molecules intrinsically have this advantage,[70]

whereas smaller-scale multiplex assays using QPCR or branched-DNA probes have become standard practice. Popular methods to analyze protein biomarkers use an immunoassay format that is bead-based,[79] microtiter plate-based[80] or planar microarray-based.[81] Of particular interest is the use of targeted mass spectrometry in which up to 100 specific analytes of proteomic or metabolomic nature are quantitated per analysis with a sensitivity that approaches that of immunoassays.[82,83] Recently, a multicenter study with eight participating laboratories showed that such targeted mass-spectrometry assays can generate highly reproducible and precise results both within and across laboratories and instrument platforms.[84]

3.4.2.2 Biomarker Test Validation

Once the biomarker is confirmed by an independent validation study, the biomarker test can then be developed for clinical application. Biomarker development will now aim to achieve increased robustness, with a focus on obtaining maximum specificity and selectivity, whilst optimizing the ease and costs in production of the final format of the test. During this process, a selected sample set will be regularly tested to verify the performance of the test. Once developed to the final format, the biomarker test will again be scrutinized using a thorough analysis of independent clinical samples, thus generating the final biomarker test performance results that will be used for marketing. Again, high-quality biosamples are imperative to be successful in this phase.

A relevant case study of clinical biomarker discovery and development that outlines the various steps mentioned above is the Mammaprint, a microarray-based biomarker test for prediction of breast-cancer recurrence (see www.mammaprint.nl). The biomarker test is based on the original discovery of a 70-gene transcript signature, identified in 98 retrospectively collected primary breast-tumor biopsies, and that distinguished patients that after five years were disease free or developed distant metastasis.[85] The selected probes enabling analysis of this gene signature were converted into a multiplex microarray test,[86] and prospectively validated in the dedicated prospective MINDACT trial.[87,88] Further validation and testing of the 70-gene Mammaprint signature indicated additional value as predictor of therapeutic efficacy in the poor outcome predicted group,[89] whereas further bioinformatics research revealed how the 70 components of the signature are related to cancer biology.[90] Finally, it was recently shown that the gene expression assay can be performed both on fresh frozen tissue as on FFPE tissue with similar results, widening the application of this approach.[91]

3.4.3 Biobanking–Biomarker Collaborations

Considering the lack of quality and standardization between biobanks and the impact on biomarker validation, one may ask if too many specimens have been collected and stored as well as if too much (public) money has been invested.

Actually, most biobanks struggle to find a suitable, sustainable business model. Moreover, looking at the success stories about biobanks but also looking at the too high numbers of biomarkers that are dying on the shelves or are not efficiently validated by lack of funds it is clear that a new business model is needed.

An interesting option is a risk–benefit sharing model where biobanks and biomarker developers will benefit.[92] It would provide biobanks with an opportunity to become a true partner with a scientific input and interest into a biomarker project, and allows the biobank to create a long-term revenue stream to establish a sustainable business. At the same time, the risk–benefit sharing model would help decrease the up-front financial requirements for the scientists developing the biomarker. Under a risk–benefit-sharing model biobanks provide access to their biospecimen collection at a substantially reduced cost, in return for a share in the revenues from a sale of the biomarker intellectual property, or royalties on the sale of the final diagnostic test. The biobank would become a partner in the biomarker research team, and be directly involved in the development program. Many biobanks have scientific specialists in their specific disease domains associated with them, who can provide valuable input into a biomarker development program, as well as access to specific patients for further testing and validation work. Such a working model will enhance collaboration between biobanks and biomarkers developers and will certainly accelerate biomarker discovery and validation.[92]

A construction to share biospecimen is particularly useful in biomarker projects by consortia funded through public–private partnerships in which multiple partners, both from industry and academia, collaborate to reach objectives that would not be achievable by the single partners alone.[93] In particular, clinical biomarker validation and development projects require multiple collaborating partners whose joint objective strongly depends on the availability of well-annotated and high-quality biosamples. Through their biosamples, biobanks can contribute well to the knowledge and output developed by the consortium, whereas they can share in potential revenue of the generated biomarker assays, creating a mutual benefit for the biobank and other partners.

Due to the large amount of public funding, it would be timely to consider coordinating a global strategy for biobanking that aligns the content and expansion of biospecimen collections with the healthcare needs in term of biomarkers. Currently, we find that too many biobanks work independently and collect the same types of biosamples, whereas it can be nearly impossible to find collections with sufficient size and content for development and validation of specific biomarkers. A global independent organization that coordinates the expansion of biobank content will avert redundancy in biobanking and decrease costs, while increasing return on investment. Equilibrium between offer and demand should exist.

On the other hand, having access to high numbers of high-quality biospecimens collected during clinical trials with drugs that failed in development

could greatly contribute to the validation of candidate biomarkers, particularly new companion or prognostic biomarkers. This source of clinical biosamples is often not accessible due to restrictions in the IRB-approved clinical protocols or the informed consent signed by the participants. Even if IRB approval and informed consent would allow such use, the sponsor who initiated these clinical trials will have to agree to make biospecimen available to third parties for biomarker research.

It is timely that all players in the biobanking field, being biobankers, biomarker scientists, research organizations, pharmaceutical and diagnostic industrial parties, funding institutions, regulatory and ethical authorities, and governments, work together in a global concerted effort to increase the amount of definitive high quality biosamples available for biomarker research and development. Ultimately, this approach could lead to the establishment of a consolidated international registry tool referencing all gold-standard biobanks aimed at expedition of biomarkers in health care.

References

1. T. Newman and J. Freitag, *Applied Clinical Trials*, July 1, 2011.
2. National Cancer Institute. Biorepositories and Biospecimen Research Branch (BBRB) Best Practices, http://biospecimens.cancer.gov/practices/2011bp.asp. 2011.
3. ''2012 ISBER Best Practices for Repositories: Collection, Storage, Retrieval and Distribution of Biological Materials for Research.'' *Biopreservation and Biobanking*, 2012, **10**(2), 79–161.
4. The Task Force on Biological Resource Centres. ''OECD Best Practice Guidelines for Biological Resource Centres.'' Organization for Economic Cooperation and Development (OECD), 2007, 3–113.
5. Common Minimum Standards and Protocols for Biological Resource Centers Dedicated to Cancer Research''). ''World Health Organization International Agency for Research on Cancer IARC Working Group Reports, Volume 2, 2007, http://www.iarc.fr/en/publications/pdfs-online/wrk/wrk2/index.php.
6. L. M. McShane, D. G. Altman, W. Sauerbrei, S. E. Taube, M. Gion and G. M. Clark, Statistics Subcommittee of NCI-EORTC Working Group on Cancer Diagnostics (2006) Reporting recommendations for tumor marker prognostic studies (REMARK), *Breast Cancer Res. Treat.*, 2006, **100**(2), 229–235.
7. http://www.stard-statement.org.
8. http://www.strobe-statement.org.
9. H. M. Moore, A. B. Kelly, S. D. Jewell, L. M. McShane, D. P. Clark, R. Greenspan, D. F. Hayes, P. Hainaut, P. Kim, E. A. Mansfield, O. Potapova, P. Riegman, Y. Rubinstein, E. Seijo, S. Somiari, P. Watson, H. U. Weier, C. Zhuand and J. J. Vaught, Biospecimen reporting for improved study quality (BRISQ), *Cancer Cytopathol.*, 2011, **119**(2), 92–102.

10. L. Miranda "Obtaining and maintaining high-quality tissue samples: scientific and technical considerations to promote Evidence-Based Biobanking Practice (EBBP)", In: M. Wigglesworth, *Management of Chemical and Biological Samples for Screening Applications*. Wiley, 2012, 59–79.
11. A. Cambon-Thomsen and G. A. Thorisson, L. Mabile for the BRIF working group, The role of a bioresource research impact factor as an incentive to share human bioresources, *Nature Genetics*, 2011, **43**, 6.
12. Manuel M. Morente, Laura Cereceda, Francisco Luna-Crespo and Maria J. Artiga, Managing a Biobank Network, *Biopreservation and Biobanking*, 2011, **9**(2), 187–190.
13. F. Betsou, D. L. Rimm, P. H. Watson, C. Womack, A. Hubel, R. Coleman, L. Horn, S. F. Terry, N. Zeps, B. J. Clark, L. B. Miranda, R. E. Hewitt and G. D. Elliott, *Biopreservation and Biobanking*, 2010, **8**(2), 81–88.
14. J. Vaught and N. C. Lockhart, *Clin. Chim. Acta*, 2012, **413**(19–20), 1569–1575.
15. R. Dash, J. Robb, D. Booker, W. Foo, D. Witte and L. Bry, *Arch Pathol Lab Med.*, 2012, **136**, 668–678.
16. D. N. Mayol-Heath, A. S. Keck and P. Woo, *Biopreservation and Biobanking*, 2011, **9**(3), 253–257.
17. S. Pathmasiri, M. Deschenes, Y. Joly, T. Mrejen, F. Hemmings and B. M. Knoppers, *Nature Biotechnology*, 2011, **29**(4), 319–323.
18. D. N. Mayol-Heath, A. S. Keck and P. Woo, *Biopreservation and Biobanking Biorepository*, 2010, 139–145.
19. K. Lindpainter, *Nature*, 2011, **470**, 175.
20. G. Poste, *Nature*, 2011, **149**, 156.
21. D. Simeon-Dubach and A. Perren, *Nature*, 2011, **475**, 454–455.
22. F. Prinz, T. Schlange and K. Asadullah, *Nature Reviews Drug Discovery*, 2011, **10**, 712.
23. L. Osherovich, *SciBX April*, 2011, **4**(15).
24. E. P. Diamandis, *J Natl Cancer Inst*, 2010, **102**, 1462–1467.
25. S. Silberman, *Wired Magazine*, (http://www.wired.com/magazine/2010/05/ff_biobanks/all/1), June 2010.
26. E. Foot, The three C"s for biomarker development, *Scrip*, 2011.
27. R. Balasubramanian, L. Muller, K. Kugler, W. Hackl, W. L. Pleyer, M. Dehmer and A. Graber A, *Biomarkers*, 2010, **15**(8), 677–683.
28. K. G. Kugler, W. O. Hackl, L. A. Mueller, H. Fiegl, A. Graber and R. M. Pfeiffer, *Journal of Clinical Bioinformatics*, 2011, **1**, 9.
29. G. Lippi, *Clin. Chem. Lab. Med.*, 2011, **49**(7), 1113–1126.
30. M. Dresse, D. Nagel, E. M. Ganser, G. Davis, B. Dowell, R. Doss and P. Stieber, *Tumour Biol.*, 2008, **29**, 35–40.
31. A. Rai, C. Gelfand, B. Haywood, D. Warunek, J. Yi, M. Schuchard, R. Mehigh, S. Cockrill, G. Scott, H. Tammen, P. Schulz-Knappe, D. Speicher, F. Vitzthum, B. Haab, G. Siest and D. Chan, *Proteomics*, 2005, **5**(13), 3262–3277.
32. H. Tammen, *Mol. Biol.*, 2008, **428**, 35–42.

33. A.-B. Halim *Drug Discovery and Development – Present and Future* ed. M. Izet Kapetanovic Publisher: InTech, 2011, Chapter 18, 401–424.
34. K. Bensalah, K. F. Montorsi and S. Shariat, *Eur. Urol.*, 2007, **52**, 601–1609.
35. R. J. Goodwin, A. M. Lang, H. Allingham, M. Boren and A. R. Pitt, *Proteomics*, 2010, **10**, 1751–1761.
36. M. E. Hammond, W. Rees, T. Belnap, B. Rowley, S. Catmull and W. Sause, *Arch Archives of Pathology & Laboratory Medicine*, 2010, **134**(4), 606–612.
37. C. Mathay, W. Yan, R. Chuaqui, A. Skubitz, J.-P. Jeon, N. Fall, F. Betsou and M. Barnes, *Biopreservation and Biobanking*, 2012, **10**(6), 532–542.
38. S. Unhale, A. Skubitz, R. Solomon and A. Hubel, *Biopreservation and Biobanking*, 2012, **10**(6), 493–500.
39. T. Steiniche, T. B. Vainer, M. Franzmann, R. Hagemann-Madsen and M. Bak, *Ugeskr Laeger*, 2008, **170**, 1050.
40. G. Hicks and B. F. Boyce, *Biotechnic & Histochemistry*, 2012, **87**(1), 14–17.
41. C. Marchio, M. Dowsett and J. Reis–Filho, *BMC Medicine*, 2011, **9**, 41.
42. I. Pinhel, F. Macneill, M. Hills, J. Salter, S. Detre, R. A'hern, A. Nerurkar, P. Osin, I. Smith and M. Dowsett, *Breast Cancer Res.*, 2010, **12**(5), R76.
43. V. Espina, C. Mueller, K. Edmiston, M. Sciro, E. F. Petricoin and L. A. Liotta, *Proteome Clin. Appl.*, 2009, **3**, 874–882.
44. S. M. Hewitt, F. A. Lewis, Y. Cao, R. C. Conrad, M. Cronin, K. D. Danenberg, T. J. Goralski, J. P. Langmore, R. G. Raja, P. M. Williams, J. F. Palma and J. A. Warrington, *Arch. Pathol. Lab. Med.*, 2008, **132**, 1929–1935.
45. Y. Miyatake, H. Ikeda, R. Michimata, S. Koizumi, A. Ishizu, N. Nishimura and T. Yoshiki, *Exp. Mol. Pathol.*, 2004, **77**, 222–230.
46. K. Sköld, M. Svensson, M. Norrman, B. Sjögren, P. Svenningsson and P. Andrén, *Proteomics*, 2007, **7**, 4445–4456.
47. A. Spruessel, G. Steimann, M. Jung, S. Lee, T. Carr, A.–K. Fentz, J. Spangenberg1, C. Zornig, F. Juhl and K. David, *Biotechniques*, 2004, **36**, 1030–1037.
48. F. Nkoy, M. Hammond, W. Rees, T. Belnap, B. Rowley, S. Catmull and W. Sause, *Arch. Pathol. Lab. Med.*, 2010, **134**, 606–612.
49. T. Khoury, T. S. Sait, H. Hwang, G. Chandrasekhar, D. Wilding and S. Ta, *Kulkarni.Mod. Pathol*, 2009, **22**, 1457–1467.
50. *American Society of Clinical Oncology: Journal of clinical Oncology*, 2010, **28**(16), 2784–2795.
51. J. Lengellé and E. Panopoulos, *F.Betsou Cytokine*, 2008, **44**(2), 275–282.
52. J. Li, C. Kil, K. Considine, B. Smarkucki, M. Stankewich, B. Balgley and A. Vortmeyer, *Laboratory Investigation*, 2012, 1–12.
53. J. Ioannidis and O. Panagiotou, *JAMA*, 2011, **305**(21), 2200–2210.
54. D. Simeon-Dubach, A. Burt and P. Hall, *J. Pathol.*, 2012, **228**, 431–433.
55. D. Ransohoff, *Journal of Clinical Epidemiology*, 2007, **60**, 1205–1219.
56. D. Ransohoff and M. Gourlay, *Journal of Clinical Oncology*, 2010, **28**(4), 698–704.

57. J. Carpenter, H. Moore, H. Juhl, G. Thomas and L. Miranda, *Biopreservation and Biobanking*, 2011, **9**(2), 129–131.
58. L. B. Miranda, K. Wyatt, I. Johnston, M. Milljanic, J. Chaffey, ''Proof of concept'' pilot study: Bioprocess chain of custody and bioresource sample management temperature observations. Sample level temperature trends and stability data obtained via utilization of bluechiip® temperature tracking technology. *Biopreservation and Biobanking*, 2013, **11**(2), in press.
59. V. Hofman, M.-C. Gaziello, C. Bonnetaud, M. Ilie, V. Mauro, E. Long, E. Selva, V. Gavric-Tanga, S. Lassalle, C. Butori, C. Papin-Michaud, N. Lerda, O. Bordone, C. Coelle, J-C. Sabourin, C. Chabannon and P. Hofman, *Annales de Pathologie*, 2012, **32**, 91–101.
60. M. Mann and N. L. Kelleher, *Proc. Natl. Acad. Sci. U.S.A*, 2008, **105**(47), 18132–18138.
61. J. C. Tran, L. Zamdborg, J. E. Lee, A. D. Catherman, K. R. Durbin, J. D. Tipton, A. Vellaichamy, J. F. Kellie, M. Li, C. Wu, S. M. Sweet, B. P. Early, N. Siuti N, R. D. LeDuc, PD. Compton, P. M. Thomas and N. L. Kelleher, *Nature*, 2011, **480**(7376), 254–258.
62. M. P. Stoop, L. Coulier, T. Rosenling, S. Shi, S. A. M. Smolinska, L. Buydens, K. Ampt, C. Stingl, A. Dane, B. Muilwijk, R. L. Luitwieler, P. A. Sillevis Smitt, R. Q. Hintzen, R. Bischoff, S. S. Wijmenga, T. Hankemeier, A. J. van Gool and T. M. Luider, *Mol. Cell Proteomics*, 2010, **9**(9), 2063–2075.
63. T. Rosenling, M. P. Stoop, A. Smolinska, B. Muilwijk, L. Coulier, S. Shi, A. Dane, C. Christin, F. Suits, P. L. Horvatovich, S.S. Wijmenga, L. M. Buydens, R. Vreeken, A.J. van Gool, T. M. Luider and R. Bischoff, *Clin Chem*., 2011, **57**(12), 1703–1711.
64. F. Betsou, S. Lehmann, G. Ashton, M. Barnes, E. Benson, D. Coppolo, Y. DeSouze, J. Eliason, B. Glazer, F. Guadagni, K. Harding, D. J. Horsfall, C. Kleeberger, U. Nanni, A. Prasad, K. Shea, A. Skubitz, S. Somiari, E. Gunter, and International Society for Biological and Environmental Repositories (ISBER) Working Group on Biospecimen Science, *Cancer Epidemiology Biomarkers and Prevention*, 2010, **19**(4), 1004–1011.
65. A. J. van Gool, B. Henry and E. D. Sprengers, *Drug Discov.Today*, 2010, **15**(3–4), 121–126.
66. J. E. Dancey, P. L. Bedard, N. Onetto and T. J. Hudson, *Cell*, 2012, **148**(3), 409–420.
67. L. Turner, J. D. Heath and N. Kurn, *Methods Mol. Biol*., 2011, **724**, 269–280.
68. O. Azimzadeh, Z. Barjaktarovic, M. Aubele, J. Calzada-Wack, H. Sarioglu, M. J. Atkinson and S. Tapio, *J Proteome Res*., 2010, **9**(9), 4710–4720.
69. B. L. Hood, M. M. Darfler, T. G. Guiel, B. Furusato, D. A. Lucas, B. R. Ringeisen, I. A. Sesterhenn, T. P. Conrads, T. D. Veenstra and D. B. Krizman, *Mol. Cell Proteomics*, 2005, **4**(11), 1741–1753.
70. M. L. Metzker, *Nat. Rev. Genet*., 2010, **11**(1), 31–46.
71. C. Gilissen, A. Hoischen, H. G. Brunner and J. A. Veltman, *Eur. J Hum Genet*., 2012, **20**(5), 490–497.

72. G. K. Geiss, R. E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D. L. Dunaway, H. P. Fell, W. S. Ferree, R. D. George, T. Grogan, J. J. James, M. Maysuria, J. D. Mitton, P. Oliveri, J. L. Osborn, T. PengT, A. L. Ratcliffe, P. J. Webster, E. H. Davidson, L. Hood and K. Dimitrov, *Nat. Biotechnol.*, 2008, **26**(3), 317–325.
73. R. D. Canales, Y. Luo, B. Willey, C. C. Austermiller, C. Barbacioru, K. Boysen, R. V. Hunkapiller, C. R. Jensen, K. Y. Knight, Y. Lee, B. Ma, A. Maqsodi, E. H. Papallo, K. Peters, P. L. Poulter, R. R. Ruppel, L. Samaha, W. Shi, L. Yang and F. M. Zhang, *Goodsaid. Nat Biotechnol.*, 2006, **24**(9), 1115–1122.
74. P. Mallick and B. Kuster, *Nat. Biotechnol.*, 2010, **28**(7), 695–709.
75. M. P. Stoop, L. J. Dekker, M. K. Titulaer, R. J. Lamers, P. C. Burgers, P. A. Sillevis Smitt, A. J. van Gool, T. M. Luider and R. Q. Hintzen, *J. Proteome Res.*, 2009, **8**(3), 1404–1414.
76. R. Seigneuric, L. Markey, D. S. Nuyten, C. Dubernet, C. T. Evelo, E. Finot and C. Garrido, *Curr. Mol. Med.*, 2010, **10**(7), 640–652.
77. H. Mischak, G. Allmaier, R. Apweiler, T. Attwood, M. Baumann, A. Benigni, S. E. Bennett, R. Bischoff, E. Bongcam-Rudloff, G. Capasso, J. J. Coon, P. D'Haese, A. F. Dominiczak, M. Dakna, H. Dihazi, J. H. Ehrich, P. Fernandez-Llama, D. Fliser, J. Frokiaer, J. Garin, M. Girolami, W. S. Hancock, M. Haubitz, D. Hochstrasser, R. R. Holman, J. P. Ioannidis, J. Jankowski, R. A. Julian, J. B. Klein, W. Kolch, T. Luider, Z. Massy, W. B. Mattes, F. Molina, B. Monsarrat, J. Novak, K. Peter, P. Rossing, M. Sánchez-Carbayo, J. P. SchanstrA, O. J. Semmes, G. Spasovski, D. Theodorescu, V. Thongboonkerd, R. Vanholder, T. D. Veenstra, E. Weissinger, T. Yamamoto and A. Vlahou, *Sci. Transl. Med*, 2010, **2**(46), 1–6.
78. N. Mattsson, U. Andreasson, S. Persson, H. Arai, S. D. Batish, S. Bernardini, L. Bocchio-Chiavetto, M. A. Blankenstein, M. C. Carrillo, S. Chalbot, E. Coart, D. Chiasserini, N. Cutler, G. Dahlfors, S. Duller, A. M. Fagan, O. Forlenza, G. B. Frisoni, D. Galasko, D. Galimberti, H. Hampel, A. Handberg, M. T. Heneka, A. Z. Herskovits, S. K. Herukka, D. M. Holtzman, C. Humpel, B. T. Hyman, K. Iqbal, M. Jucker, S. A. Kaeser, E. Kaiser, E. Kapaki, D. Kidd, P. Klivenyi, C. S. Knudsen, M. P. Kummer, J. Lui, A. Lladó, P. Lewczuk, Q. X. Li, R. Martins, C. Masters, J. McAuliffe, M. Otto, G. P. Paraskevas, L. Parnetti, R. C. Petersen, D. Prvulovic, H. P. de Reus, R. A. Rissman, E. Scarpini, A. Stefani, H. Soininen, J. Schröder, L. M. Shaw, A. Skinningsrud, B. Skrogstad, A. Spreer, L. Talib, C. Teunissen, J. Q. Trojanowski, H. Tumani, R. M. Umek, B. Van Broeck, H. Vanderstichele, L. Vecsei, M. M. Verbeek, M. Windisch, J. Zhang, H. Zetterberg and K. Blennow, *Alzheimers Dement.*, 2011, **7**(4), 386–395.
79. R. J. Fulton, R. L. McDade, P. L. Smith, L. J. Kienker and J. R. Kettman, Jr, *Clin. Chem.*, 1997, **43**(9), 1749–1756.
80. N. Mattsson, M. Axelsson, S. Haghighi, C. Malmeström, G. Wu, R. Anckarsäter, S. Sankaranarayanan, U. Andreasson, S. Fredrikson,

A. Gundersen, L. Johnsen, T. Fladby, A. Tarkowski, E. Trysberg, A. Wallin, H. Anckarsäter, J. Lycke, O. Andersen, A. J. Simon, K. Blennow and H. Zetterberg, *Mult Scler*., 2009, **15**(4), 448–454.
81. M. Hartmann, J. Roeraade, D. Stoll, M. F. Templin and T. O. Joos, *Annals of Bioanal. Chem*., 2009, **393**(5), 1407–1416.
82. L. Anderson and C. L. Hunter, *Mol. Cell Proteomics*, 2006, **5**(4), 573–588.
83. S. Gallien, E. Duriez and B. Domon, *J. Mass Spectrom*., 2011, **46**(3), 298–312.
84. T. A. Addona, S. E. Abbatiello, B. Schilling, S. J. Skates, D. R. Mani, D. M. Bunk, C. H. Spiegelman, L. J. Zimmerman, A. J. Ham, S. C. Hall, S. Allen, R. K. Blackman, C. H. Borchers, C. Buck, H. L. Cardasis, M. P. Cusack, N. G. Dodder, B. W. Gibson, J. M. Held, T. Hiltke, A. Jackson, E. B. Johansen, C. R. Kinsinger, J. Li, M. Mesri, T. A. Neubert, R. K. Niles, T. C. Pulsipher, D. Ransohoff, H. Rodriguez, P. A. Rudnick, D. Smith, D. L. Tabb, T. J. Tegeler, A. M. Variyath, L. J. Vega-Montoto, A. Wahlander, S. Waldemarson, M. Wang, J. R. Whiteaker, L. Zhao, N. L. Anderson, S. J. Fisher, D. C. Liebler, A. G. Paulovich, F. E. Regnier, P. Tempst and S. A. Carr, *Nat. Biotechnol*., 2009, **27**(7), 633–641.
85. L. J. van't Veer, H. Dai, M. J. van de Vijver, Y. D. He, A. A. Hart, M. Mao, H. L. Peterse, H. K. van der Kooy, M. J. Marton, A. T. Witteveen, G. J. Schreiber, R. M. Kerkhoven, C. Roberts, P. S. Linsley, R. Bernards and S. H. Friend, *Nature*, 2002, **415**(6871), 530–536.
86. A. M. Glas, A. Floore, L. J. Delahaye, T. Witteveen, R. C. Pover, N. Bakx, J. S. Lahti-Domenici, T. J. Bruinsma, M. I. O. Warmoes, R. Bernards, L. F. Wessels, L. J. Van't Veer. Converting a breast cancer microarray signature into a high-throughput diagnostic test, *BMC Genomics*, 2006, **7**, 278.
87. S. Mook, L. J. Van't Veer, E. J. Rutgers, M. J. Piccart-Gebhart and F. Cardoso, *Cancer Genomics Proteomics*, 2007, **4**(3), 147–155.
88. F. Cardoso, M. Piccart-Gebhart, L. Van't Veer and E. Rutgers, *TRANSBIG Consortium. Mol. Oncol*., 2007, **1**(3), 246–251.
89. M. E. Straver, A. M. Glas, J. Hannemann, J. Wesseling, J. E. J. van de VijJ, M. J. Vrancken Peeters, H. van Tinteren, L. J. Van't Veer and S. Rodenhuis, *Breast Cancer Res. Treat*., 2010, **119**(3), 551–558.
90. S. Tian, P. Roepman, L. J. Van't Veer, R. Bernards, R. F. de Snoo and A. M. Glas, *Biomark Insights*, 2010, **28**(5), 129–138.
91. L. Mittempergher, J. J. de Ronde, M. Nieuwland, R. M. Kerkhoven, R. M. I. Simon, I. E. J. Rutgers, L. F. Wessels and L. J. Van't Veer, *PLoS One*, 2011, **6**(2), e17163.
92. www.trans-hit.com.
93. B. M. Knoppers and T. J. Hudson, *Hum Genet*., 2011, **130**, 329–332.