

# Potential Applications for the Biospecimen Research Database and Proposed Future Development

Andrew W. Breychak B.A.<sup>1</sup>, Kelly B. Engel Ph.D.<sup>2</sup>, Ian Fore D.Phil.<sup>3</sup>, Helen M. Moore Ph.D.<sup>4</sup>

<sup>1</sup>Sapient Government Services, <sup>2</sup>Preferred Staffing Group, <sup>3</sup>National Cancer Institute, Center for Bioinformatics and Information Technology,

<sup>4</sup>National Cancer Institute, Office of Biorepositories and Biospecimen Research

## Informatics

### BRD 2.0

- **Support enhances search and curation features**
- **Standard Operating Procedures**
- **Experimental Data**
  - BRN data
  - Curated supplemental information where available
- **Community collaboration**
- **caBIG compatibility**
  - Syntactic (systems for grid query) and semantic (vocabulary) integration
- **Digital Object Identifiers**
- **Linking to other systems**
  - caTissue (and other systems) subscribing to SOPs, PubMed Linkout

<http://biospecimens.cancer.gov/brd>

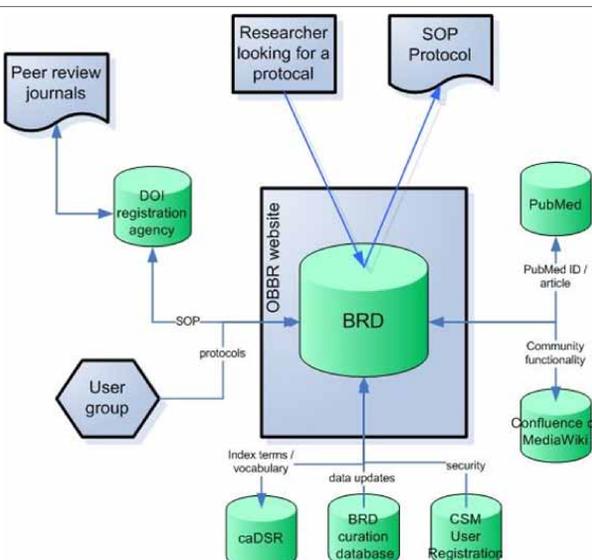
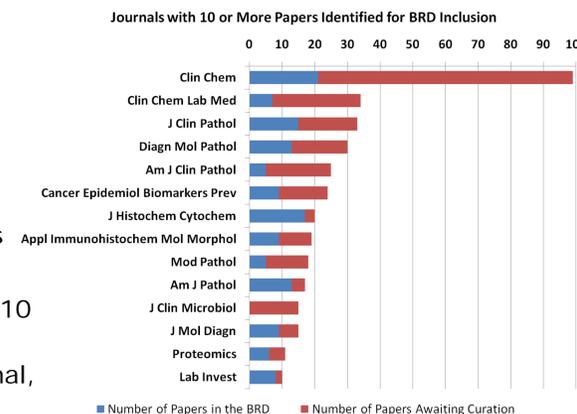
[biospecimens@mail.nih.gov](mailto:biospecimens@mail.nih.gov)

## Curation



### Papers identified for future inclusion...

- Number 707 research and review articles
- Represent 315 journals
- Span the publication dates of 1925 to 2010
- Are a product of literature searches, monthly literature screens, and institutional, organizational, and individual referrals



**Add Standard Operating Procedures (SOP's) as PDF files into the BRD...increase the number of existing papers & studies.**

- Enable SOPs to be:**
- Exported & Imported
  - Structured
  - Unique
  - Versioned
  - Actionable (searchable)

- Starting a new development cycle...**
- Support Standard Operating Procedures (and Protocols)
    - Exchange structured protocols with biorepositories systems
  - Integration with caBIG™ tools (caTissue, caDSR, CSM)
  - Linkage to data from Biospecimen Research Network (BRN) studies

- Supporting BRD...**
- Continue work to support data curation
  - Realigning BRD scope with OBBR vision and linking to other initiatives
  - Engaging the community at the BRN Symposium

## Future

Standard Operating Procedures

Guidelines

>1,000 Papers

Meta-analysis

500-1,000 Papers

Biospecimen Research Database

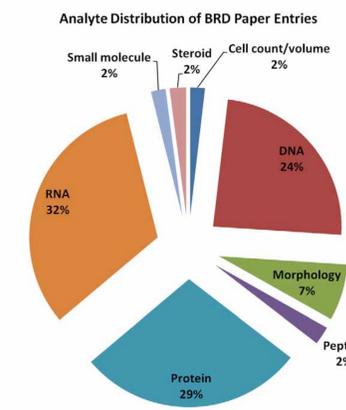
303 Papers & 539 Studies

## Present

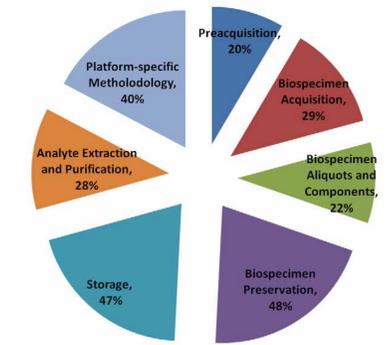
Add infrastructure to support Standard Operating Procedures...

Expert workshops

Data mining

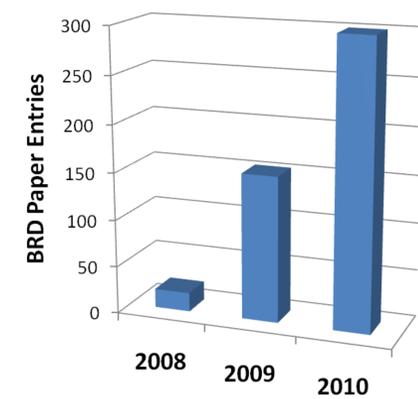


BRD Paper Distribution based on the Experimental Factor Investigated



### Papers contained within the BRD...

- Have published findings specific to human biospecimen science in peer-reviewed journals
- Have been categorized and annotated by a PhD level scientist to facilitate searchability and highlight key findings
- Represent 14 analyte classifications
- Represent 95 experimental factors encountered within the lifecycle of a biospecimen
- Represent 119 journals, with 62% of journals having a single BRD paper entry
- Are continually added



Since its inception by the National Cancer Institute's Office of Biorepositories and Biospecimen Research (OBBR) in 2008, the Biospecimen Research Database (BRD) (<http://biospecimens.cancer.gov/brd/>) has been in a state of continual growth and development. Developed by OBBR as an annotated reservoir of peer-reviewed and published research and review articles of specific relevance to human biospecimen science, the BRD has increased in both content and design, with additional curation tools and experimental factor search options. In an effort to match the diversity observed among biospecimen science articles, the number of journals represented by BRD curated publications has grown to 115 in the last year, as has the diversity of encompassed analytes and experimental variables. The future of the BRD is robust, with 700 additional publications identified for curation (150 of which are a product of organizational referral), targeted implementation of user accounts, evaluation as a vehicle for meta-analysis, and platform expansion to accommodate electronic Standard Operating Procedures (SOPs). The BRD will become a definitive source of structured SOPs that can be subscribed to by caBIG® and non-caBIG applications. SOPs will exist as structured data that can be queried in caGrid-enabled data services. The BRD is intended to be used by researchers, tissue expeditors, biorepository managers, and pathologists who wish to search, subscribe, and comment on SOPs and the supporting annotated published literature. As a central goal of the OBBR is for the BRD to serve the needs of the biospecimen science community, feedback and paper referrals are welcome and can be submitted at [biospecimens@mail.nih.gov](mailto:biospecimens@mail.nih.gov).