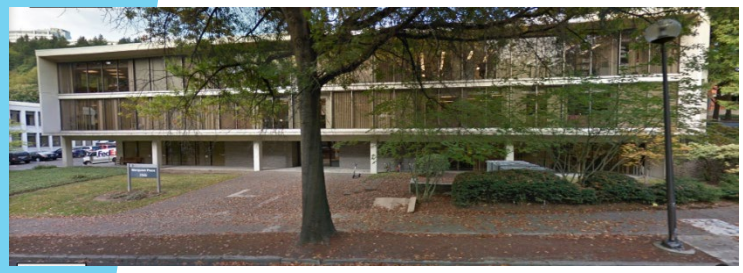
OCTRI Research Forum

Research Repositories:
Sample and Data Management Best Practices

DATE: May 20, 2021 PRESENTED BY: Danielle Galipeau

Research Repositories Sample and Data Management Best Practices

- Introduction to the BioLibrary
- Where to start
- Initiation of a repository
  - Study
  - Samples
  - Data
- Coded vs. De-Identified
- Questions



BioLibrary

The BioLibrary is a IRB approved, CAP accredited biorepository. We provide tissue, sample management, pathology services and scanning services to researchers.

Chris Corless M.D/Ph.D- Director
Danielle Galipeau – Operations Supervisor
Rosemary Makar M.D./M.B.A - Pathologist
Aletha Lesch – Program Technician
Leah Brower – Program Technician
Milly Seeley – Study Coordinator
Mary Katherine Henry – Study Coordinator
Christian Denagnon – Research Assistant II
Isaac Ross – Study Coordinator (Pathology)
Nick Nepochatov – Bioinformatics Developer
Grant Roesler – System Analyst

Knight BioLibrary

- Study Coordinator
  - Identify qualifying participants
  - Consent
  - Prospective collections
  - Delivery
- Study Coordinator in Pathology
  - Track cases and assist in collecting tissue
  - Triage collections in Pathology
- Specimen Management
  - Facilitate request from retrospective tissue
  - Management samples for 17 different studies including SMMART
  - Slide scanning
- Pathology Services

Plan, Plan, Plan!!!!

- What is driving the need for a repository
  - Collection for a specific research project
  - Disease specific
- What regulatory or oversight committees will be involved
  - IRB
  - Clinical Research partners
- Samples and Data
  - Repository Steward: A **guardian** is a person who has primary control of data and human biospecimens and maintenance of the repository.
  - Data Steward: A data steward is an oversight or data governance role within an organization, and is responsible for ensuring the quality and fitness for purpose of the organization's data assets, including the metadata for those data assets
- Who are the stakeholders and other departments you need support from
  - Clinical Research partners
  - Clinics, providers, pathology, OCTRI
- Infrastructure
  - What do you need
  - Where will it go
- Sharing

IRB Study

- Repository Study
  - What is the purpose of the biorepository
  - What types of samples and Data will be collected and what type or research could happen with the samples
  - Sharing of samples and data
  - Consent form and/or WOA
- Clinical Research
  - Transferring of samples
  - There can be specific restrictions that a sponsor can place on use of both samples and data.
    - Protocol
    - Consent
    - Assays

Standard Operating Procedures

- Equipment
  - Validation, monitoring, maintenance
- Collections
  - Participant identification to consenting
  - Collection to distribution
- Staff
  - Training and Competency programs
- Auditing
  - Internal and External
- Quality Management Program
  - Formalized QA/QC policies that minimize circumstances that could adversely affect scientific results

Samples

Sample: A specimen that has been collected
Sample Type: The preservation method of the sample

- Samples
  - Collections
    - Types of sample
    - Sample types
  - Processing
    - Special equipment
  - Metadata
    - Minimum data elements required
  - Labeling of samples
- Storage
  - Long term and short term
  - Best Practices
  - Monitoring
  - Fees
- Transportation
  - Hill, CHH, RLSB and KCRB

Data

- Data
  - Data Plan
    - What is needed now and in the future
    - Data Entry
  - Origin of Data
  - Storage
    - REDCap, Advarra, Eterprise Imaging, X-drive, Box, Research data Storage, encrypted personal computer, computers attached to lab equipment, OMERO, BEMS, LabKey
  - Cost of data storage

CODED VS. DE-IDENDTIFIED

Coded

- Data are separated from personal identifiers through use of a code. As long as a link exists, data are considered indirectly identifiable and not anonymous, anonymized or de-identified.
  - Indirect Identifiable: Data that do not include personal identifiers, but link the identifying information to the data through use of a code. These data are still considered identifiable by the Common Rule.
  - Anonymized: Previously identifiable data (indirectly or individually identifiable) that have been de-identified and for which a code or other link no longer exists. An investigator has NO means for linking anonymized data back to a specific subjects.
  - Anonymous: Data that was collected without identifiers and that were never linked to an individual. Coded data are not anonymous. See also de-identified data.

De-Identified

- A record in which identifying information is removed.
- Under the HIPAA Privacy Rule, data are de-identified if either:
- an experienced expert determines that the risk that certain information could be used to identify an individual is "very small" and documents and justifies the determination, or
- the data do not include any of the 18 identifiers (of the individual or his/her relatives, household members, or employers) which could be used alone or in combination with other information to identify the subject. Note that even if these identifiers are removed, the Privacy Rule states that information will be considered identifiable if the covered entity knows that the identity of the person may still be determined.

Definitions retrieved from https://irb.ucsf.edu/definitions

HIPPA Identifiers

You can get the full list on ITGs HIPAA and Research page

- Any other unique identifying number, characteristic or code, including any code that includes or is derived from any of the identifiers on this list.
  - Identifying Number
    - Study participant ID
  - Identifying Code: A code corresponds to a value that is derived from a non-secure encoding mechanism.
    - Secure hash function without a secret key
    - Hospital assigned barcodes
  - Identifying Characteristics
    - A *characteristic* may be anything that distinguishes an individual and allows for identification.

Definitions retrieved from https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Methods for De-identification

Expert Determination

- Apply statistical or scientific principles
- Very small risk that anticipated recipient could identify individual

Safe Harbor

- Removal of 18 types of identifiers
- A covered entity may assign a code or other means of record identification to allow information de-identified under this section to be re-identified by the covered entity, provided that:
(1) *Derivation*. The code or other means of record identification is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual; and
(2) *Security*. The covered entity does not use or disclose the code or other means of record identification for any other purpose, and does not disclose the mechanism for re-identification.

Definitions retrieved from https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Repository Resources

- IRB Policies and Forms Webpage
  - Repositories (Help Sheet)
  - Repository Protocol Checklist
  - Sharing, submittal agreements, and tracking forms
  - Consent and Authorization – Repository Only
- Knight Cancer
  - Knight Repositories
  - Data Governance
  - Tech Transfer Office
- NIH
  - Final NIH Policy for Data Management and Sharing (effective 1/25/2023)
  - Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research
  - NCI Best Practices for Biospecimens
- OHRP
  - Issues to Consider in the Research User of Stored Data or Tissues
- Sample Acquisition Network (SAN)
  - Resource for guidance

Next OCTRI Research Forum:
Sex and Gender Considerations in Clinical Research
June 22, 2021 12:00-1:00
Registration in Compass

Thank You