Current Topics in Biobanking
Informatics
Standards, Security, and Other Discussions

ISBER Informatics Working Group
Agenda

• List of Trends in Biobanking Informatics
• Best practices 101 for Biobank Systems
• Standards and Interoperability
• Accreditation: an Informatics perspective
• Discussion
List of Trends in Biobanking Informatics

Kevin Meagher

Information Management Services
List of Trends in Biobanking Informatics

- Continuing Trends
  - Barcoding & IDs
  - Fewer custom written tracking applications
  - Implementation of COTS IT systems as collection grows
  - Focus on security and privacy
  - Satisfaction with IT not improving
List of Trends in Biobanking Informatics

• Emerging Trends
  – Collections of Collections
    • IT Sharing
    • Data Sharing
  – Data Integration (#1)
  – Data Standards
List of Trends in Biobanking Informatics

• Some Things Never Change
  – Want IT to be
    • Easier
    • Faster
    • Cheaper
  – Want IT to Anticipate the Future
  – Willing to Use a Standard…. Unless…
Continuing Issues

• Still hand entering data
• Forced to leave silo system, but central system not ready for prime time
• Needs of high level vs needs of the biobank/lab
  – Barrier is not IT, but reluctance to sharing
  – Not to share unless it is of value to them (politics)
  – Privacy conflicts with sharing
BEST PRACTICES 101 FOR BIOBANK SYSTEMS
Labelling and Global Unique Identifiers (GUIDs)

A Rose by any other Name...

Piper Mullins, MIS
Smithsonian Cryo-Initiative
Labels
A Rose by any other Name...

- Why
  - linkages

- What
  - ID #s

- How
  - print
Global Unique Identifiers
A Rose by any other Name.../ Unique Snowflakes

• Why - traceable
• What - #s
• How - system
Managing Change

Mark Cada

LABVANTAGE
Laboratory Knowledge. Delivered.
What is *Change Management*?
Change management is a structured approach to transitioning individuals, teams, and organizations from a current state to a desired future state. The current definition of Change Management includes both organizational change management processes and individual change management models, which together are used to manage the people side of change. – Wikipedia

Minimizing resistance to organizational change through involvement of key players and stakeholders. - BusinessDictionary.com

What are some important issues to address in *Change Management* as it Applies to Informatics of Biobanking?
Lorenzi and Riley define eight categories for potential pitfalls when managing change during the implementation of new software.

- Communication
- Culture
- Underestimation of complexity
- Scope creep
- Organizational
- Technology
- Training
- Leadership/Ownership

Data Migration

Ashokkumar A. Patel
Reasons for Data Migration

• Current system over extended and need additional functionality
• Enterprise wide biobank initiatives
• Existing biobanks forming new networks

• Good reasons:
  – Ability to have your biobank get an **accurate accounting** of your specimens/data
  – Ability to **SHARE** your data with **OTHERS**

• Not so good reasons:
  – “Better and Newer” technologies
Retrospective vs. Prospective Data Collection

- Legacy data problems
  - Similarities vs. Differences
  - use of non-standardized vocabularies
  - mismatch of IDs when merging multiple files
  - Multiple data formats [i.e. date and units (cm vs. mm)]
  - Poor or multiple data models implemented
    - Patient->Specimen->Sample
    - Specimen->Sample
    - Collections->Patient->Specimen->Sample
    - Collections->Specimen->Sample
    - Sample data entry (TMA vs. paraffin block)
  - Free text fields

- Prospective data collection issues:
  - Workflow
  - Data standardization (different textual terms/coding)
  - Change management/Human Factors
  - Expectations of End Users
  - Centralized vs. Distributed/Federated Models
Take Home Message

• Past experience has shown that data migrations benefit from several steps that were described:

  – Analysis of the Old Data
  – Review of the New Data Structures
  – Creation of a Conversion Plan
    • Conversion of Coded Values
    • Conversion of IDs
    • Linkage to the Original Data (if needed)
    • Cleaning of Patient Identifiable IDs (if needed)
  – A Post-Migration Analysis and Validation
  – Consider Human Factors with Change Management
  – Time and Cost of Migration can vary (i.e. size of bank, quality/integrity of existing data, personnel, and complexity of required modules)
BIOBANKING STANDARDS
Helena Ellis
Director, Duke Biobank

“Our standards are very high. We even have high double standards.”
Biobanking at Duke

• Nearly every one of over 2000 human subjects protocol has a biobanking component
• Biobanking activities are diverse and siloed
• Biobanking Informatics are also siloed
  – No interoperability (no connection and no common data dictionary)
  – Commercial biobanking system(s)
  – Home grown database(s) (Oracle, SQL Server)
  – Microsoft Products
  – Paper
  – Some combination of above
LIMS for Biospecimen Management

- Centrally managed - Office of Research Informatics
  - Hardware/servers, application support
  - Customer support and training
  - Hands on data cleaning and migration from legacy system as they are retired
- Policy that all users must adopt and use the standard terminology
- Users request new terms through Duke Biobank (cannot add their own terms or permissible values)
Five Definitions for “Sample”

- Single unit of biological material
- Individual who provided material
- To collect material from an individual
- A set of samples that reflect parent/child relationships
- Several units of bio material collected at same time
How many samples do you have in your bank?

Five Different Terms

Sample
Sample Family
Sample Set
To Collect
Participant
Solution - Establish and Enforce Biobanking Standard Terminology

- Scope: Lifecycle of the biospecimen (NCI)*

*From Office of Biorepositories and Biospecimen Research of the NCI
Purpose was

- Accurately track and manage samples within a bank and between banks for studies shared across more than one bank
- Central support team learns only one language
- Training is more straightforward, both for the trainers and the trainees
- Interoperability between other clinical systems, such as data warehouse, eIRB and Epic

Purpose was NOT

- To promote sharing of the samples, in fact that would have been a problem
- We have another system used for ‘advertising’ the existence of samples
  – The Index of Biospecimens (Based on NCI’s Common Biorepository Model and Specimen Resource Locator)
Organization & Leadership

Oversight Committee

Facilitators

Informaticists

Group 1
Collection & Storage

Group 2
Tracking & Non-Chemical Handling

Group 3
Chemical Handling & Derivatives

Group 4
Complex / Omics Data

Group 5
Clinical Data

Working Groups
Authoritative Sources


Authoritative Sources, cont.

• Laboratory Data Management System (LDMS) Glossary
• CAP Biorepository Checklist Terms
• ISBER Glossary
• NCI’s
  – Best Practices for Biospecimen Resources- Minimal Clinical Data Set
  – Data Elements from NCI’s Common Biorepository Model
  – CaDSR
  – Thesaurus
Duke’s Standard Terminology
Common Data Elements

• Nearly 500 CDEs-
• Term, Definition, Data Type, Permissible Values

http://coreresearch.duke.edu/bbt
How Standards Proliferate:
Standards are great, everyone should have their own!
Other Biobanking Standards

1. **SPREC** - short-hand code of 7 important pre-analytical variables that the sample has experienced

2. **BRISQ** data elements regarding participants, samples, processing and storage recommended for reporting (publication or for regulatory purposes)

3. **CBM** - Meta Data about participants and sample collections for the purposes of visibility, collaboration and sharing

4. **MIABIS/ OMIABIS** - minimum data elements and their relationship about types of biobank for sharing

5. **CCB Standard Terminology** - expands on MIABIS, to enable searching across different biobanks/systems for sharing

6. **CAP Pre-Analytical Variables** - identifies key pre-analytical variables for tracking, and scientific impact if not tracked
Other Standards - SPREC

• SPREC provides a short-hand code of 7 important pre-analytical variables that the sample has experience

• Facilitates inter-laboratory and intra-laboratory specimen use by collaborating researchers

• SPREC is stored in sample tracking system and each sample has a SPREC

• Two samples of same type, that have had identical collection, handling and processing and storage procedures would have identical SPRECs

SPREC Fluid Specimens


1st: Type of sample  1. SER - Serum specimen

2nd: Type of primary container  2. SST - Collected with a serum collection tube

3rd: Pre-centrifugation  3. A – Pre-centrifugation delay is <2 hrs at room temp

4th: Centrifugation  4. E – Centrifugation at ambient temp at 3K-6K g w/braking

5th: Second centrifugation  5. N – One centrifugation step

6th: Post-centrifugation  6. A – Delay between centrifugation and freezing was <1 hr at 3°C to 7°C

7th: Storage condition  7. G - Serum was stored in straws at a temperature between −85°C and −60°C.
SPREC- Solid Tissue

TIS-BPS-N-B-RNL-A-A

1st: Type of sample 1. TIS - Tissue specimen
2nd: Type of collection 2. BPS - That was collected as a biopsy
3rd: Warm ischemia time 3. N - No warm ischemia time
4th: Cold ischemia time 4. B - Cold ischemia of <10 minutes
5th: Fixation type 5. RNL - Fixed in RNA Later
6th: Fixation time 6. A - Fixative time <15 minutes
7th: Storage condition 7. A - Stored in a 0.5- 2-mL polypropylene tube at a temp between −85°C and −60°C
Other Standards – BRISQ

“When performance is measured, performance improves. When performance is measured and reported back, the rate of improvement accelerates.”
Thomas S. Monson

Biospecimen Reporting for Improved Study Quality (BRISQ)

• Not a code

• Provides information to researchers and regulatory agencies consistent and standardized information to better evaluate, interpret, compare, and reproduce the experimental results.

• Recommended data elements to report when publishing or providing data to regulatory agencies (FDA)
  – Patient characteristics that might influence the biospecimens, such as vitals and disease states
  – Tissue type and the pathology of the sample
  – Collection and handling of the biospecimens, stabilization
  – Shipping, and storage conditions
BRISQ - Three Tiers

**Must have:** First tier, items crucial to report,
- anatomical site from which the biospecimens were derived
- manner in which the biospecimens were collected, stabilized, and preserved

**Should have:** Second tier, Items helpful to know, but less crucial to the science
- the time from biospecimen acquisition to stabilization.

**Nice to have:** Third tier, might be useful to know concerning the biospecimens but are not known to be as likely to influence research results or are unlikely to be available to researchers,
- environmental factors to which patients were expose
- type of storage container in which the biospecimens were kept.
<table>
<thead>
<tr>
<th>Data Elements</th>
<th>Tier 1 Should Have</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease Status of Patients</td>
<td>Controls or individuals with the disease of interest</td>
<td>Diabetic, Healthy Control</td>
</tr>
<tr>
<td>Clinical Characteristic of Patients</td>
<td>Available medical information known or believed to be pertinent to the condition of the biospecimens</td>
<td>α-Menopausal Breast Cancer Patients</td>
</tr>
<tr>
<td>Vital State of Patients</td>
<td>Alive or deceased patient when biospecimens were obtained</td>
<td>Postmortem</td>
</tr>
<tr>
<td>Clinical Diagnosis of Patients</td>
<td>Patient clinical diagnoses (determined by medical history, physical examination, and analyses of the biospecimen) pertinent to the study</td>
<td>Breast Cancer</td>
</tr>
<tr>
<td>Biospecimen Type</td>
<td>Solid tissue, whole blood, or another product derived from a human being</td>
<td>Serum, Urine</td>
</tr>
<tr>
<td>Anatomical Site</td>
<td>Organ of origin or site of blood draw</td>
<td>Liver, Antecubital Area of the Arm</td>
</tr>
<tr>
<td>Collection Mechanism</td>
<td>How the biospecimens were obtained</td>
<td>Fine Needle Aspiration, Pre-Operative Blood Draw</td>
</tr>
<tr>
<td>Type of Stabilization</td>
<td>The initial process by which biospecimens were stabilized during collection</td>
<td>Heparin, On Ice</td>
</tr>
<tr>
<td>Constitution of Preservative</td>
<td>The make-up of any formulation used to maintain the biospecimens in a non-reactive state</td>
<td>10% Neutral-Buffered Formalin, 10 USP Heparin Units/mL</td>
</tr>
<tr>
<td>Pathology Diagnosis</td>
<td>Patient pathology diagnoses (determined by macro and/or microscopic evaluation of the biospecimen at the time of diagnosis and/or prior to research use) pertinent to the study</td>
<td>Her2-Negative Intraductal Carcinoma</td>
</tr>
<tr>
<td>Type of Long-Term Preservation</td>
<td>The process by which the biospecimens were sustained after collection</td>
<td>Formalin Fixation, Freezing</td>
</tr>
<tr>
<td>Storage Temperature</td>
<td>The temperature or range thereof at which the biospecimens were kept until distribution/analysis</td>
<td>-80 °C, 20 to 25 °C</td>
</tr>
<tr>
<td>Storage Duration</td>
<td>The time or range thereof between biospecimen acquisition and distribution or analysis</td>
<td>8 Days, 5 to 7 Years</td>
</tr>
<tr>
<td>Shipping Temperature</td>
<td>The temperature or range thereof at which biospecimens were kept during shipment or relocation</td>
<td>-170 °C to -190 °C</td>
</tr>
<tr>
<td>Composition Assessment and Selection</td>
<td>Parameters used to choose biospecimens for the study</td>
<td>Minimum 80% Tumor Nuclei &amp; Maximum 50% Necrosis</td>
</tr>
</tbody>
</table>
### Other Standards - Common Biorepository Model

#### Specimen Availability
- Available to External Organizations
- Available to Commercial Organizations
- Available to non-US Organizations
- Collaboration Required

#### Annotation Availability
- Patient Demographics
- Exposure History
- Histopathologic Data
- Non-Omics Lab Data
- Longitudinal Specimens
- Matched Specimens
- Family History
- Study Treatment Data
- Follow-up Permissible
- Outcome Information

#### Collection Protocol
- Name
- Start Date
- End Date

#### Participant Collection
- Gender
- Ethnicity
- Participant Count

#### Specimen Collection
- Anatomical Source
- Specimen Type
- Participants Count
- Preservation

#### Institution
- Homepage URL
- Organization

#### Contact Person
- Name
- Email, Phone
- Address

#### Race
- Ethnicity

#### Patient Age Group

#### Diagnosis
Other Standards

**Minimum Information About Biobank data Sharing - MIABIS**

- Built on work done by BBMRI*, 52 data attributes, describing a biobank’s content, minimum data set for biobanks and studies using human biospecimens.
- Standardized data elements describing a biobank at the aggregate level, to facilitate data discovery
- Not intended to standardize data on an individual sample or participant level
- Minimum biobank and sample collection attributes that will help researchers initiate collaborations on biospecimen resource

*BBMRI- Biobanking and Biomolecular Resources Research Infrastructure*
**Other Standards**

**O-MIABIS**

**Ontology** is a formal naming and definition of the types, properties, and interrelationships of entities or concepts in a particular domain.

- Search for tissue samples from donors diagnosed with nemaline myopathy to determine:
  - Age group
  - Sample storage conditions
  - Possible detailed information about the biopsies
  - Whether myoblast cell cultures have been grown from these samples

Other Standards - Confederation of Cancer Biobanks (CCB) in UK

• Goal: to help create a national catalogue (federated) of samples to help researchers find samples based on the *individual* characteristics of patients.

• Key focus
  – seek agreement on the data that should be collected and how this should be structured
  – the structure and format of the data terms within the standard rather than defining the specific terms
  – make data model extendable for specific diseases

• Biobanks can continue to use their existing databases and ontologies; standard would not dictate the data terms to be used; it would ensure that the biobank supplies the current data definitions in use.

A Data Standard for Sourcing Fit-for-Purpose Biological Samples in an Integrated Virtual Network of Biobanks. P. Quinlan et al, Biopreservation and Biobanking. Volume 12, Number 3, 2014
Other Standards- CCB Standards

• Built on the MIABIS standard
  – Same: describes the biobank and a collection of samples
  – Different: allows the description of individual participant, and individual sample aliquot

• Allows two different, independent databases (with different terminology) to communicate about the samples available at each site
Standards Development Timeline

- **SPREC**: March 2010
- **BRISQ**: May 2011
- **MIABIS**: Nov 2012
- **O-MIABIS**: April 2013
- **Duke**: Dec 2013
- **CAP**: April 2014
- **CCB**: Nov 2014
<table>
<thead>
<tr>
<th></th>
<th>Duke’s CDEs</th>
<th>SPREC</th>
<th>BRISQ</th>
<th>CBM</th>
<th>O-MIABIS</th>
<th>CCB-STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Details</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Sample Meta Data</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Participant Details</td>
<td>YES</td>
<td>NO</td>
<td>YES (limited)</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Participant Meta Data</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Biobank Meta Data</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Created for sharing</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Considers quality</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Coded</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Created for tracking</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Created for reporting</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>
Acknowledgments

Oversight Committee
• Helena Ellis
• Mary-Beth Joshi, MPH
• Aenoch Lynn, MS
• Anita Walden, MS

Working Group Members
• Diane L. Satterfield
• David Layfield
• Lawrence Whitley
• Mike Leonard
• Tom Burke, PhD
• Dawn E. Bowles, PhD
• Eric Lipp, MD
• Lori Hudson, PhD
• Michelle Smerek
• Gary Archer, PhD
• Seth Fehrs
• Jessie Tenenbaum, PhD
• Carol Hill, PhD
• Pankaj Agarwal, PhD
• Paul Debien

“You can't always get what you want, but if you try sometimes, you just might find, you get what you need.”
-Mick Jagger
Accreditation for Biorepositories

Cheryl Michels
President, Freezerworks
ISBER Member since 2000
Founding co-Chair Informatics Working Group
Former ISBER Secretary-Treasurer
Available Accreditation

• ISO15189/ISO17025
  — Quality and competence for medical laboratories and testing laboratories

• CLIA (Clinical Laboratory Improvements Amendments)
  — Laboratory testing on humans, excludes research

• CAP (College of American Pathologists)
  — First to specifically accredit biorepositories
CAP Accreditation for Biorepositories

• What is it?
• Why is it important?
• What does it have to do with Informatics?
Workflow Management

- Sample Collection, Accession, and Processing
- Sample Storage and Retrieval
- Temperature Monitoring
- Shipping
Example from CAP Checklist
BAP.XXXX Specimen Identification Phase II

• There is a system to positively identify all participant specimens, specimen types, and aliquots through all phases of the analysis, including specimen receipt, nucleic acid extraction, nucleic acid quantification, hybridization, detection, documentation, and storage.
Security, Privacy, Regulatory Issues

• System Security – hardware, software, Presidential Directive on access
• Audit Trail – what, who, when, why
• Protection of PHI/PII – who can see it? logging that they saw it;
System Validation

• How often?
• How deep?
• If it’s not written down, you didn’t do it.
February 2014 ISBER Newsletter Article
Johns Hopkins University

• Specimen Security
• Data Security
• 30 Spot Checks
More Information?

• Contact me via ISBER 2015 meeting app
• Stop by Freezerworks booth #215
Get Involved

People like you are the backbone of our work!

ISBER is looking for volunteers with strong organizational and leadership skills to get involved.

Learn more…
Go to http://isber.org/getinvolved