An NCI Perspective on Creating Sustainable Biospecimen Resources

Jimmie Vaught, Joyce Rogers, Kimberly Myers, Mark David Lim, Nicole Lockhart, Helen Moore, Sherilyn Sawyer, Jeffrey L. Furman, Carolyn Compton

Correspondence to: Jimmie Vaught, PhD, Office of Biorepositories and Biospecimen Research, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, 11400 Rockville Pike, Ste 700 Rockville, MD 20892 (e-mail: vaughtj@mail.nih.gov).

High-quality biospecimens with appropriate clinical annotation are critical in the era of personalized medicine. It is now widely recognized that biospecimen resources need to be developed and operated under established scientific, technical, business, and ethical/legal standards. To date, such standards have not been widely practiced, resulting in variable biospecimen quality that may compromise research efforts. The National Cancer Institute (NCI) Office of Biorepositories and Biospecimen Research (OBBR) was established in 2005 to coordinate NCI’s biospecimen resource activities and address those issues that affect access to the high-quality specimens and data necessary for its research enterprises as well as the broader translational research field. OBBR and the NCI Biorepository Coordinating Committee developed NCI’s “Best Practices for Biospecimen Resources” after consultation with a broad array of experts. A Biospecimen Research Network was established to fund research to develop additional evidence-based practices. Although these initiatives will improve the overall availability of high-quality specimens and data for cancer research, OBBR has been authorized to implement a national biobanking effort, cancer HuMan Biobank (caHUB). caHUB will address systematically the gaps in knowledge needed to improve the state-of-the-science and strengthen the standards for human biobanking. This commentary outlines the progressive efforts by NCI in technical, governance, and economic considerations that will be important as the new caHUB enterprise is undertaken.

J Natl Cancer Inst Monogr 2011;42:1–7

With the rapid growth in knowledge of the molecular basis of both disease and health status, biomedical research and health-care delivery are shifting toward a vision of “personalized medicine.” The vision of “personalized medicine” is to improve the standard of medical care by including an individual’s genetic and molecular information in the clinical decision-making process. This molecular information holds the potential to highlight individual differences among patients with the same disease and enable health-care providers to predict a patient’s response to therapy, thereby prospectively guiding the plan for treatment and eliminating trial-and-error approaches.

As the source of molecular information, human biospecimens, such as tissue or blood, are both the foundation of personalized medicine and the fuel that drives the basic and translational research needed to achieve this vision (Figure 1) (1–5). High-quality human biospecimens and associated data are required to define both the biology of the patient and the biology of his or her disease. Recently, the general public has become more aware of the importance of biospecimen resources and the key role they play in the move toward personalized medicine. In fact, in a special March 2009 edition of Time magazine, biobanks were listed among the “Ten Ideas Changing the World Right Now” (6).

Because human biospecimens are important to personalized medicine and translational research, they must be collected and processed following standards that safeguard quality and annotated with the appropriate patient-related and biospecimen-specific information (7). High-quality well-annotated biospecimens for research are currently in high demand but can be very limited in availability (8,9). Before the promise of personalized medicine can be realized, myriad challenges, including biospecimen technical, scientific, and economic factors, need to be resolved.

The NCI Biorepository Experience

Worldwide, millions of biospecimens are collected, processed, stored, and analyzed each year at a significant cost. The National Cancer Institute (NCI), for example, spends what was conservatively estimated in an internal assessment in 2004 to exceed $50 million per year on its basic biospecimen resource infrastructure. Moreover, this figure does not include many resources operated under individual extramural grants. Studies involving biospecimens at NCI include large collaborative projects such as The Cancer Genome Atlas (10) and Clinical Proteomics Technologies Initiative for Cancer (11), as well as a multitude of translational research efforts, clinical trials, and epidemiological studies that collect and process biospecimens under increasingly stringent quality requirements to meet the demands of the latest molecular analysis technologies. The 2004 assessment also indicated that NCI biospecimen resources used a variety of nonstandard practices to collect, process, store, and track specimens.

Although the trend has slowed since 2007, the pharmaceutical industry also invests heavily in research and development projects.
(12–14) requiring a wide variety of tissue samples, often with strict specifications for biospecimen type, quality, and annotation. Over the past two decades, industry’s need for human samples has increased significantly as the importance of biospecimen uses in the drug-discovery pipeline has been recognized (8). With high-quality biospecimens in short supply, biospecimen procurement networks and commercial biobanking brokers have been established in an attempt to meet the demand.

Although the pressing need for quality biospecimens has become increasingly apparent (15), there has been a parallel growing recognition of the problem of extreme variation among biospecimen resources (16,17). Realizing that biospecimens were the cornerstone of future research endeavors, NCI leadership identified human biospecimens as critically important to post genomic cancer research and began to implement a strategy to create sustainable standardized biospecimen resources. In 2002, using surveys and community forums, the NCI undertook an unprecedented internal and external review of biospecimen resources that revealed significant heterogeneity in technical and operational practices. Based on the review, the NCI Biorepository Coordinating Committee was created to sponsor workshops and gather and evaluate authoritative best practices in biobanking that the NCI would then provide as a baseline for operations to the resources that it funded. These efforts to survey the biobanking landscape and identify solutions to the outstanding challenges were part of the stimulus for early discussions about building a National Biospecimen Network with unified and standardized operations and transparent access policies (18). In 2005, the NCI established Office of Biorepositories and Biospecimen Research (OBBR) to provide a focal point for these activities and issues. Since its inception, OBBR has convened several national workshops (19) on “Best Practices for Biorepositories that Support Cancer Research,” “Biospecimen Ethical, Legal, and Policy Issues,” and other specific topics such as the use of pediatric biospecimens in research, biospecimen custodianship/ownership issues, and biorepository economics. In 2007, it created the Biospecimen Research Network (BRN) (20), which funds research in biospecimen science, with goal of strengthening the evidence base for biobanking practices. Last, OBBR oversees the Innovative Molecular Analysis Technologies (21) program, which fosters technology development from inception through development for the market place. The Innovative Molecular Analysis Technologies program specifically solicits technological solutions to biobanking challenges among its targeted applications.

NCI’s efforts ultimately led to development of the NCI “Best Practices for Biospecimen Resources” (22) first published in 2007 after extensive expert input and public comment. The document identifies salient guiding principles that define state-of-the-science biospecimen resource practices, promote biospecimen and data quality, and support adherence to ethical and legal requirements. NCI’s current best practices do not include detailed laboratory operating procedures; rather they consist of principles by which such procedures should be developed. The recommendations contained within that document are intended to be adapted, as appropriate, based on the mission and scientific needs of individual biospecimen resources. Although compliance with the NCI Best Practices for Biospecimen Resources is voluntary, the NCI believes that the principles outlined in that document support the goal of optimizing biospecimens for cancer research. Recognizing the need for best practices to evolve with the field, the NCI Best Practices for Biospecimen Resources document has been updated to a revised online version (23). New appendices to the NCI Best Practices for Biospecimen Resources include tools for the biobanking community such as a template biospecimen resource governance plan and a suggested minimal common dataset for specimen annotation.

The remainder of this article will briefly address other biospecimen resource issues, including emerging technologies; the role of standard setting in improving quality and value; and the potential for new partnerships to realize the vision of personalized medicine in an era of shrinking government, academic, not-for-profit, and industry budgets. The article will conclude with discussions about centralized biospecimen resources and current NCI efforts in this direction.
Biospecimen Resources: Background and Emerging Technologies

Biospecimens generally are organized into collections created for particular research purposes, such as a specific clinical trial or epidemiological study. For the purposes of this article, we will use the term biospecimen resource to describe such collections, although the terms biorepository, specimen bank, biobank, or biological resource center are often used. The NCI’s current best practices define a biospecimen resource “as a collection of human specimens and associated data for research purposes, the physical structure where the collection is stored, and all relevant processes and policies” (22). Biospecimen resources vary considerably in scope, ranging from formal government, academic, and commercial organizations to informal collections of materials in an individual researcher’s freezer.

Within the context of personalized medicine, the quality of human biospecimens is directly proportional to the richness of the associated biospecimen data profile (7,24), as well as researchers’ confidence in the validity and completeness of the information. The biospecimen data profile, or annotation, is the aggregate of information available for each biospecimen, including patient demographics and medical history to make meaningful clinical correlations; biospecimen collection and processing details to account for variability in the biospecimen collection and storage procedures; the type, nature, and composition of the biospecimen including the pathological assessment details; the data yielded by molecular analyses to better understand disease biology; and quality control data for both specimens and clinical data (Figure 2). This biospecimen profile can be further improved by incorporating radiological, pathological-imaging, and clinical laboratory data. The number of layers of information is limited only by the available technology to capture, store, and integrate it, and the scope of the ethical and legal framework within which it is permitted to be used.

Issues Related to Quality, Reproducibility, and Evidence-Based Protocols

The latest generation molecular analysis technologies require rigorous biospecimen quality and pose enormous operational challenges for existing research biospecimen resources, many of which are carved out of existing academic medical institutions without separate operating budgets, objectives, or performance measures (25). Evidence-based and validated standard operating procedures and ongoing quality control programs are needed to ensure human biospecimens are handled in ways that maximize their usefulness and availability for molecular research (1,26).

In 2006, as a way to begin meeting some of these challenges, OBBR created the BRN to systematically address the impact of specific variables (Figure 3) in individual specimen types on molecular data from given analysis platforms. The goal of the BRN is to address these issues by sponsoring, conducting, and collaborating on studies to assess the effects of human specimen preanalytical variables on the outcome of genomic and proteomic studies conducted for clinical diagnosis and cancer research purposes. The BRN program has engaged in public outreach to define issues and needs in biospecimen standardization and consulted on biospecimen issues for programs within NCI and the broader National Institutes of Health. The BRN hosts an annual biospecimen research symposium (9) and has awarded a series of contracts as part of its mission to advance the field (20). By communicating the results of such research to the scientific community, and deploying the data to strengthen the evidence base of biobanking standard operating procedures, the BRN aims to significantly improve the quality of NCI-funded biospecimen-based research. With intramural and extramural collaborators, the BRN research program addresses several challenges in biospecimen science, including bridging gaps between existing clinical practice and emerging technologies, defining significant variables for prospective collections, and developing evidence-based protocols and biospecimen quality indicators.

As part of this biospecimen research effort, OBBR worked with the RAND Corporation and the NCI Center for Bioinformatics and Information Technology to create the Biospecimen Research Database (BRD), a searchable Web-based tool to make biospecimen research data more accessible to the community. By surveying and curating the existing scientific literature for data that define the precise relationships between biospecimen handling and the quality and reproducibility of molecular analysis data for cancer research, OBBR is populating the Biospecimen Research Database (BRD) with published biospecimen research data. Mechanisms to post unpublished biospecimen research datasets as well as evidence-based standard operating procedures are part of the BRD development plan. Information is organized within the BRD in a fashion that allows easy identification and retrieval of variable-specific, biospecimen-specific, or molecular analysis relevant information of interest.
Ethical, Legal, and Social Issues

Biospecimen resources must also consider how ethical, legal, and social issues, such as informed consent and privacy protection, affect the availability and downstream use of biospecimens. Often, informed consent documents for research use of biospecimens do not provide sufficient information about possible future uses of biospecimens, which may unintentionally constrain future research. This problem may be further exacerbated by differences between two different Federal laws: the “Common Rule” (28), which governs the process of informed consent and allows consent for unspecified future research use of specimens and data, and the Health Insurance Portability and Accountability Act Privacy Rule (29), which protects personal health information but does not permit patient authorization of unspecified future use of health information (28–30). Although these laws serve important functions for protecting patients and individuals, they can also create roadblocks to research that could ultimately benefit many people. Additionally, informed consent documents often limit sharing of biospecimens and/or data by either failing to describe or directly precluding use of biospecimens by researchers outside the primary institution, especially commercial or industry partners. Restrictive informed consent language can become a significant hurdle in large-scale team-science projects that require large numbers of appropriately consented biospecimens that must be amassed from multiple diverse sources (7,31,32).

Additionally, the lack of standardized approaches to informed consent and privacy protection leads not only to possible underutilization of valuable biospecimen collections but also to confusion and duplication of effort. If a biospecimen collection is not initially consented for downstream research use, then further institutional review board action may be required, such as a waiver of consent or approval to recontact and re-consent living research participants. This process could slow the progress of the proposed research, necessitate additional staff effort, and ultimately, increase costs. Although the NCI Best Practices for Biospecimen Resources and other similar documents have provided guidance on these issues, much work remains to be done to standardize ethical and legal approaches in ways that facilitate research and accelerate medical progress for the benefit of patients. OBBR hosted a series of workshops concerning these issues, including specimen ownership and custodianship and pediatric consent issues (33), as well as a workshop in 2010 about the return of research results to study participants.

Economic Issues: Biospecimen Resources “Cost” and Value

To a certain extent, all biospecimen resources are established in unique setting-specific ways, which presents challenges in understanding the true costs of these resources. A biospecimen resource can be a freezer in a hallway, a fully staffed and independent operation, or an extensive virtual network of resources, such as the NCI Cooperative Human Tissue Network, the European Human Frozen Tumour Tissue Bank network (TuBaFrost), and the European Biobanking and Biomolecular Resources Infrastructure (BBMRI) (34–36). Most of the larger biospecimen resources in the United States and abroad are not-for-profit and receive financing through a central or shared funding mechanism as their objective is to support research at a local institution or within a large integrated program (37). But the lack of a uniform model, the diversity in funding mechanisms, and the dearth of publicly available financial data complicate the development of metrics that can be used to assess a biospecimen resource’s true cost or true value.

To better understand the economic issues associated with biorepositories, OBBR convened a special workshop in 2008 dedicated to the topic of “Biorepository Economics” (38). The workshop identified a number of alternative public–private partnership models that may represent the most sustainable biorepository approach. Gaps were also identified in understanding the costs involved in the establishment and ongoing operations of large-scale biorepositories. In 2009, OBBR commissioned a study on the economic considerations for developing a national biospecimen resource. Described in greater detail in another article in this monograph (13), Booz Allen Hamilton’s economic and business analysis team surveyed the current for-profit biobanking market, developed a detailed cost model for
In 2009 and 2010, the Biospecimen Research Network (BRN) awarded contracts to study the following:

- Research studies on the effects of intraoperative ischemia time on gene and protein expression patterns in liver and colon tissue: Preliminary data indicate that intraoperative ischemia time has a dramatic effect on gene expression patterns. Additional experiments will examine ischemia-dependent changes in immunohistochemical targets, and assay for key phosphoproteins in cell lysates, and ischemia-dependent proteins such as hypoxia-inducible factor 1 and heat shock protein 27.
- Investigations into the effects of blood specimen handling procedures on protein integrity: Multiplexed highly reproducible protein assays for cancer-related proteins are being developed and tested on biological samples in preparation for testing the supported collection; preliminary liquid chromatography-mass spectrometry studies are being performed to assess proteolysis and posttranslational modifications resulting from specific biospecimen preanalytical variables.
- Credentialing plasma and serum biospecimen banks for proteomics analyses: Several analytical approaches are under way to examine the effects of preanalytical variables on the proteolysis-driven changes in plasma and serum proteins.
- Intrinsic and extrinsic controls for formalin-fixed paraffin-embedded tissue: A robust antibody validation process has been developed and two potential predictors of tissue integrity have been identified (beta-actin and histone H4). Additional studies are planned to assess the effects of postoperative ischemic time on clinically relevant markers (such as estrogen receptor, progesterone receptor, human epidermal growth factor receptor 2, Ki67, tumor protein 53, and cytokeratin) as well as phosphor-signaling molecules.
- Effects of biospecimen integrity, intratumoral heterogeneity, and analytical variance on microarray-based pharmacogenomics tests of breast cancer.

In 2009 and 2010, the Biospecimen Research Network (BRN) awarded contracts to study the following:

- Research studies on the effects of intraoperative ischemia time on gene and protein expression patterns in liver and colon tissue: Preliminary data indicate that intraoperative ischemia time has a dramatic effect on gene expression patterns. Additional experiments will examine ischemia-dependent changes in immunohistochemical targets, and assay for key phosphoproteins in cell lysates, and ischemia-dependent proteins such as hypoxia-inducible factor 1 and heat shock protein 27.
- Investigations into the effects of blood specimen handling procedures on protein integrity: Multiplexed highly reproducible protein assays for cancer-related proteins are being developed and tested on biological samples in preparation for testing the supported collection; preliminary liquid chromatography-mass spectrometry studies are being performed to assess proteolysis and posttranslational modifications resulting from specific biospecimen preanalytical variables.
- Credentialing plasma and serum biospecimen banks for proteomics analyses: Several analytical approaches are under way to examine the effects of preanalytical variables on the proteolysis-driven changes in plasma and serum proteins.
- Intrinsic and extrinsic controls for formalin-fixed paraffin-embedded tissue: A robust antibody validation process has been developed and two potential predictors of tissue integrity have been identified (beta-actin and histone H4). Additional studies are planned to assess the effects of postoperative ischemic time on clinically relevant markers (such as estrogen receptor, progesterone receptor, human epidermal growth factor receptor 2, Ki67, tumor protein 53, and cytokeratin) as well as phosphor-signaling molecules.
- Effects of biospecimen integrity, intratumoral heterogeneity, and analytical variance on microarray-based pharmacogenomics tests of breast cancer.

Access to high-quality biospecimens depends on the ability to develop funding models that generate sufficient revenue to at least partially recover costs associated with the resource’s operation, namely the costs of collecting, processing, storing, and distributing biospecimens (39–41). Although in the past organizations have created separate resources, unique opportunities for alternative partnerships models exist among Federal organizations such as the NCI, not-for-profit organizations such as patient advocacy groups, and for-profit entities, namely biotechnology and pharmaceutical companies. These partnerships may provide financially sustainable funding models that help each participant or partner realize significant returns on its financial and scientific investments through centralized resources.

The establishment of public–private partnerships has become an increasingly popular approach to supporting translational research (40). Establishing a business model that unifies all stakeholders around a common infrastructure can be both highly beneficial and highly complex. However, to sustain biospecimen resource partnerships, issues concerning the structure of collaborations and biospecimen access must be addressed (41).

Several questions become apparent when considering a public–private partnership model for financing and supporting biospecimen resources:

1. What are the incentives for participation in a shared resource?
2. What policies and procedures create transparency and inspire an environment of mutual trust and accountability?
3. How does each entity contribute to the cost of running such an enterprise?
4. Who has access, and do nonmonetary forms of payment exist for access to biospecimens, such as sharing information obtained from research with biospecimens with partners and/or the research community?

Federal and regulatory communities have realized the importance of biospecimen resources and have produced guidelines and recommendations for their optimal operation, such as NCI’s Best Practices for Biospecimen Resources (22). Given federal interest (12) in this issue and given that biospecimen resources represent key, albeit labor-intensive, infrastructures that are not likely to be profitable, the Federal Government could be an important partner in funding and management models that are developed to support biospecimen resources. Expert opinion gathered at the 2008 biorepository economics workshop (38), as well as extensive NCI-sponsored market research, suggests that the United States, with appropriate Government financial and management support, should consider a national biobanking system (42,43). The concept of such a national resource garnered broad support from the market research participants, who also noted several challenges: sustainability, standardization of data and processes, clearly defined access policies, and intellectual property concerns.

National initiatives are already under way in many other countries. Broadly speaking, the national biobanking efforts are of two distinct types: population-based research banks that collect normal biospecimens from all individuals in a defined geographical region and translational research banks that collect disease specimens from affected individuals (patients) in a defined catchment area. Less commonly, biobanking efforts may encompass both approaches.
The UK Biobank is working to improve the prevention, diagnosis, and treatment of illness through the formation of a national population-based biospecimen resource (44). Funded by numerous government and private organizations, this effort aims to collect blood samples from 500,000 people between the ages of 40 and 69 years. In Singapore, the Singapore Tissue Network collects tissue and DNA samples to support translational and population research (45). The Pan-European BBMRI is an effort to develop a network of existing and new biospecimen resources for translational research (36). The success of national biobanking efforts depends on a number of factors, including a clear governance model, a quality management program, a business model that promotes long-term sustainability, and adherence to a set of best practices. A variety of circumstances can make it difficult for national biobanking programs to survive in the long term. For example, onCore UK decided to change its approach to a more localized tissue collection effort after 4 years of attempting to create a centralized national biobank (46). Such lessons will be important to consider as NCI further develops its own national biospecimen resource program.

Future Directions for Biospecimen Resources

OBBR is creating a biospecimen resource, the cancer HUman Biobank (caHUB), that will take advantage of foundations already laid by the NCI, including the BRN and the NCI Best Practices for Biospecimen Resources (47). caHUB provides an opportunity to implement concepts from the National Biospecimen Network Blueprint, an NCI-developed consensus document published in 2003 (48) that outlined the vision of a national system for providing standardized biospecimens of the requisite quality for advanced analytical technologies and standardized associated data of the requisite type and amount to assure clinical correlation. caHUB is to be the realization of that vision. caHUB will address systematically the gaps in knowledge needed to improve the state-of-the-science and strengthen the standards for human biobanking.

Lessons learned from other countries indicate that national biobanks have been based on models that afforded little control over quality and greatly increased the costs of operation. In addition, projects managed by the NCI, including the Cancer Genome Atlas, have been hampered by the lack of consistent centralized control over specimen and data quality parameters. For these reasons, caHUB specimen and data collection will be managed under strictly controlled prospective collection protocols and a comprehensive quality management program. This approach should prevent most of the inconsistencies in quality that have been experienced by previous biospecimen resource efforts.

During its pilot phase, caHUB aims to modernize the field of biobanking by creating evidence-based standard operating procedures and biospecimen quality standards for the translational and product development communities. To do this, caHUB will adhere to the following key tenets:

- Rich standardized data profile for each sample.
- Contribution to the state of the science through peer-reviewed publications.

Achieving the vision of personalized medicine depends on the availability of high-quality biospecimens for translational research and, ultimately, for the medical management of individual patients. To fully realize the possibility of this personalized approach, the technical, scientific, ethical, and economic hurdles facing the creation of sustainable biospecimen resources need to be addressed.

References


**Funding**

National Science Foundation’s Science of Science Policy Program (0738394) to J.F.

**Notes**

In June 2008, the NCI hosted a “BioEconomics Workshop.” Many of the themes developed in this article were initially discussed at the workshop. The contributions of government, academic, and industry experts who contributed to those discussions are greatly appreciated (49). The authors appreciate the excellent editorial assistance of Andrea Kelly, PhD, of Rose Li and Associates.

**Affiliations of authors:** Office of Biorepositories and Biospecimen Research, National Cancer Institute, NIH, DHHS, Rockville, MD (UV, KM, MDL, NL, HM, SS, CC); Kelly Services, Rockville, MD (JR); School of Management, Boston University, Boston, MA, and the National Bureau of Economic Research, Cambridge, MA. (JLF).