Comprehensive Data Resource

Introduction

Helen Moore and Ping Guan
Biorepositories and Biospecimen Research Branch
Cancer Diagnosis Program
Division of Cancer Treatment and Diagnosis
National Cancer Institute

July 23, 2015
Agenda

• Introductions
• Opening Remarks – BBRB
• Collaboration Mechanism – NCI Tech Transfer
• CDR Overview – Leidos Biomedical Research
• CDR Demo – Leidos Biomedical Research
• Questions and Answers
Biorepositories and Biospecimen Research Branch (BBRB)

- BBRB provides leadership, tools, resources, and policies in biobanking for the global biomedical research community, to enable translational research and precision medicine for patients.
Current Initiatives

- Biospecimen Preanalytical Variables Program (BPV)
- Genotype-Tissue Expression Program (GTEx)
- NCI Best Practices for Biospecimen Resources
- Biospecimen Evidence-Based Practices (BEBPs)
- Biospecimen Research Network (BRN)
- Biospecimen Research Database (BRD) – online literature and SOPs db
- Biobank economics research and online tools
- ELSI research in biobanking
- Patient brochures
Comprehensive Data Resource (CDR)

• CDR is part of the caHUB program.

• caHUB was not developed as a national biobank as initially envisioned, due to funding changes.

• However, two biospecimen programs were conducted under caHUB and are ongoing:
  – The Biospecimen Preanalytical Variables program (BPV) – a study of preanalytical variation in tissue processing and storage (FFPE and frozen tissues) and the effects of such variation on downstream molecular analysis.
  – The Genotype-Tissue Expression Program (GTEx) – a NIH Common Fund study of genomic variation and tissue-specific expression, analyzing up to 30 tissues per donor in 900 deceased donors.

• CDR serves as the Comprehensive Data Resource for these two programs.
Comprehensive Data Resource (CDR)

• CDR is being adopted for other NCI programs including the CPTAC program (Clinical Proteomic Tumor Analysis).

• The CDR code was posted last year.

• Goal of the NCI Collaborative Announcement:
  – See if the community would find this useful
  – Identify one or more collaborative partners who wish to adopt CDR at their institution.

• This is not a funding opportunity.
Program Needs Driving CDR Development

• Biospecimen science project: BPV
  – Cancer patients (primary)
  – Surgical specimens
  – Predefined preanalytical conditions

• NIH Common Fund Project: GTEx
  – Normal/Non-diseaseased
  – Postmortem
  – 30+ tissue types per donor
  – 900 donors
What Types of Information Do We Want to Capture?

Specimen is viable and biologically reactive

Molecular composition subject to further alteration/degradation

Pre-acquisition

Post-acquisition

Patient
Medical/Surgical Procedures
Acquisition
Handling/Processing
Storage
Distribution
Scientific Analysis
Restocking Unused Sample

Time 0

Patient

Scientific Analysis

National Cancer Institute
Examples of Information to Capture

- Consent
- Enrollment
- Blood and tissue collection and processing data
  - Blood tube type, time stamps for processing
  - For resected tissues: surgical clamp times, time placed in fixative or frozen, time placed in tissue processor, etc.
  - Storage conditions
- Pathology QC
- Pathology reports
- Clinical data about the donor
## Functional Requirements

### Required functional areas

<table>
<thead>
<tr>
<th>Patient</th>
<th>Medical/Surgical Procedures</th>
<th>Acquisition</th>
<th>Handling/Processing</th>
<th>Storage</th>
<th>Distribution</th>
<th>Scientific Analysis</th>
<th>Restocking Unused Sample</th>
</tr>
</thead>
</table>

### Not required functional areas
CDR Built After Trying Other Tools

• CaTissue for data collection at BSS
  – Not suitable for real-time data entry
  – Required data fields overwhelmed the system

• OpenClinica for sample annotation
  – Difficult for data integration and management

• CDR: custom-built to meet the challenging needs of GTEx and BPV collection efforts
Program requirements for CDR Functions

- Development of CDEs to thoroughly annotate the biospecimen life cycle to support the goals of the project
- Development of workflow-based annotation with live data entry at BSS when possible
- Record data at the BSS and transmit to project homepage (annotation, gross pathology images)
- Monitor shipping between different program sites
CDR – Can This be Useful to the Biobanking Community?

- Unmet needs for management software in biobanking community
- Facilitate Best Practices and annotation of biospecimen collection and processing steps
- Collaborative Announcement:
CDR – Collaborative Proposal

• Voluntary collaboration:
  – No funding to individual collaborator(s)
  – Collaborator(s) must have their own IT capacity to customize the software for their needs

• Interested parties should provide:
  – Description of unmet biobanking need that CDR could meet
  – Intended area of research that CDR could facilitate
  – IT experience and expertise in the proposed adoption