National Institutes of Health
National Cancer Institute
Office of Biorepositories and Biospecimen Research

SUMMARY

National Cancer Institute Biospecimen Best Practices Forum

The Bell Harbor Conference Center
Seattle, Washington

January 28, 2008

Prepared by:
Rose Li and Associates, Inc.
February 2008

Table of Contents

I. Introduction 1

II. Part 1: Overview and Discussion of *NCI Best Practices* 2
NCI Best Practices for Biospecimen Resources 2
Overview of Technical and Operational Best Practices 3
Question-and-Answer Session 4
Overview of Ethical, Legal, and Policy Best Practices 5
The Importance of Best Practices to Patients, Survivors, Advocates, and the General Public 7
Question-and-Answer Session 8

III. Part 2: Informatics Best Practices and Economic Issues for Biospecimen Resources 9
caBIG™, caTissue, and Achieving Silver-Level Compatibility 9
Cost Recovery Models and Other Economic Issues Involved in the Implementation of the *NCI Best Practices* 11
Question-and-Answer Session 14

IV. Part 3: Next Steps for Updating the *NCI Best Practices* 15
Assessing the Effects of Preanalytical Variables on Molecular Research: The Biospecimen Research Network 15
The Biospecimen Research Database 16
Question-and-Answer Session 16

V. PART 4: CONCURRENT BREAKOUT SESSIONS 17

V. Part 5: Closing and Adjournment 17
Breakout Session Reports 17
Discussion of Cost Recovery Models and Other Economic Issues Involved in the Implementation of the *NCI Best Practices* 17
Demonstration of caBIG™ Biospecimen Resource Management Tools 18
Biospecimen Science 18
Closing Remarks 19

Appendix—Breakout Session Summaries 20
Discussion of Cost Recovery Models and Other Economic Issues Involved in Implementing the *NCI Best Practices* 21
Question-and-Answer Session 22
Demonstration of caBIG™ Biospecimen Resource Management Tools 24
Question-and-Answer Session 25
Biospecimen Science 27
Question-and-Answer Session 29

I. INTRODUCTION

Cancer research in the 21st century is moving toward a vision of personalized medicine where clinical and molecular data are used to treat individual patients with greater specificity, reduce the frequency of adverse events, and determine disease predisposition to allow early detection and prevention. In today's cancer medicine, the analysis of human specimens supports diagnosis, staging, and prognosis. In addition, these materials provide a critical link between molecular and clinical information for the personalized medicine of the future. The collection of accurate molecular data to inform the development of personalized medicine depends heavily upon the quality and consistency of the biospecimens analyzed in the translational and validation research arenas.

Over the past several years, the National Cancer Institute (NCI) has undertaken an intensive due diligence process to understand the state of its funded biospecimen resources and the quality of biospecimens used in cancer research. Based on extensive input from cancer research experts including clinicians, scientists, ethicists, biotechnology and pharmaceutical professionals, patients, survivors, advocates, and numerous authoritative sources and regulatory bodies, the NCI developed the *NCI Best Practices for Biospecimen Resources*.[1] The purpose of the *NCI Best Practices* is to define state-of-the-science practices for acquiring tissues and fluids from research participants to promote biospecimen and data quality and consistency and to encourage adherence to the highest ethical and legal standards to support the development of new cancer interventions.

The purpose of this forum was to inform and obtain feedback about the *NCI Best Practices* from intramural and extramural research communities in and around Seattle, WA. This forum was the fourth in a series of public meetings to be held across the United States.[2] The forums were designed to address major areas of stakeholder concern and interest based on public comments received on an earlier draft of the document. The forum included NCI and non-NCI speakers to offer different perspectives on the practical impact of the *NCI Best Practices* on the cancer research and patient communities and provided time for questions and feedback from the audience. In addition to presenting external perspectives about the *NCI Best Practices* during the plenary presentations, non-NCI speakers had an opportunity to offer their opinions in response to questions and comments from the audience. The NCI intends to use feedback gathered from the non-NCI speakers and audience participants at these forums to inform, update, and plan for future versions of the *NCI Best Practices*.

[1] Full-text document available at http://biospecimens.cancer.gov/practices/.
[2] Further information on public meetings available at http://biospecimens.cancer.gov/practices/forum/.

II. PART 1: OVERVIEW AND DISCUSSION OF *NCI BEST PRACTICES*

NCI Best Practices for Biospecimen Resources
Why Do We Need Biospecimen Best Practices?
Carolyn C. Compton, M.D., Ph.D., Director, Office of Biorepositories and Biospecimen Research, NCI

Dr. Carolyn Compton is the Director of the Office of Biorepositories and Biospecimen Research (OBBR), with responsibility for developing a common biorepository infrastructure that promotes resource sharing and team science, and establishing biobanking as a new area of research. She came to the NCI from McGill University where she served as the Strathcona Professor and Chair of Pathology and the Pathologist-in-Chief of McGill University Health Center. Prior to this, she had been Professor of Pathology at Harvard Medical School and Director of Gastrointestinal Pathology at Massachusetts General Hospital. Dr. Compton holds leadership positions in several professional organizations such as the College of American Pathologists, the Cancer and Leukemia Group B, the American Joint Committee on Cancer, and the American Society of Clinical Oncology. She is a member of the editorial boards of *Cancer*, *Cell Preservation Technology*, and *Clinical Proteomics*.

Dr. Compton opened the forum by stating that cancer research is at an inflection point as exponential technological advances enable new understanding of cancer biology. High-quality biospecimens are essential for taking full advantage of new techniques and moving into an era of personalized medicine. The dearth of large quantities of high-quality, clinically annotated biospecimens is the main roadblock to advances in translational medicine.

The biospecimen banking community faces significant challenges from varying processes and procedures that compromise molecular research:

- Methods of collection, processing, and storage can affect the physical or biologic state of the biospecimen.
- Biospecimen-associated data elements influence how much the researcher knows about the character and nature of the specimen.
- Clinical information determines what the researcher knows about the patient and, therefore, the biologic context of the biospecimen.
- Legal, ethical, and policy restrictions—such as those relating to informed consent documents—dictate what the researcher may do with the biospecimen or the data.

To address these issues, the NCI, in conjunction with the biospecimen research community, has identified key requirements for biospecimen resources.[3] These requirements include best practice–based, data-driven technical and operational standards to ensure quality and enable reproducible molecular analysis; consistent, high-quality biospecimen annotation, encompassing pathological and clinical data; biospecimen access through a timely, centralized peer-review

[3] The NCI defines biospecimen resource as "a collection of human specimens and associated data for research purposes, the physical entity where the collection is stored, and all relevant processes and policies."
Source: *National Cancer Institute Best Practices for Biospecimen Resources* (http://biospecimens.cancer.gov/practices/).

process; ethical and privacy compliance through a well-defined chain of trust; state-of-the-art informatics systems to track biospecimens, associated data (clinical, pathological, and quality control), and patient consents; and communication with the public.

The *NCI Best Practices* was published with the dual objectives of unifying policies and procedures for NCI-supported biospecimen resources for cancer research and providing a baseline for operating standards on which to build as the state of the science evolves. It is a living document that will be updated in response to evidence-based recommendations. Periodic revision of the *NCI Best Practices* will occur with input from researchers, biospecimen resource managers, advocates, policymakers, and related stakeholders as changes in science, law, and policy occur. New tools and supplemental guidance in key areas will be added as appendices and/or posted to the OBBR Web site. Further, the NCI and, specifically, the OBBR are committed to developing biospecimen research as a valid area of scientific investigation worthy of funding and will be developing evidence-based standard operating procedures (SOPs).

Overview of Technical and Operational Best Practices
State-of-the-Science Biospecimen Handling: Real-World Perspective
Janet Warrington, Ph.D., Vice President, Standards and Government Policy, Affymetrix, Inc.

Dr. Janet Warrington is Vice President of Standards and Government Policy for Affymetrix. She also has served as Vice President of Research and Development for Emerging Markets and Molecular Diagnostics at Affymetrix, at which time she was the originator and lead scientist for the first Food and Drug Administration (FDA)–cleared microarray system for a diagnostic test. Dr. Warrington has extensive experience in the development of clinical applications for molecular analysis tools and serves in advisory roles in both public and privately led consortia, standards development organizations, and committees.

Dr. Warrington began by describing the long, costly, and often unsuccessful path to developing drugs and diagnostics. Cancer researchers would like to realize personalized medicine more efficiently, which will require analyses of high quality clinical, molecular, and family history information. This goal has been hampered by difficulties with data quality, which is dependent on balancing sources of variability with sample sizes sufficient to enable statistically significant results. The greatest source of variability in cancer research is the biological sample. Even the most painstakingly detailed analysis will lead to inaccurate or irreproducible results if such variables as biospecimen handling specifications, protocol deviations, incomplete or ambiguous records, and others are not recognized and minimized. The goal of the OBBR in producing and supporting the *NCI Best Practices*, and others in developing related guidelines, is to improve consistency and standardization for all aspects of biospecimen science by establishing meaningful SOPs and agreed-upon metrics and terminology.[4]

[4] Other guidelines are available through the Clinical and Laboratory Standards Institute (http://www.clsi.org), the International Society for Biological and Environmental Repositories (http://www.isber.org), the Marble Arch Working Group (http://www.oncoreuk.org/pages/MarbleArchWorkingGroup.html), the Organization of Economic Cooperation and Development (http://www.oecd.org), and the International Agency for Research on Cancer (http://www.iarc.fr/News/RecommendationsBRC.pdf).

Dr. Warrington then reviewed the technical and operational recommendations in section B of the *NCI Best Practices* that address biospecimen collection and processing, monitoring and storage, biosafety, packaging and shipping, collecting and managing clinical data, and recordkeeping. She concluded by stating that success in obtaining qualified biospecimens relies upon establishing goal-appropriate, well-defined protocols and employing well-trained personnel.

Question-and-Answer Session

A participant asked about the feasibility of following a patient's clinical outcome after he/she has donated a biospecimen and whether the NCI is conducting any studies in which this is being done. Dr. Warrington responded that this would be possible if it were built into the study design. Dr. Compton added that such an addition to the annotation would greatly increase the value of the specimen. The NCI is in the process of planning a national-level biorepository, a consideration of which has been how to obtain and integrate clinical outcome information with biospecimen data. The open-access cancer Biomedical Informatics Grid (caBIG™) software is designed for interoperability to facilitate the integration and exchange of data from all sources. The tracking of clinical outcome information is one promise of the electronic medical record. Dr. Ian Fore, Associate Director for Biorepository and Pathology Informatics at the NCI Center for Bioinformatics and a member of the NCI OBBR, added that a key issue to interoperability is breaking down barriers involving limitations in information technology (IT) and the territoriality of some researchers, both of which the caBIG™ development team is working to address.

Another attendee observed that the recommendations in the *NCI Best Practices* are general and asked about resources for specific SOPs. Dr. Compton emphasized that the recommendations focus on principles; details will be published in appendices, which are currently under development. Additionally, packages of guidance concerning all levels and elements of best practices will be available on the OBBR Web site in the future. She further added that data on specific biomolecules as analyzed on specific platforms are difficult to find, often buried in the "materials and methods" sections of studies performed with other goals in mind. This situation will change as the OBBR legitimizes biospecimen science. For example, a partnership with the College of American Pathology (CAP) enables data from NCI-sponsored research to flow to CAP members, who will employ their expertise to develop granular benchtop SOPs. Dr. Compton requested feedback from audience members regarding the areas toward which it would be the most useful to direct these efforts.

A participant inquired about an OBBR contact for investigators establishing or updating a biorepository. Dr. Compton encouraged anyone seeking such input to e-mail the OBBR. She described self-evaluation tools in development that will help biorepository management identify strengths and weaknesses. Someone establishing a new biorepository has an advantage over those retrofitting an existing facility; the OBBR can direct investigators to members of the extramural community who can help ensure that the resource fits the needs of the users.

A Webcast participant asked whether biorepositories should require that data from rare or valuable samples be deposited in a central database to help prevent duplication of efforts between laboratories. Dr. Compton responded that one of the goals of the *NCI Best Practices* is to break down "information silos," making important information accessible to the entire

research community. While one goal of the *NCI Best Practices* is that data generated from biospecimen research be high quality, another is that the data be accessible. The labor intensiveness of the data collection process inspires some investigators to hoard samples and information. The OBBR addressed this issue by requiring that biorepositories funded with tax dollars have a transparent sharing policy. Other investigators might claim that they do not have the computer technology to enable sharing. The caBIG™ team is developing tools that will enable valuable data and valuable biospecimens to be shared for the collective good.

A researcher inquired as to the best way to ensure that assays conducted in different laboratories in a collaborative study generate results that are comparable. Dr. Warrington responded that to rely on data from multiple sites, it is important to establish SOPs, identify nonexperimental tissue that can be used in cross-training, and exchange technicians between sites to ensure alignment in training and experimental procedure. Dr. Compton asserted that national standardization of ethical, legal, and policy documents such as consent forms and material transfer agreements would significantly streamline multisite studies.

The final comment involved the lack of consistency between institutional review boards (IRBs), and frustration with the need to satisfy the requirements of multiple IRBs. The participant commented that consistency across institutions would be welcomed. Dr. Compton vehemently agreed.

Overview of Ethical, Legal, and Policy Best Practices
Ethical, Legal, and Policy Implications of Using Human Specimens in Research: What You Need To Know
Karen Thiel, Ph.D., J.D., Patton Boggs, LLP

Dr. Karen Thiel is an attorney at Patton Boggs and formerly an academic researcher, policy consultant, and Federal policy analyst. Dr. Thiel participated in the Surgeon General's Task Force on Drug Importation, served as Chief Evaluator in the Department of Health and Human Services (HHS) Office of Population Affairs, and was a senior consultant for the Joint Committee on Science and Technology in the California State Legislature. In addition to her work as an attorney, she partners with professional associations, medical institutions, healthcare providers, and patient advocacy groups to develop their legislative agendas and strategies related to proposed regulations, with a focus on Medicare, Medicaid, and State Children's Health Insurance Program issues.

After a presentation of the *NCI Best Practices* on informed consent, privacy protection, custodianship, intellectual property, and access (found in section C of the *NCI Best Practices*), Dr. Thiel noted that two issues illustrate the interaction of ethics, law, and policy: (1) The statutory and regulatory maze affecting biospecimens and (2) the unresolved issue of "custodianship," or "ownership," of biospecimens. The NCI has attempted to address both of these issues in turn through (1) the 2004 NCI 50-State Survey of Laws Regulating the Collection,

Storage, and Use of Human Tissue Specimens and Associated Data for Research and (2) a recent symposium, Custodianship and Ownership Issues in Biospecimen Research.[5]

Dr. Thiel discussed the relationship between State statutes and regulations and Federal law. While some State statutes incorporate the Common Rule and the Health Insurance Portability and Accountability Act (HIPAA) privacy protections provisions, they may be more rigorous. State laws that are more stringent in these areas are not preempted by Federal laws. In addition, agency guidance (i.e., FDA, Office for Human Research Protections [OHRP]) does not have the force of law. Although courts may look to these documents, they ultimately decide cases based on statutes, regulations, and legal precedent. Dr. Thiel reported that as of 2004, nearly half of the States had research exceptions that permit disclosure of medical information to researchers, and 21 States allow research use of genetic information with conditions that also vary from State to State. She expressed the opinion that a State-by-State summary of these varying statutes and regulations must be updated annually to serve as a useful tool for researchers. Regarding biospecimen custodianship, Dr. Thiel noted that this issue has yet to be resolved effectively in statute, regulation, or case law. Legal challenges to ownership of biospecimens are fact specific, and decisions made to date in a small number of cases have not yielded a robust body of law.

Dr. Thiel then provided some history on how the U.S. courts have interpreted Federal regulations and guidelines on issues surrounding biospecimens. Most importantly, from the landmark *Moore* v. *The Regents of the University of California* to the recent and highly publicized *Washington University* v. *Catalona* cases, the courts have denied claims of biospecimen ownership based on common law property theories, essentially applying gift laws to biospecimens. Current case law shows that courts scrutinize informed consent forms and any additional documents provided to research participants at the time of consent. Dr. Thiel supports standardizing consent forms as a means of documenting biospecimen donation and mentioned the concept of a biospecimen equivalent to the Uniform Anatomical Gift Act. She concluded that courts often do not address biospecimen issues directly because of the lack of clear guidance in the Federal regulations. In addition, courts are protective of the research process and view biospecimen donation as an altruistic process.

Dr. Thiel closed by outlining several options for researchers given the current legal environment:

- Perform a methodical analysis of applicable statutes and regulations as illustrated in the NCI's 50-State Survey;
- Encourage the NCI or another entity to compile and maintain a real-time listing of State statutes and regulations affecting all aspects of biospecimen-related research to facilitate standard State-by-State analyses;
- Rely on IRB review of the informed consent process and its reflection of Federal and State requirements; and
- Closely monitor court decisions and their applicable jurisdiction.

[5] The NCI publication 50-State Survey of Laws Regulating the Collection, Storage, and Use of Human Tissue Specimens and Associated Data for Research can be found at http://www.cancerdiagnosis.nci.nih.gov/specimens/50_state_survey/index.htm.

The Importance of Best Practices to Patients, Survivors, Advocates, and the General Public
It's All About the Patient: Putting Biospecimen Research in Perspective
Deborah Collyar, Patient Advocates in Research

Ms. Deborah Collyar is a two-time cancer survivor and has been a leader in cancer patient advocacy since 1991. She founded the Patient Advocates in Research (PAIR) international network and initiated the patient advocacy component in Cancer and Leukemia Group B. She has served as a patient advocate in multiple other capacities: Program Director for the Specialized Programs of Research Excellence Patient Advocate Research Team Program grant; as a member on many NCI/NIH committees; as faculty for American Association for Cancer Research and American Society of Clinical Oncology (ASCO) workshops; and as a member of ethical and executive advisory boards for companies, cancer centers, and advocacy organizations.

Ms. Collyar focused her presentation on the patients from whom biospecimens are derived and who ultimately benefit from biospecimen research. From their perspective, the promises of cancer research are great, and yet there have been few successes in treatment development to date. Collaborations that are critical to these successes are advancing slowly, and the pathology paradigm is shifting as the value of high-quality biospecimens is being understood and appreciated among researchers.

Ms. Collyar underscored the need to educate both the scientific and public communities to ensure that patients make informed choices. Currently, the general public does not think about biospecimens in the context of advancing cancer research. In addition, although most patients are willing to contribute biospecimens to research, they often neglect to consider the ramifications. Patients may not understand how research affects their privacy and confidentiality in this genomic age; specifically, that DNA has the potential to be used as a unique barcode identifying the individual. Terms used frequently in the scientific community, such as "positive results," also have differing and occasionally opposite meanings to patients and their doctors. Finally, most research participants have common expectations: That patient care comes first (e.g., research specimens will be available for future clinical use); that researchers will adhere to informed consent requirements; that biospecimens and data will be shared for the greatest common good; and that investigators will communicate with participants and keep them abreast of research advances. Understanding and accommodating these expectations will enable investigators to ensure continuing public trust.

Ms. Collyar then explained the role of patient advocates in research. Advocates operate as liaisons, sharing goals with researchers while acting as proponents of participants' expectations. Several examples of biospecimen partnerships were offered in which researchers and patients are brought together and resources and/or collaborators found, with the help of advocates, to meet research goals.

In closing, Ms. Collyar provided links to a number of publications on biospecimens and their use in research.[6] She also encouraged the biospecimen scientists in the audience to advance patient–

[6] Publications to inform the general public about biospecimens and their use in research include the NCI tissue brochure (http://www.cancer.gov/clinicaltrials/resources/providingtissue), Research Advocacy Network materials

researchers relationships by sharing biospecimens, data, and knowledge with colleagues; training themselves and their coworkers on human relations; treating individuals with respect; explaining to the research participant what specimens and data are collected and how they will be used; and explaining their options and commitments. Ms. Collyar concluded by stating that the biospecimen and research systems must be repaired now to realize the promise of personalized medicine in the future.

Question-and-Answer Session

An attendee asked how privacy advocates should be engaged. Ms. Collyar replied that privacy advocates generally focus on Internet privacy and are not as likely to be versed in research and medical privacy issues. She recommended that medical advocates reach out to them.

Regarding the *Washington University* v. *Catalona* decision, a participant asked what Dr. Catalona might have done differently to have legal permission to take the biospecimens collected at Washington University with him to his new institution. Dr. Thiel replied that court decisions would factor in donor intent when considering informed consent forms. Tiered consent that allowed for various eventualities might have been useful. Ms. Collyar added that Dr. Catalona might have been able to access the biospecimens after leaving the university if the consent form had included a check box that would enable sharing with researchers at other institutions. Dr. Thiel mentioned that Washington University had argued that Dr. Catalona could access the biospecimens via a materials transfer agreement like any other researcher outside of the university.

An attendee followed up on the previous question by observing that Office for Protection from Research Risks guidance discourages use of the word "donation" for biospecimens. Dr. Thiel replied that "donate" can imply giving up all legal rights to the biospecimen, whereas patients retain the right to withdraw consent. Dr. Compton added that the problem of inconsistencies between guidances from various Governmental agencies has been discussed at the highest levels within the HHS; harmonization is a work-in-progress. Ms. Collyar further added that such harmonization is an issue patient advocacy groups have been working on advancing.

A discussant who works primarily with underserved populations expressed concern that individuals who have limited education and generalized fear of the research enterprise are refusing to participate in research because of their uncertainty about the information presented in informed consent forms. She called for high-level language, such as discussed in this forum, to be translated into laymen's terms when it is communicated to potential research participants. Dr. Thiel applauded this individual's efforts to perform the consent process with hard-to-reach populations, noting that there are other populations that are never involved in research because they do not have access to an institution that (1) collects biospecimens and (2) has culturally sensitive staff trained in consenting such participants. Ms. Collyar added that models are in development for improving communication with patients, building trust, and correcting

(http://researchadvocacy.org/publications/pdf/tissue_WhatIsTissue.pdf and http://researchadvocacy.org/publications/pdf/tissue_ConsiderDonating.pdf), the Dana-Farber Cancer Institute brochure (http://dana-farber.org/res/tissue-banking.html), and The Wesley Research Institute brochure (http://www.wesleyresearch.com.au/docs/TISSUE_BANK.pdf).

misinformation. She mentioned a community workshop that she helped develop focused on explaining research and clinical trials and called for extending that dialogue with communities to include biospecimen donation and informed consent.

Finally, a discussant asked, in light of greater transparency and the drive to prevent duplication of research efforts, whether more negative findings are likely to be published. Dr. Compton replied that negative results rarely get published; thus, similar experiments get repeated by different investigators, which detracts from research advancement. As the gatekeepers of data dissemination, journal editors must recognize that negative findings merit publication, while studies employing poor-quality biospecimens do not.

III. Part 2: Informatics Best Practices and Economic Issues for Biospecimen Resources

caBIG™, caTissue, and Achieving Silver-Level Compatibility

Informatics Solutions to Biospecimen Management: Finding the Right Tools for Your Resource

Ian Fore, D.Phil., Associate Director for Biorepository and Pathology Informatics, NCI Center for Bioinformatics

Dr. Ian Fore is Associate Director for Biorepository and Pathology Informatics at the NCI Center for Bioinformatics and a member of the OBBR team. Prior to joining the NCI, he worked on drug discovery informatics at Wyeth Pharmaceuticals and Johnson and Johnson. More recently, Dr. Fore was at Celera Genomics as Senior Product Manager for a toolkit to integrate the company's databases into customer bioinformatics systems.

Dr. Fore began his presentation with a discussion of the functionality, integration, and security required in informatics tools to achieve compatibility with existing informatics systems. Functionality of biorepository software was highlighted in terms of supporting key operational activities such as biospecimen tracking. Dr. Fore explained that the best informatics software tools for tracking treat each biospecimen as a distinct and identifiable entity while maintaining the appropriate parent-child relationship throughout processing steps. The linkage between the physical biospecimens and the informatics system is best maintained through the use of barcoded samples. The ability to integrate these tracking systems with disparate databases containing clinical and outcome data on biospecimens is also critically important for implementing regulatory and sharing requirements. Finally, as trusted custodians of biospecimens, there is a responsibility to provide adequate security of these precious resources through appropriate safeguards. Access to informatics systems must be limited by physical means as well as technical means, including login protections and role-based security. Systems also must be backed up regularly to prevent possible data loss.

Dr. Fore offered the following recommendations for biospecimen resource managers aiming to build or buy an informatics system:

- Look for systems that use structured information as opposed to free text.
- Consider the true cost of system development, installation, and maintenance when deciding whether to build or buy.

- Plan for the future—seek out a software platform that is robust and will likely still be in existence years later when the biospecimen resource realizes its value.
- Consider open-source software with accessible code that can be customized or adapted to individual needs if you have software support resources available.

If building a software system, Dr. Fore stressed the importance of getting the users involved early to provide input into the development of use cases. A unified software development process that is iterative and incremental also should be followed. Again, it is important to determine the true cost of building in terms of time and financial resources that will need to be invested.

Dr. Fore then shifted to a discussion of NCI's informatics infrastructure, caBIG™, the mission of which is to develop a truly collaborative information network that accelerates the discovery of new approaches for the detection, diagnosis, treatment, and prevention of cancer, ultimately improving patient outcomes. It comprises a community of over 1,000 individuals working in the domains of clinical trial management systems, integrative cancer research, biospecimen banks and pathology tools, and imaging.

In an effort to avoid penalizing systems based on size, caBIG™ relied on a modular approach to span the cancer research landscape, which ranges from institutions with integrated, IT staff–supported systems to those with informal or no information systems. The three prongs of the caBIG™ approach are the development of modules that address specific needs, connection through defined electronic interfaces, and use of international data standards. Dr. Fore emphasized that the focus of caBIG™ is not on the particular content of systems but rather on providing the necessary middleware infrastructure to bind together disparate and ever-evolving applications for interoperability. Within caBIG™, vocabularies and common data elements as well as architecture are developed to support other research domains.

Within the area of biospecimen research, caBIG™ has a number of objectives based on its high level of interoperability, including creating virtual biorepositories and supporting multisite studies, among others.[7] Dr. Fore noted that institutions may follow any of multiple pathways to caBIG™ compatibility: Adopt caBIG™ tools, map an existing tool to caBIG™ tools, or make an existing tool caBIG™ compatible for standard reports only. He briefly highlighted two documents available at the OBBR Web site that are useful for implementing caBIG™ at biospecimen resources:

- *Implementing caBIG™ for Biospecimen Resources: An Overview*
- *Implementing caBIG™ for Biospecimen Resources: Next Steps*

Dr. Fore then briefly described three core caBIG™ biorepository and pathology tools:

- *caTissue Core*: Biorepository management infrastructure that supports the key functions of biospecimen resources; i.e., inventory management

[7] These caBIG™ objectives are further explained in biospecimen resource–specific materials available at http://biospecimens.cancer.gov/practices/forum/.

- *cancer Text Information Extraction System (caTIES)*: Supports importing free-text information from pathology reports into a structured biospecimen resource system through automated extraction of free text
- *caTissue Clinical Annotation Engine (CAE)*: Supports the addition of clinical information associated with biospecimens

Each of these tools is open-source software available as a free download at the caBIG™ portal under the Tissue Banks and Pathology Tools Workspace domain.[8] Some specialized IT skills are required to adopt caBIG™ tools or to make an existing tool caBIG™ compatible, but installation and use do not require hiring a full-time staff or investing in an IT laboratory.

In closing, Dr. Fore informed participants that successful development, adoption, and use of new informatics systems at biospecimen resources is complex and requires a high level of commitment. To this end, the NCI is augmenting its traditional customer support with the caBIG™ Enterprise Support Network, a collection of four new offerings: Service Providers, Knowledge Centers, Program Offices, and Enterprise Adopters. In addition to ongoing tool development, adoption, and workspace participation, these new programs will form a support network that will expedite and increase the integration of caBIG™ technology into scientific and clinical workflows at cancer and academic medical research centers and pharmaceutical and biotechnology companies.

Cost Recovery Models and Other Economic Issues Involved in the Implementation of the *NCI Best Practices*
Jim Vaught, Ph.D., Deputy Director, OBBR, NCI
Lisa Miranda, Technical Director, Tumor Tissue and Biospecimen Bank, University of Pennsylvania

Dr. Jim Vaught has consulted internationally on the development of biobanking networks and was a founding member of the International Society for Biological and Environmental Repositories (ISBER), serving as its second president and helping to develop ISBER's Best Practices for Repositories. Since 2005 he has been working in the OBBR on the development of the *NCI Best Practices for Biospecimen Resources* and has taken a leading role in other OBBR and National Institutes of Health (NIH) initiatives. Dr. Vaught is a member of several professional societies; Senior Editor for Biorepository and Biospecimen Science for *Cancer Epidemiology, Biomarkers & Prevention*; and a member of the editorial board of *Cell Preservation Technology*.

Ms. Lisa Miranda is Technical Director for the Tumor Tissue and Biospecimen Bank (TTAB) at the University of Pennsylvania. Ms. Miranda has served as the prime architect for the development of this new core facility and is responsible for all daily operations. She also has contributed significantly to the development of NCI's caBIG™, working with the biomedical informatics group at the University of Pennsylvania along with the NCI adopters-developers team. In 2006, Ms. Miranda developed an economic business model for biospecimen resources

[8] Open-source software is available as a free download at the caBIG™ portal under the Tissue Banks and Pathology Tools Workspace domain (https://cabig.nci.nih.gov/workspaces/TBPT/).

that promotes cost recovery and focuses on sustainable development. She is currently active in ISBER as a member of the strategic planning and education and training committees.

Background and Overview

Dr. Vaught explained that the OBBR is presently seeking to better understand several economic issues related to biospecimen resources based on public comments received about the *NCI Best Practices*. The OBBR is trying to determine the overall economic value of biospecimen resources that are accessible to the research community. This effort is timely because current NIH budget limitations are encouraging the NCI leadership to employ a more comprehensive approach for controlling costs for intramural and extramural biospecimen resources. There are also concerns from the scientific research community about how to effectively recover additional costs associated with implementing the *NCI Best Practices*. Dr. Vaught explained that the need to control costs must be balanced against the value of preserving annotated biospecimen collections over long periods of time.

Dr. Vaught noted that an economic analysis of biological resource centers (BRCs), which are somewhat analogous to biospecimen resources, demonstrated that BRCs amplified the impact of scientific research by enabling future generations to build upon past discoveries.[9] Similar to biospecimen resources, BRCs were found to serve several important functions including:

- Authenticating materials to ensure quality
- Preserving materials over long periods of time that may have future value
- Providing access to materials for the research community
- Creating economies of scale for larger biospecimen resources

However, it also was shown that maintaining BRCs may be challenging and costly when viewed from the perspective of the individual institutions that house them.

Although cost recovery models are often proposed as a viable mechanism for individual institutions to maintain valuable biospecimen resources, there are several challenges associated with their use. Determining and monitoring the true costs associated with operating biospecimen resources employing a variety of funding models is often complex. As such, it is difficult to ensure that user fees do not exceed cost recovery, which has the potential to limit access to biospecimens for future generations of researchers.

Biobanking Cost Recovery

Ms. Miranda opened her presentation on cost recovery with a brief description of TTAB, a newly established core facility created with the intention to centralize the institution's biospecimen resources. After 2.5 years of existence, the facility is servicing approximately 20 protocols, a number that will expand with the opening of a new laboratory space anticipated in February/March 2008. Ms. Miranda explained that to become a core facility, the resource must prove financial viability; thus, with only modest startup funding, cost recovery is vital to the resource's continued existence. The resource provides a full range of services with as much as 90 to 100 percent of its costs recovered in user fees. Resource user types are individual investigators, departmental banks, and external institutions for which TTAB functions as a

[9] Furman JL, Stern S. "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research." 2006. National Bureau of Economic Research (NBER) Working Paper 12523.

virtual resource. TTAB is using NCI's caTissue as its primary biospecimen inventory system to support these users.

After reviewing the basic elements and value of cost recovery as a business tool for biospecimen resources, Ms. Miranda focused on four critical questions:

- What does fee-for-service really mean?
- What does it take to be revenue neutral?
- What are the pros and cons of cost recovery?
- How can the implementation of cost recovery support the *NCI Best Practices*?

At most nonprofits, a fee-for-service model of social enterprise is used to provide services directly to clients.[10] The social enterprise achieves financial self-sufficiency through the fees charged for these services. The income received is used as a cost recovery mechanism for the organization to recoup its expenses. For many biospecimen resources, expenses often include salaries and benefits, capital depreciation, service contracts, and supplies and consumables.

Some biospecimen resources, such as TTAB, are required to achieve revenue neutrality in their fiscal year budgets. This requires them to be able to accurately track, allocate, and recoup all direct and overhead costs so that expenditures are in line with revenues. Ms. Miranda emphasized that cost analysis is paramount to recovering costs with this aim. She outlined a 12-step method for performing the analysis.

Cost recovery can be an effective tool to support resource planning and aid biospecimen resource growth. It also can be used to assist with financial and executive reporting. However, there are some potential pitfalls with use of a cost recovery model. With proactive planning, these pitfalls can be avoided. It is important to ensure that the pace of biospecimen research is not impeded through protracted business processes. There is also a risk that access to biospecimens may become more limited to those in the research community who are unable to pay user fees. One solution is to encourage transparency with end users about the reasons for cost recovery and to encourage stakeholders to promote a culture of understanding and support in this regard.

Cost recovery also can be used to support the *NCI Best Practices* by aiding in the quantitative evaluation of three critical questions:

- How effectively has the biospecimen resource performed?
- What impact has the biospecimen resource had on research?
- Is there a continuing need for the biospecimen resource?

The answer to each of these questions contains a financial-analysis component. Without accurate cost recovery models, it is impossible to provide stakeholders with the necessary information to assess existing biospecimen resources operations or proactively plan for future facility growth.

Ms. Miranda recognized that cost analysis is labor intensive but argued that the time demands are well worth the benefits. She also acknowledged challenges in billing and described success with transparent invoicing, building quality assurance into the price structure, and providing

[10] Source: http://www.virtueventures.com/setypology.pdf (accessed Feb. 9, 2008).

mechanisms for collaborators to continue funding activities that the resource otherwise would be unable to do. Ms. Miranda closed by endorsing this form of sustainable development as a way to achieve financial security for biospecimen resources in uncertain economic times.

Questions-and-Answer Session

One participant commented that he often feels conflicted when asked to share his expertise in examining tissue samples without charge. He asked for suggestions on possible responses that are in line with a cost recovery model. Ms. Miranda responded that the first step would be to determine what the actual costs associated with his service would be so that any concern about overcharging would be eliminated. It also would be important to make department chairs aware that this type of uncompensated activity takes time away from clinical responsibilities. Dr. Vaught added that the 100-percent cost recovery model discussed earlier might not be appropriate at every institution and in every situation. He stressed that even partial cost recovery would be a big step forward. Ms. Miranda concurred with Dr. Vaught that any form of cost recovery can be helpful and a best practice.

Another participant asked how competing biospecimen resources are dealt with and the ramifications to the cost recovery budget. Ms. Miranda replied that within her institution there are no other large competing biospecimen resources but her institution does have a tissue procurement service, the Eastern Division of the Cooperative Human Tissue Network or CHTN. One of her main challenges is recruiting customers who are accustomed to using the various principal investigator banks for free. In the case of the Cooperative Human Tissue Network, a network established to provide increased access to human cancer tissue for basic and applied scientist from academia and industry to accelerate the advancement of discoveries in cancer diagnosis and treatment. She suggested a few strategies on how to coexist with competing resources: Aiming to provide complementary services and to harmonize goals and resources for the greater good.

A Webcast participant asked if there is a plan to have a single common application form for caBIG™ so that researchers can simultaneously request tissue samples across multiple biospecimen resources. At present, researchers are required to fill out many and varied applications to different biobanks to acquire enough samples for a single research project. Dr. Fore responded that future iterations of caTissue available on the Web should be able to address this problem through their ability to execute federated queries to various biobanks. He also commented that a next generation biospecimen resource locator is being developed by the OBBR that would leverage Google for connecting researchers with various tissue specimen resources throughout the world.

Another Webcast participant asked how caTissue would prioritize similar specimen requests submitted simultaneously. Dr. Fore replied that caBIG™ software does not replace specimen custodians' local autonomy but offers mechanisms for federated queries and brokering between different centers to help fill requests.

IV. PART 3: NEXT STEPS FOR UPDATING THE NCI BEST PRACTICES

Assessing the Effects of Preanalytical Variables on Molecular Research: The Biospecimen Research Network

Helen Moore, Ph.D., Biospecimen Research Network Program Manager, OBBR, NCI

As a member of the OBBR, Dr. Helen Moore serves as Director of the Biospecimen Research Network (BRN), a new NCI research program whose primary aim is to sponsor, conduct, and collaborate on scientific studies of how biospecimen collection, processing, and storage variables influence the molecular integrity of those biospecimens. Dr. Moore has a broad background in research and product development and worked in the Human Genome Project at Celera Genomics before joining the NCI.

Dr. Moore asserted that translational research will advance molecular medicine and lead to personalized patient care. High-throughput technologies critical to translational research, such as genomics and proteomics, require high-quality, well-annotated human biospecimens. She reiterated the myriad adverse consequences of poor biospecimen quality and enumerated many preacquisition and postacquisition variables that can detract from biospecimen quality. The BRN takes a comprehensive approach to improving biospecimen quality by developing, promoting, and implementing evidence-based best practices. The BRN is improving accessibility to existing evidence on how biospecimen variables affect molecular analyses through the Biospecimen Research Database and the upcoming 2008 BRN symposium, Advancing Cancer Research Through Biospecimen Science.[11,12] Through the BRN, biospecimen research needs are being identified for new extramural programs, and associated research is being conducted by the BRN Intramural Laboratory. Molecular analysis technology development also is being supported by the NCI Innovative Molecular Analysis Technologies, or IMAT, Program via a request for applications on innovative technologic solutions for cancer sample preparation. Finally, the BRN is establishing strategic partnerships with organizations such as the CAP to develop data-driven, specimen-specific, platform-appropriate SOPs that will be incorporated into the CAP's accreditation programs.

The BRN will conduct comprehensive studies to identify the impact of tissue preservation variables on a biospecimen's molecular profile. To improve prospective biospecimen collections, the BRN aims to define the most significant variables for the collection of tissues, blood, and body fluids. One such study, conducted in collaboration with the NCI Clinical Proteomic Technology Assessment for Cancer program, will address issues involved in prospective blood collection and plasma processing as well as the development of evidence-based biospecimen quality indicators to assess the usability of archival plasma collections.

Dr. Moore concluded by stating that the BRN contributes to the evolution of biospecimen resources by developing and implementing state-of-the-science processes that ensure the molecular integrity and clinical relevance of human biospecimens used in cancer research and

[11] More information on the Biospecimen Research Database is available at http://brd.nci.nih.gov/BRN/brnHome.seam.

[12] More information on the Advancing Cancer Research Through Biospecimen Science symposium is available at http://www.brnsymposium.com/meeting/brnsymposium/.

clinical medicine. She also invited attendees to respond to a recently issued request for information regarding tissue preservation variables.[13]

The Biospecimen Research Database
Ian Fore, D.Phil., NCI Center for Bioinformatics

Dr. Fore described the Biospecimen Research Database, a Web-based tool that tabulates information regarding the effect of biospecimen handling variables on experimental results within various analytical platforms. The database is being populated with evidence that has been expertly curated from published studies, unpublished results, and ongoing BRN experimentation. Future steps for the database include meta-analyses of biospecimen handling data to help define the state-of-the-science in biospecimen research, development of evidence-based SOPs, and an online library of biospecimen protocols.

The Biospecimen Research Database matrix comprises specimen types on one axis and analysis platforms on the other. Cells at the intersection of these two axes show the number of pertinent studies that have been entered in the database to date. By clicking on a cell, researchers can access citation information and PubMed links to relevant publications, structured study data, and free-text entries about the purpose and conclusions of the study as well as individual study findings. Users also can locate data by searching the structured values in the database for particular terms.

Dr. Fore requested input from the audience on the level of interactivity of the database. One possibility would be to enable Web 2.0 mechanisms; i.e., open-access Wiki- or forum-like input capability with minimal oversight. Another option would be controlled access to data entry, which would provide more concise analyses of the evidence. He requested that participants contact the OBBR to recommend key scientific papers and protocols and/or to volunteer their assistance.

Question-and-Answer Session

A participant asked whether the Biospecimen Research Database incorporates information about the number of times an experiment was replicated or other measures of study quality. Dr. Fore replied that such parameters are not currently part of the database; however, as more meta-analysis is performed, that depth of analysis may be incorporated. Dr. Moore added that if a paper has relevant findings, it will be entered into the database. The OBBR anticipates more community input as protocols are disseminated via the Biospecimen Research Database.

A Webcast participant submitted a question concerning the recommended time between tumor excision and flash freezing for RNA analysis. Dr. Moore responded that this should be as brief as possible and requested input from the audience. One attendee replied that biospecimen researchers would appreciate hard numbers from the OBBR in answer to such questions to help ensure consistency between different institutions. Another postulated that the shortest time or the optimum time will depend upon the tissue in question; it will be useful to pathologists to know

[13] The RFI Notice of Request for Information on Cancer and Normal Tissue Acquisition and Processing Variables is available at http://grants.nih.gov/grants/guide/notice-files/NOT-CA-08-002.html

the acceptable timeframes for different tissues. Ms. Miranda added that input is needed regarding how the NCI could best disseminate any SOPs they develop.

Another Webcast participant asked about sources for consensus protocols. Dr. Moore responded that many are available in the public domain and that the OBBR intends to offer them for download from the Biospecimen Research Database in the future.

V. PART 4: CONCURRENT BREAKOUT SESSIONS

At this point in the program, participants were asked to attend one of three breakout sessions on the following topics: (1) Cost recovery models and other economic issues involved in implementing the *NCI Best Practices*, (2) caBIG™ biospecimen resource management tools, or (3) biospecimen science. Detailed summaries of the breakout session presentations and discussions are found in the appendix.

VI. PART 5: CLOSING AND ADJOURNMENT

Breakout Session Reports

Discussion of Cost Recovery Models and Other Economic Issues Involved in the Implementation of the NCI Best Practices

Jim Vaught, Ph.D., OBBR, NCI; Martin L. Ferguson, Ph.D., Consultant, Pharmaceutical and Life Sciences, Consultant and Cofounder of Ardais Corporation; Lisa Miranda, Tumor Tissue and Biospecimen Bank, University of Pennsylvania

Presented by Jim Vaught

Biospecimen resources may employ a range of cost recovery models, partial to full, and different public or private funding models may be appropriate, depending upon the institutional goals. The most appropriate cost recovery model will depend on the biospecimen resource model. Regular communication with resource customers is critical for effective implementation of cost recovery.

Resource managers would appreciate a compilation from the NCI of practical case studies featuring institutions implementing the *NCI Best Practices*, especially those institutions that have adopted cost recovery business practices. Social science research indicates that biospecimen resources have substantial economic value.[14] The NCI should consider conducting a cost-benefit study of NCI-supported resources to evaluate whether those adopting the *NCI Best Practices* add to the accumulated knowledge base more effectively. Logic dictates that resources that adopt standard practices are more cost effective and valuable; perhaps the NCI could publish concrete examples of this, with confirmation from economists. Conversely, it may also be helpful to publish case studies that exemplify the adverse effects and costs incurred by failing to follow the *NCI Best Practices*.

[14] Furman JL, Stern S. "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research." 2006. National Bureau of Economic Research (NBER) Working Paper 12523.

Discussants reached a consensus that future iterations of the *NCI Best Practices* should provide more guidance on economic issues and cost recovery, which is the best insurance of long-term repository survival. OBBR staff members are considering holding a workshop to focus on these issues.

Ms. Collyar added from the audience that pairing the NCI with senior extramural investigators could lead to a win-win situation: Investigators share their evidence-based SOPs with the NCI and the Institute provides, through a formal review process, funding to both maintain their biospecimen resources and to have these repositories become implementation models of the *NCI Best Practices*.

Demonstration of caBIG™ Biospecimen Resource Management Tools
Ian Fore, D.Phil., NCI Center for Bioinformatics

Released in June 2007, caTissue version 1.2 addressed usability issues and was more user friendly than earlier versions. The key feature is that each user has access to different system parameters depending on his or her user status. Approximately 20 institutions employ caTissue for day-to-day resource management. The next generation, caTissue Suite, will be released in March 2008 and allows users to decide which modules to install. Other innovations include allowing the addition of fields, integrating caTissue Core with other informatics components, such as caTIES and the CAE; enabling consent tracking; saving and repeating complex queries for nonexpert users; and mechanizing biospecimen ordering, shipping, and receiving.

It is not the intent of the caBIG™ community to make caTissue use mandatory, nor is it intended to replace more complex systems such as FreezerWorks. Rather, caTissue is intended to interoperate with such systems. It was designed to allow more sophisticated biospecimen tracking than simple database programs such as Excel. The caBIG™ community is an open one, seeking participation from all stakeholders and input on new features, such as adding the ability to enter the dimensions of tissue blocks. Dr. Fore invited input in advance of an upcoming meeting at which it will be decided what features the next iteration will include.

Biospecimen Science
Richard G. Hegele, M.D., FRCPC, Ph.D., Professor and Head, Department of Pathology and Laboratory Medicine, University of British Columbia; Beatrice Knudsen, M.D., Ph.D., Affiliate Assistant Professor of Pathology, University of Washington School of Medicine; Lawrence D. True, M.D., Professor of Pathology, University of Washington School of Medicine

Presented by Beatrice Knudsen

The objective of the biospecimen science discussion session was to identify ways in which biospecimen science can be used to inform the *NCI Best Practices.* The conclusions will help elucidate the next steps for the NCI and the rest of the research community.

The brief presentations of the discussion leaders focused on different variables influencing biospecimen quality. Preacquisition and postacquisition variables can be recorded and their effects on the experimental outcome identified. For example, factors such as participant diet,

biospecimen warm-ischemia time, tissue fixation time, and immunohistochemical technique can have a major impact on gene expression. Identifying the influence of such factors will allow researchers to determine which observations are artifacts due to extraneous sources of variation and subtract such findings from relevant scientific observations.

The presentations also covered tools to advance biospecimen science and cancer research. The British Columbia BioLibrary is an excellent example of how a collection of biospecimens can be integrated with bioinformatics to yield a high-quality resource. The Minimum Information Specification For *In Situ* Hybridization and Immunohistochemistry Experiments represent the minimum published information needed for an independent researcher to reproduce experimental results.[15]

The resultant discussion concerned the most pressing needs that must be addressed to enable improved quality of banked biospecimens in light of limited resources and the urge to generate publishable data. Discussants also identified barriers to the implementation of *NCI Best Practices* protocols and ways to circumvent them.

Closing Remarks

Dr. Compton thanked attendees for participating in this landmark event, stating that biospecimen science is at a tipping point: The vision of personalized medicine cannot be realized until the issues discussed today are resolved. The culture of cancer research cannot be changed without the cooperation of stakeholders to address the multifactorial issues that comprise obstacles. She likened personalized medicine to personal computers: At this time personalized medicine seems like an unattainable goal, but someday it will be the norm, just as a decade or two ago the personal computer seemed like a luxury. Standardizing biospecimen handling will be as transformational to medicine as standardizing the semiconductor was to information technology.

Dr. Compton thanked attendees for their participation and urged them to submit feedback via e-mail or phone. She also recognized members of the OBBR staff whose hard work saw this round of forums to fruition.

[15] More information on the Minimum Information Specification For *In Situ* Hybridization and Immunohistochemistry Experiments http://is available at scgap.systemsbiology.net/standards/misfishie/.

APPENDIX

BREAKOUT SESSION SUMMARIES

Discussion of Cost Recovery Models and Other Economic Issues Involved in Implementing the NCI Best Practices

Jim Vaught, Ph.D., Deputy Director, OBBR; Martin Ferguson, Ph.D., Pharmaceutical and Life Sciences Consultant and Cofounder of Ardais Corporation; and Lisa Miranda, Technical Director, Tumor Tissue and Biospecimen Bank, University of Pennsylvania

Dr. Martin Ferguson served as Director of Bioinformatics at Axys Pharmaceuticals and Senior Vice President of Bioinformatics and Cofounder of Ardais Corporation, a company whose mission was to build and deploy annotated biospecimen collection systems. Dr. Ferguson is now an independent consultant to pharmaceutical and life science informatics companies and Federal agencies such as the National Cancer Institute (NCI).

Dr. Jim Vaught opened the discussion by asking if any participants had questions relating to the morning session dealing with cost recovery. Since there were no questions, he introduced Dr. Ferguson, an NCI consultant who also cofounded Ardais Corporation, one of several for-profit tissue brokers established in the early 2000s. Dr. Ferguson has extensive experience in commercial biobanking and brings an industry perspective to the issue of cost recovery.

Dr. Ferguson told participants that for-profit entities typically apply two basic models when funding biospecimen resources through cost recovery mechanisms: (1) Grants in which they have no expectations about what they will obtain for their investment and (2) contracts with specific expectations about work to be performed or tissues to be delivered without intellectual property (IP) concerns.

If the more common grant model is applied, the research and data analysis is typically done onsite (e.g., hospital or clinic) with the funder having the right of first review to the data and an option to the IP. Publications are ultimately permissible after providing time for any patents to be filed. When the second contract model is used, tissues specimens are purchased outright at the institution's cost without any IP reach through. In this second, more direct model of cost recovery, managing ethical issues becomes critical. Cost recovery is permissible, although not for salaries of full-time employees who are in decisionmaking roles governing access to biospecimens. Biospecimen resources also should be careful to apply true cost recovery consistently and fairly across nonprofit and for-profit entities. In the model used by Ardais, profits were allowed to flow back to the academic institution as long as those profits did not go to support any administrative entity related to the biospecimen resource operation. Profits could be used, however, to fund the general mission of the academic institution.

Ardais divided the biospecimens into three tiers to incentivize academic medical centers to participate. The first tier, the clinical site pool, was made up of biospecimens paid for by the company but specifically reserved for the sole use of the institution. The second tier, the shared contributing pool, was made up of biospecimens that were made available only to nonprofits that were contributing biospecimens to the pool. The third tier, the commercial pool, was the largest and operated on a first-come, first-served basis. There was a single price for biospecimens, and academic medical centers could purchase samples from the commercial pool as well.

Dr. Ferguson concluded his presentation by telling participants that there are many opportunities for academic medical centers to work with industry to provide biospecimens because the demand for tissues in preclinical use is increasing rapidly. These types of business relationships can be managed from a bioethical perspective and can augment the funding of existing institutional biobanking programs.

Ms. Lisa Miranda gave a brief overview of the cost recovery process at the Tumor Tissue and Biospecimen Bank (TTAB) at the University of Pennsylvania. She explained that in the first year of operation, the cost model was applied to users on a per-sample basis. It later evolved to include a cost-per-project model in which costs are customized based on user and project type. Written quotes indicating total costs are provided to investigators who wish to contract with the TTAB. Billing is done monthly as biospecimens are actually collected.

User fees typically include charges for:

- Labor, both direct and general/administrative
- Direct materials
- General laboratory supplies
- Service contracts
- Capital depreciation

Question-and-Answer Session

A participant asked Dr. Ferguson about the review process used at Ardais for obtaining biospecimens. Dr. Ferguson explained that the standard processes involving institutional review board (IRB) approval and tissue utilization committees were used when biospecimens were transferred from academic institutions to Ardais. Once inside the corporate biorepository, however, samples were distributed on a first-come, first-served basis without a review of scientific merit.

Another participant asked Dr. Ferguson which types of biospecimens were brokered by Ardais, whether different prices were used for different sample types, and if a sliding scale was employed. Dr. Ferguson said that toward the end of the company's life, it brokered tissue and blood for DNA purposes but not serum or plasma for protein purposes. He explained that no sliding scale was used, but prices varied based on the type of tissue and degree of annotation.

A participant asked Dr. Ferguson what a customer actually received when paying Ardais $2,000 to $3,000 for a biospecimen. Dr. Ferguson stated that the price included fresh-frozen material composed of two to fifteen 500 mg cubes, a hematoxylin and eosin–stained section from what would have typically been a diagnostic block, and a snapshot of the patient's clinical information and medical history.

One participant asked Dr. Ferguson to comment on why Ardais failed. Dr. Ferguson responded that the company raised $65 million in venture funding over 5 years and generated millions in revenue during its 7 years of operation. He clarified that Ardais had profitable quarters but never a profitable year. Dr. Ferguson said that he thought the biggest problem with the company was the fact that it was structured as a for-profit entity. Had it been a nonprofit, it might have been

successful. Although it would have been unable to raise venture capital, it would have avoided the mixed feelings many of those at academic medical centers experienced upon routing biospecimens through a private company. If the company had focused on collaborations with the cancer centers in which a clinical question was being addressed, as opposed to simply being a broker, it would have had a better chance of surviving because researchers might have spoken about the company in a more positive voice.

One participant asked Ms. Miranda if patients are required to sign a consent form indicating that their biospecimens can be sold to industry. Ms. Miranda explained that the TTAB currently does not provide services to industry. At some point, it will be possible to foster industry collaboration, at which time costs may be recovered, but never will specimens be sold, and the informed consent form specifically states that specimens will not be sold.

A participant asked Ms. Miranda if she services the needs of investigators who are interested in collecting specific types of tissues as well as providing specific types of tissues that already are part of the collection to other investigators, both on a cost recovery basis. Ms. Miranda indicated that she currently provides both collection and tissue allocation services. The model used is that of a combined biobank and service center.

Another participant asked Ms. Miranda which consent form is used at the TTAB. Ms. Miranda stated that they have their own umbrella protocol, which includes consent for identified collection of biospecimens as well as permission to collect deidentified biospecimens without consent.

A participant asked Ms. Miranda whether the money to fund the TTAB comes entirely from projects or is subsidized by other institutional funds. Ms. Miranda stated that approximately one-half of this year's revenues will be generated by the service center, and the remainder will be provided by stakeholder support. In the future, more revenue will need to be generated by the service center operation.

A participant asked Ms. Miranda to explain the ownership of biospecimens paid for by investigators and collected by the TTAB. She also asked about the existence of oversight regarding conservation of rare tissues. Ms. Miranda stated that the policy of the TTAB is that if an investigator pays for a biospecimen, he or she has control over its use, although there is a tissue utilization committee and peer review process for reviewing the scientific merit of the study and monitoring access decisions. As a followup question, the participant asked whether the consent form addresses data sharing and placement of data in a national biorepository. Ms. Miranda stated that the present repository consent form aims to protect patient's privacy by specifying who will be able to view the data, but it does not directly address the issue of data sharing beyond the institution. She indicated that the TTAB would like to develop a separate data-sharing agreement for this purpose.

A participant asked Ms. Miranda if pathologists are reimbursed through cost recovery for their time in making decisions on whether there is excess tissue available for research use after the clinical needs have been met. Ms. Miranda indicated that there is not presently money available to give small stipends to pathologists as an incentive to make the decisions and collect tissues for

research. She also noted that there are bioethical issues that must be addressed when granting financial incentives to those participating in research efforts. As a followup question, the participant asked whether the NCI has a policy for allowing inclusion of cost reimbursement for pathologists providing tissue for research as a part of the grant budget. Dr. Vaught added that NCI grants do not routinely cover the costs of collection and allocation of tissues.

Dr. Vaught concluded the discussion session by presenting two summary slides developed from presentations and comments made at previous forums held in Boston and Chicago.

Demonstration of caBIG™ Biospecimen Resource Management Tools
Ian Fore, D.Phil.

Dr. Ian Fore opened the demonstration session by informing participants that three versions of caTissue Core had been released to date, with the most recent (version 1.2) released in June 2007. This version, which was developed by Washington University in St. Louis, underwent testing at four funded "adopter" universities and is now in use at approximately 20 institutions. Enhancements from the previous version are primarily around usability and include easy access to edit data from search screens, support for a study calendar, the ability to propagate collection values for all biospecimens in a group, and a more intelligent storage system.

Before beginning a step-by-step demonstration of caTissue Core using mock data, Dr. Fore explained its general organization. A patient will be entered into the system then registered to a protocol and a biospecimen collection group—one patient might be part of many collection groups. From there, biospecimen data can be entered. Biospecimen events (e.g., centrifugation or slicing) are recorded by the software such that the parent-child relationships of each biospecimen are maintained for the resultant units. The software also will track the physical location in which the biospecimen is stored and the transfer of biospecimens to other researchers.

Dr. Fore explained that caTissue Core is a Web-based application built on Java programming language that runs on MySQL and Oracle database management systems. The software runs on a central server at the institution and can be accessed by registered users at their personal computers with usernames and passwords. caTissue is now caBIG™ compatible at the gold level, as it has achieved computable semantic interoperability and is grid compatible. Dr. Fore then initiated a step-by-step caTissue demonstration. Each user logs into the system with a user name and password and will have access to different system parameters depending on user status. Administrators, for example, are super users; they are able to define collection protocols and the types of biospecimens and processing events that will be logged for a particular study. Data entry personnel would not have access rights to define protocols but would be able to enter information about individual patients, biospecimens, and events. Screen by screen, Dr. Fore showed participants the vast array of biospecimen data that can be recorded and tracked with caTissue Core.

Dr. Fore concluded that caTissue version 1.2 is user friendly and practical for managing day-to-day operations at a biospecimen resource. He then mentioned several caTissue Suite features, which will be available in late 2007 or early 2008, including informed consent tracking; advanced methods for saving routine queries; the ability to add data entry fields; integration with

cancer Text Information Extraction System (caTIES); and tracking of biospecimen ordering, shipping, and receiving.

Question-and-Answer Session

In response to questions on caBIG™ and achieving caBIG™ compatibility, Dr. Fore explained that the caBIG™ project is establishing an IT infrastructure for sharing data—including biospecimen data—across a network of cancer researchers. It supports the development of information systems based on input from a community of over 1,000 individuals working in the domains of clinical trial management systems, integrative cancer research, tissue banks and pathology tools, and *in vivo* imaging. caBIG™ also is developing vocabularies and common data elements and architecture to support other research domains. The caBIG™ project is engaging a variety of stakeholders ranging from institutions with integrated, IT staff–supported systems to those with informal or no information systems. The caBIG™ approach is to develop modules that address specific needs, connect through defined electronic interfaces, and use international data standards. caBIG™ focuses on the boundaries and interfaces between applications rather than the use of specific types of applications. Dr. Fore noted that there are multiple pathways to caBIG™ compatibility for an institution: To adopt caBIG™ tools, to map an existing tool to caBIG™ tools, or to make an existing tool caBIG™ compatible for standard reports only. Workshops also are available to help interested parties achieve caBIG™ compatibility. Furthermore, a caBIG™ Enterprise Support Network is being formed to provide greater support and services to biomedical research organizations interested in using caBIG™ technology.

A participant inquired whether a data dictionary exists that is updated regularly. Dr. Fore replied that there are two components to such a data dictionary: The NCI Thesaurus and the NCI Metathesaurus (both of which list terms and definitions) and the cancer Data Standards Repository, a database and tool set used to create, edit, and deploy common data elements.[16] He also added that, to achieve caBIG™ semantic interoperability, a tool is available to the community that reviews the attributes in one's database and compares it to caBIG™ data standards.

Several specific software-operation questions were asked throughout the presentation. Dr. Fore demonstrated entering several biospecimens at one time; labeling and tracking multiple aliquots derived from a single specimen; accessing information (amount and type) based on user role and responsibility; clicking on a sample in the software specimen location map to access information on a particular biospecimen; and illustrating biospecimen storage information in three dimensions; e.g., which freezer shelf and the position in X-Y coordinates on that shelf. With respect to the tracking of multiple specimen aliquots, Dr. Fore pointed out that the system currently does not support the recording of formalin-fixed, paraffin-embedded block dimensions and slice thickness and cannot derive the number of remaining slices. It does, however, support the entry of freezer specifications that allow the customization of specimen location information. A participant inquired about specimen identification codes. Dr. Fore replied that caTissue Core generates identification codes automatically but that the labeling scheme can be customized

[16] For more information on these vocabulary and common data element tools, visit the NCI Center for Bioinformatics Web site at http://ncicb.nci.nih.gov/NCICB/infrastructure/cacore_overview/vocabulary and http://ncicb.nci.nih.gov/NCICB/infrastructure/cacore_overview/cadsr.

through some programming at the user end. Such customizations will hold in future versions of the software as long as caTissue programming rules are followed. Dr. Fore also indicated that an improvement to the software would consist of providing template choices as in Excel where programming is no longer required to modify the default label setting. He also clarified that the specimen label need not be the same as the barcode; caTissue Core has both a label and a barcode data entry field.

Another discussant asked how duplicity of barcodes is avoided across different biospecimen resources. Dr. Fore recommended that preprinted barcodes, which are then scanned into the system, be used. An IT expert involved in the Early Detection Research Network (EDRN) also commented that multicenter studies within the EDRN are coordinated from a single site to ensure that each study has different identifiers. In many cases, specimens already exist and, therefore, already have identifiers. In those cases, some mapping is conducted and identifiers changed. Dr. Fore reminded participants that in addition to the barcode, caTissue generates an internal identifier that is unique within the system. He also informed participants of the possibility of providing caTissue users with an initial set of identifiers. Once those are exhausted, additional identifier sets could be obtained from an established Web service, an approach similar to that used to obtain chemical compound identifiers.

Dr. Fore then clarified that clinical data associated with a specimen can be entered into a series of predetermined fields in the Web application. Furthermore, caTissue Suite will support the addition of fields to define parameters of a particular study. He also indicated that output data (e.g., query results) into other programs, such as SAS and Excel, for analysis purposes is supported.

Other attendees asked about how caTissue Core compares to commercially available products. Dr. Fore replied that, while it is hoped that such commercial applications will become caBIG™ compatible, caTissue Suite is not intended to replace them. Rather, caTissue Core meets the needs of some institutions for a mature, robust biospecimen resource management tool to replace simpler systems such as a Microsoft Access database or an Excel spreadsheet.

In closing, Dr. Fore stated that caTissue Suite will address many of the concerns raised by participants. For example, it will permit local customization and the addition of unique data fields to a system without compromising caBIG™ compatibility. caTissue Core also allows the user to load information from a delimited file (i.e., Excel spreadsheet or other legacy system) as well as automatically load information into the biobanking application when patients are registered to a linked clinical trial system.

Biospecimen Science
Richard G. Hegele, M.D., FRCPC, Ph.D., Professor and Head, Department of Pathology and Laboratory Medicine, University of British Columbia; Beatrice Knudsen, M.D., Ph.D., Assistant Member, Fred Hutchinson Cancer Research Center and Affiliate Assistant Professor of Pathology, University of Washington School of Medicine; Lawrence D. True, M.D., Professor of Pathology, University of Washington School of Medicine

Dr. Beatrice S. Knudsen is an anatomic pathologist, an assistant member at the Fred Hutchinson Cancer Research Center, and an affiliate assistant professor at the University of Washington. Her current research focuses on identifying cancer biomarkers. She applies her expertise in tissue banking as a coleader of the specimen core of the Pacific Northwest Ovarian Specialized Program of Research Excellence (SPORE), a member of the Pacific Northwest Prostate SPORE tissue utilization committee, and a member of the Gynecologic Oncology Group tissue bank committee.

Dr. Knudsen introduced the objectives of the discussion:

- To examine problems with collected and banked biospecimens and how to overcome those problems
- To examine how preanalytical conditions affect biospecimen quality and to discuss the ongoing efforts toward understanding, recording, and standardizing conditions that affect measurements in biospecimens
- To prioritize the improvements that are needed for the collection and banking of high-quality biospecimens

Postacquisition Variables

Dr. Knudsen showed a schematic of the lifecycle of a biospecimen, from acquisition through handling and processing, storage, and distribution, to analysis. While researchers have little control over preacquisition variability, careful records can help clarify their influence on the final results. Researchers *can* control many postacquisition variables. Dr. Knudsen is involved in a project with prostate SPOREs examining how fixative composition, fixation time, and immunohistochemical processing affect analytes. Several institutions recorded specimen processing time (which varied significantly between institutions) from ethanol to paraffin embedding. Samples were then sent to a central location and processed concurrently. Although results showed good correlations among sites for some normal prostate gland markers, significant variability was observed in tumor markers between institutions, which could have been accounted for by differential processing.

Preanalytical Variables

Dr. Richard Hegele is a professor and Head of the Department of Pathology and Laboratory Medicine at the University of British Columbia (BC). His service responsibilities are based at St. Paul's Hospital, where he is a Principal Investigator of the James Hogg iCAPTURE Centre for Cardiovascular and Pulmonary Research. Prior to his current post, Dr. Hegele served as Program Director of the Anatomical and General Pathology residency training programs at the University

of British Columbia and as Acting Head of the university's Department of Pathology and Laboratory Medicine.

Dr. Hegele discussed information dissemination using examples of private knowledge, a targeted audience, and an expanded audience. First he described a study he had done on lung disease in guinea pigs in which he determined that processing and handling the lungs differently will lead to differential RNA preservation. However, he had difficulty finding a journal that would publish this example, so the experimental outcome remained private knowledge.

Next he discussed an experiment looking at the influence of storing breast cancer tissue on ice for zero to 24 hours prior to freezing at -70º. He found that c-*myc* and estrogen receptor mRNA levels remained relatively unchanged for the first 3 hours then degraded significantly up to 24 hours. The conclusion was that timing of collection can influence measured levels of mRNA expression. Therefore, documenting timing and standardizing protocols were recommended. This information reached a targeted audience when it was published in a specialized journal.

The BC BioLibrary was designed to expand the audience reached by such experimental results as those in the above examples.[17] An evaluation of biobanks in British Columbia revealed just what has been found in biobanks in other regions of the world: Variable access to adequately collected and annotated biospecimens, inadequate synergy between biobanks and research collections, lack of standardization in collection of materials or material quality, and a gap between the pace of research and improved health outcomes. One solution was formation of a BioLibrary, defined as a collection of biospecimens organized to provide access to biospecimens, associated data, and related biobanking knowledge to a target group—biobanks, clinical trials, and research studies—and modeled after a public library. As opposed to a super biobank, it is a facilitator, enhancing the quality and accessibility of biospecimens and hence the dissolution of data silos, helping pathologists streamline and improve banking protocols, and contributing to sustainability of biobanking in British Columbia by developing and upholding public trust. The BC BioLibrary comprises operational units that integrate pathology, clinical, and biobanking concerns and catalog units listing inventory matched to research participant intent as well as a standard operating procedure (SOP) library. Information technology (IT) maintains transparency through public engagement and its oversight structure includes several committees.

Dr. Hegele concluded that it is inadequate for investigators to create information without a strategy for disseminating it. The BC BioLibrary platform provides a means to expand the research audience, accelerate the research process, and fulfill the promise of translational research.

Assessing Tissue Quality

Dr. Lawrence D. True is Director of Urologic Pathology at the University of Washington School of Medicine. His research program focuses on male genitourinary pathology and biospecimen science. Since coming to the University of Washington, Dr. True has codirected the Program Project in Mechanisms of Prostate Cancer Metastases and the Pacific Northwest Prostate Cancer

[17] More information on the BC BioLibrary is available at http://www.bcbiolibrary.ca.

SPORE Specimen and Tissue Core, which encompasses a standardized system of biospecimen collection, storage, and distribution, and related clinical/research information dissemination.

Dr. True discussed an experiment designed to evaluate the effects of preacquisition variables on biospecimens. In this example, patients diagnosed with prostate cancer were asked to consume either a high-glycemic, high-caloric diet or a normal diet for 6 weeks. Evaluation of differential gene expression in frozen prostate tissue biopsies revealed up to a 13-fold increase in the most highly expressed genes in the high-glycemic, high-caloric patients over the patients consuming the control diet. In another experiment, the investigators evaluated warm-ischemia time on laser capture–microdissected tumor tissue and observed differential downregulation of certain genes in a time-dependent manner. Studies such as these will yield fingerprints of gene expression that allow for correction of variation based on preprocessing variables such as diet and warm-ischemia time, the result being a truer data set of the difference between normal tissue and cancer cells.

These and other experiments aimed to assess the influence of preanalytical variables on the preservation of macromolecules and clearly show that a number of variables can influence the experimental outcome. Recording those data will enable others to replicate investigators' results. A method to evaluate the reproducibility of published results, called Minimum Information Specification For *In Situ* Hybridization and Immunohistochemistry Experiments (MISFISHIE), encourages investigators to record such variables. A MISFISHIE evaluation of 32 journal articles revealed that a large percentage included insufficient details for an independent investigator to replicate the study.

Dr. True concluded that to determine if a molecular profile of experimental tissue is abnormal, for uncontrollable variables a molecular fingerprint can be developed then subtracted from the experimental gene expression profile. For controllable variables**,** optimized biospecimen handling protocols can be developed to minimize the effects of variables that most influence gene expression profiles. Obtaining all relevant information from all experimental data sets—i.e., MISFISHIE specification—would help to achieve these goals.

Question-and-Answer Session

The first question concerned difficulties in determining how much of an excised tumor would be required for clinical evaluation. Dr. True replied that most of the samples in his work are the result of radical prostatectomies. From there, six blocks of tissue are frozen in precooled isopentane, which will preserve histology, RNA, and proteins. Two sets of frozen sections are taken from each block, one set for the patient's diagnostic files and the other into the institution's research files. This results in as much diagnostic tissue histology information as would be obtained had no portions been devoted to research. Dr. Hegele responded that his institution does not currently use SOPs, and variable protocols result in a lack of consistency.

Another participant asked about the mean warm-ischemia time from excision in the operating room to preservation in a freezer. Dr. Hegele replied again that this is not consistent between institutions but that in general 90 percent of the samples are frozen within 15 minutes and 99 percent within an hour. Dr. True added that at his institution, the individual responsible for

freezing the tissue is in contact with and paid by the surgical team, and that samples are frozen within 15 minutes. The participant replied that he had been unable to freeze the samples he works with more rapidly than 15 minutes from excision; warm-ischemia time is recorded as a method of tracking biospecimen quality. The timing in his protocol is sufficient for the preservation of the bioanalytes of interest.

An audience member asked how the BC BioLibrary was funded. Dr. Hegele explained that the funding was provided by the Michael Smith Foundation for Health Research, a provincial government body in British Columbia that was in part modeled on characteristics of the Alberta Heritage Foundation for Medical Research. The BC BioLibrary is funded for 3 years with the possibility of renewal for up to 5 years. Funding is undetermined thereafter. BC BioLibrary staff have to produce progress reports and other measures of their deliverables for accountability and continuation of funding.

Dr. Compton commented on obstacles to implementing high-quality tissue collection. She stated that one obstacle is the surgeon, who can either facilitate the process or present the largest hurdle. The surgeon should have the mindset that the biospecimen is an extension of the patient and is critical to patient care. Therefore, he or she should take care of the biospecimen as he or she would take care of the patient.

A respondent described the formation of a biospecimens and clinical trials committee consisting of surgeons, oncologists, pathologists, and nurses at his institution. This committee meets monthly and works to educate everyone in the hospital about processing and storing biospecimens. They grapple with many issues, such as where to annotate fixation times and how to fix tissue within a 48-hour window. He said this is a huge task without standardized forms. Dr. True advocated minichip technology, with which a biospecimen can be tagged and which will record every biospecimen event; e.g., excision or freezing. Dr. Compton added that anesthesiologists record some preacquisition data but do not record factors such as clamp time, which need to be considered a part of basic patient care.

An attendee asserted that annotation of preanalytical factors is uncommon, and specifics are rarely reported in journals. Dr. Compton agreed that the contrast between the rigor expected in experimental measurements and the lack of rigor in reporting materials and methods is surprising. Another attendee replied that time-to-collection data might be reported as 30 percent of biospecimens were snap frozen within 30 minutes, another 30 percent within 60 minutes, and the rest within 3 hours. There is a lack of awareness in the scientific community of the influence such variability could have upon the experimental results.

An investigator asked how to evaluate quality of samples and whether it would be worthwhile to normalize for gene or protein expression. Dr. Compton suggested he could normalize for housekeeping genes, but the investigator replied that he was uncertain whether stability of the levels of housekeeping genes is a good assumption. Dr. True replied that he has observed expression of housekeeping genes change threefold to fivefold in some cell lines. He suggested that one possibility would be to use the patient as his or her own control, comparing the first biospecimen to subsequent biospecimens. The investigator added that another caveat would be whether the tumor is necrotic because necrosis causes altered gene expression.

A Webcast participant commented that variability in processing or storage of peripheral mononuclear cells can cause a great deal of variability in the signal-to-noise ratio in downstream assays. She asked whether anyone could recommend best practices for peripheral blood cells. Dr. Moore recommended that the participant might like to search the National Cancer Institute (NCI) Division of Cancer Epidemiology and Genetics Web site for that information. Another attendee added that at her institution, some of those issues are addressed by asking technicians to record processing times, but there have been problems with falsified records.