

Consensus-Driven Development of a Terminology for Biobanking, the Duke Experience

Helena Ellis,¹ Mary-Beth Joshi,² Aenoch J. Lynn,³ and Anita Walden⁴

Biobanking at Duke University has existed for decades and has grown over time in silos and based on specialized needs, as is true with most biomedical research centers. These silos developed informatics systems to support their own individual requirements, with no regard for semantic or syntactic interoperability. Duke undertook an initiative to implement an enterprise-wide biobanking information system to serve its many diverse biobanking entities. A significant part of this initiative was the development of a common terminology for use in the commercial software platform. Common terminology provides the foundation for interoperability across biobanks for data and information sharing. We engaged experts in research, informatics, and biobanking through a consensus-driven process to agree on 361 terms and their definitions that encompass the lifecycle of a biospecimen. Existing standards, common terms, and data elements from published articles provided a foundation on which to build the biobanking terminology; a broader set of stakeholders then provided additional input and feedback in a secondary vetting process. The resulting standardized biobanking terminology is now available for sharing with the biobanking community to serve as a foundation for other institutions who are considering a similar initiative.

Keywords: biobank, biorepository, terminology, data elements, interoperability, standards

Introduction

CLINICAL BIOBANKS PROVIDE both the basic and translational research communities with human biological samples for basic research, biomarker discovery, and validation. Although institutions have collected and managed biological specimens for decades, the term “biobank” was not coined until 1996.¹ In the early 2000s, biobanking became recognized as an emerging field with the establishment of the International Society of Biological and Environmental Repositories (ISBER), a forum of biospecimen researchers.² Since then, the discipline has continued to grow and mature, and as a result, other international forums and conferences have emerged.^{3–6}

In 1991, RAND Corporation⁷ estimated there were 146 large tissue banks, repositories, core facilities, and longitudinal biobanks comprising over 120 million biological samples in the United States. In a 2013 survey of 456 U.S. biobanks, 78% ($n=356$) were affiliated with an academic institution.⁸ However, since biobanks are inconsistently defined in the research community, with variations in size, mission, or extent of associated data,^{1,9,10} one could argue that any collection of bi-

ological samples for research is a biobank, and therefore a large academic organization could house hundreds to thousands of biobanks. At the lead author’s institution, 96% of the roughly 2500 active human subjects’ research protocols included a biobanking component at the time of this analysis. If we estimate that the other 356 academic institutions surveyed by Henderson have only half the active human subjects research protocols (1250), we conservatively speculate that there could be at least 427,200 academic biobanks in the United States ($356 \times 96\% \times 1250$). This proliferation of biobanks has led to a call for standardization as researchers have recognized that biobanking often represents an intersection between clinical care and research and the tracking chain of custody and preanalytical variables in a standardized searchable manner is essential for quality biobanking, and hence, good science. As previously asserted by Carolyn Compton of the National Cancer Institute (NCI), biospecimens are the “center of the personalized medicine universe,” and poor and/or unreproducible results can lead to the wrong answers with unprecedented speed.¹¹ Without quality specimens and well-documented annotation, one can neither differentiate between

¹Biobanking Without Borders, LLC, Durham, North Carolina.

²Department of Surgery, Duke University, Durham, North Carolina.

³Buck Institute for Research on Aging, Novato, California.

⁴University of Arkansas for Medical Sciences, Little Rock, Arkansas.

a null result and noise nor be confident of the veracity of test results,¹² wasting both the precious donations made by participants as well as financial resources.

In the last decade, interest in these issues has intensified and publications emphasizing the need to standardize biobanking practices and procedures have proliferated.^{13,14} ISBER and the NCI developed and published best practice documents that included guidelines regarding sample collection, processing and storage, quality, as well as legal and ethical issues related to consent and governance.^{15,16} In 2012, the College of American Pathologists (CAP) began the United States' first Biorepository Accreditation Program, aiming to "improve the quality and consistency of facilities that collect, process, store, and distribute biospecimens for research."¹⁷ However, the biobanking terminology (i.e., data elements and their definitions) is not specifically prescribed by these U.S. resources. In Europe, the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI¹⁸) developed a comprehensive source of information about existing biological sample resources and developed the informatics infrastructure to link existing biobanks. They defined 52 attributes to serve as the minimum data set for biobanks and studies using human biospecimen, called MIABIS¹⁹ (Minimum Information About Biobank Data Sharing).

In the last several years, numerous companies began offering software to track and manage biospecimens. Before these platforms were available, biobanks relied on homegrown databases, spreadsheets, and paper logbooks that were perceived as more affordable and accessible than relational databases and commercial products. However, as requirements for chain of custody, inventory management, and sample documentation have become more formally defined, more robust tools have become the norm. Commercial software products that allow users to customize data fields are attractive to users, but sometimes result in idiosyncratic definitions, leading to difficulty with queries and reports. Cancer centers and academic medical centers are beginning to invest in enterprise software solutions to consolidate biobank databases; nonstandard, inconsistent data elements and definitions can lead to misunderstandings that confound research results. Thus, rigorous data standards are becoming expected and are being encouraged, if not required by government sponsors in solicitations. Although standardization of disease-specific data elements has progressed,^{20,21} at the time this project was undertaken, very little work had been done to standardize data elements for biobanking and biospecimen science.

The Biobanking Data Element Standardization Project took place in the context of a larger centralized biobanking effort that began in 2012, in which our institution sought to strengthen and harmonize its many diverse biobanking entities. Institutional funding was secured to purchase, configure, and implement an enterprise-wide biobanking information management system (BIMS). At the outset, a policy was established that required all users to use a common set of data elements with standard definitions and prespecified valid values for discrete data elements, regardless of which BIMS was purchased. This was important for several reasons shown in Table 1. To enforce this policy, user roles and privileges in the BIMS were designed such that only central administration could add new terms that had been fully vetted and approved. In this study, we describe the consensus-driven process and how multidiscipline stakeholder engagement ensured that the resulting terms met users' needs.

TABLE 1. THREE IMPORTANT REASONS TO USE A COMMON TERMINOLOGY

-
1. Searching for appropriate samples across the legacy biobanks was difficult at best and impossible at worst, without the use of a common terminology.
 2. Reporting was identified as a critical requirement by principle investigators, biobank managers, and sponsors. The disparate and nonstandard terminology in the legacy systems had already been proven to be an impediment to querying and reporting across existing biospecimen databases.
 3. The system would be centrally supported by a team that provides training, data migration, and ongoing support; thus, the data captured by the legacy biobanks needed to be standardized.
-

Materials and Methods

The methodology established by a data standards project from the National Institutes of Health (NIH) Roadmap Program was leveraged for the Biobanking Data Element Standardization Project. The NIH projects described by Nahm et al.²² developed a methodology for identifying, defining, and standardizing therapeutic area data elements. The process involved creating an expert team to review clinical content and an informatics team to provide structure and develop data elements based on International Organization for Standardization (ISO) 11179 standards.²³ For this project, a data element is described as a standard term, its definition, and set of allowable values. These methods were adapted and a comprehensive project plan was designed over several months that defined the organizational and leadership structure, communication plan, data element approval process, and member responsibilities. The plan served as a reference document and helped team members maintain focus toward an achievable goal. In the first phase of the project, described in this article, the terms were identified and defined. The identification and defining of allowable values are to be addressed in a later phase as individual biobanks are implemented on the BIMS platform.

Organization and leadership

A terminology Oversight Committee led and managed the overall initiative (Fig. 1) and consisted of four individuals: the Director of the Duke Biobank served as the chair and the other three individuals were the BIMS product manager, a terminologist, and an experienced biobank manager. The Oversight Committee directed the vision and established and governed the Biobanking Data Element Standardization Project, was responsible for final decisions and deliverables, and championed the project with institutional leadership.

Five working groups (WGs) were established by the Oversight Committee by inviting qualified individuals to participate based on their biobanking, scientific, or informatics experience, planned future use of the BIMS, and ability to work as a member of a team. Each WG was appointed an informatics lead and a facilitator. Facilitators scheduled meetings, recorded meeting minutes, documented decisions, and handled communication with the other WGs and the Oversight Committee. The facilitators from each WG met regularly to review progress and resolve problems such as unclear or overlapping scope, duplicate terms, and conflicting

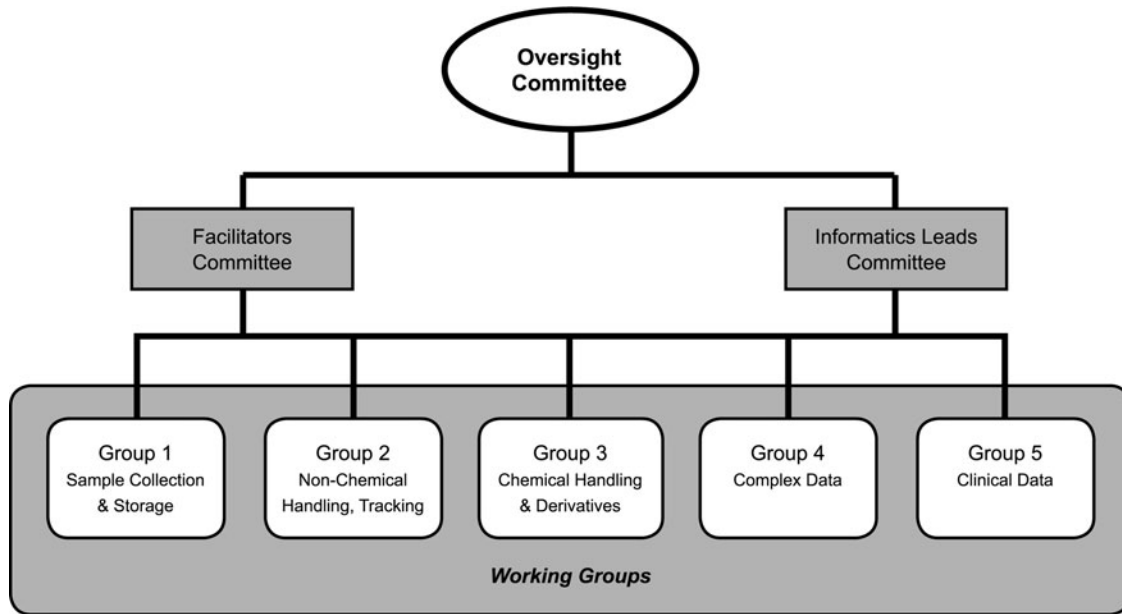


FIG. 1. Diagram of project organization and leadership.

definitions. They then reviewed all the terms approved by the WGs before sending them to the Oversight Committee. The informatics leads provided advice consistent with informatics conventions, such as proper data management techniques, data element structure, and existing authoritative sources.

Definition of scope

The Biospecimen Lifecycle (Fig. 2), as established by the Biospecimen Research Network of the NCI,²⁴ was used to define the general scope for the project. The Lifecycle is defined by all the activities and processes between consenting of a participant, through sample collection, processing, storage,

tracking, analysis, and restocking. For practical purposes, the scope for each WG needed to be defined in such a way as to allow the effort of each WG to leverage, but not duplicate, the work of the others. The cumulative effort of the five WGs encompassed the Biospecimen Lifecycle; however, the processes focused on by each WG differed slightly from those defined in the Lifecycle (Table 2).

Data element identification, process, and communication

The Oversight Committee provided the potential sources of data elements after consulting with in-house

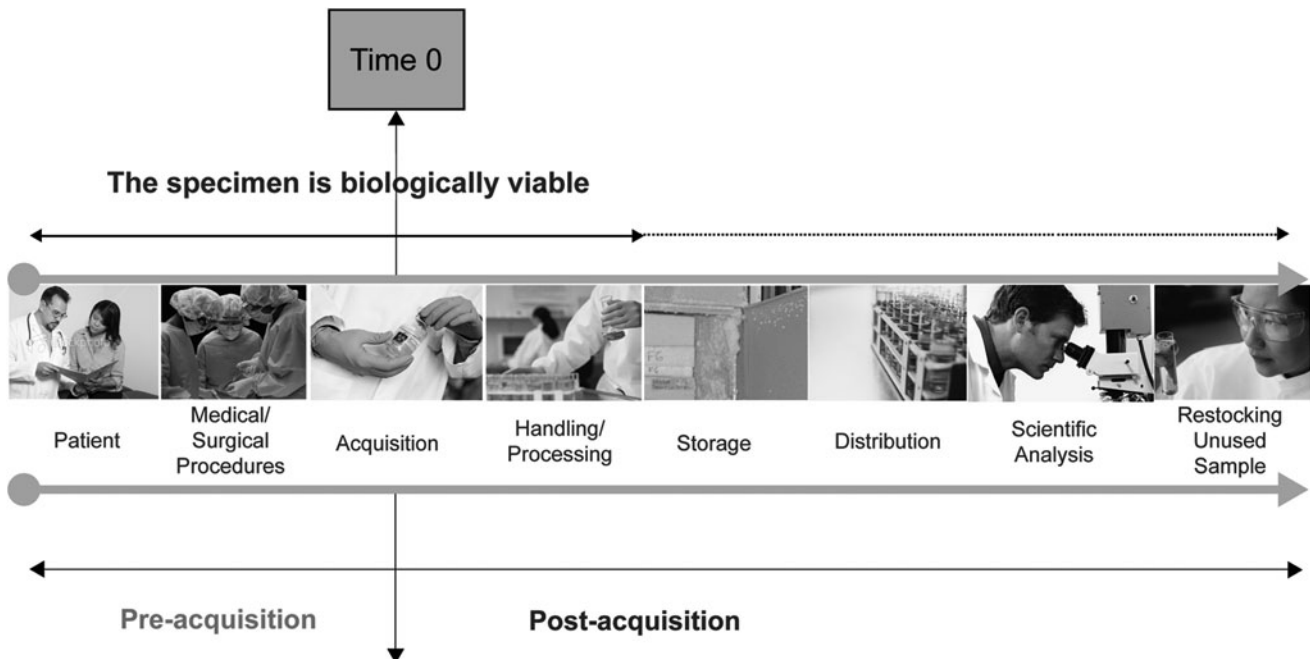


FIG. 2. Lifecycle of the biospecimen as defined by the biospecimen research network of the National Cancer Institute (reprinted with permission).

TABLE 2. SCOPE OF WORK FOR EACH OF THE FIVE WORKING GROUPS

| | |
|---|--|
| 1. Sample collection and storage: data elements related to the collection and storage of biological material | <ul style="list-style-type: none"> a. Collection event information (dates, times, temperatures, study site, physical position, etc.) b. Collected material information (collection procedure, sample type, body site, quantities, etc.) c. Material acquisition information (container, identifiers, participant demographics, shipping information, etc.) d. Accessioning information (biobank identification, sign in, storage units, temperatures, freezer locations, etc.) e. Material handling information (handling instructions, quantities, dates, temperatures, times, quantities, etc.) |
| 2. Tracking and nonchemical sample processing: data elements related to nonchemical handling of collected biological material (e.g., separation into smaller units), tracking material in and out of the biobank, and storage device monitoring | <ul style="list-style-type: none"> a. Biobank location, building, room, and personnel for the biobank b. Handling and nonchemical processing information (identifiers, aliases, barcodes, methods, and procedures, etc.) c. Storage unit information (temperature/time logs, make, model, repair history, monitoring, and asset number, etc.) d. Material storage information (storage unit type, storage unit position, and storage conditions and temperature, etc.) e. Study/protocol descriptors and information (IRB number and status, consent status, title, principal investigator, material use restrictions, data collection parameters, and data sharing restrictions, etc.) |
| 3. Chemical handling and derivatives: data elements associated with chemical handling, manipulation, and production of derivatives and products | <ul style="list-style-type: none"> a. Stabilization information b. Derivative types (RNA, DNA, protein, and IHC, etc.) c. Bench-top protocols utilized (methods for extraction and detection) d. Chemical handling information (kit types, lot numbers, and method names, etc.) e. Concentration and quality metrics (units, methods) |
| 4. Complex data: data elements associated with complex data, such as “omics” type analyses and resulting data | <ul style="list-style-type: none"> a. Data types available (SNP, gene array, sequencing, and raw data vs. normalized data, etc.) b. Methodology/platform information (chip type, etc.) c. Analyses information (type of analyses and dates performed, etc.) d. Location/link to file and size of file e. Data describing primary and secondary data f. Analysis techniques and processes and how results are stored |
| 5. Clinical data: data elements related to clinical outcomes and demographics | <ul style="list-style-type: none"> a. Standard of care data b. Clinical laboratory data c. Diagnoses d. Disease stage e. Clinical follow-up/survival information (date of death, last contact, or disease recurrence, etc.) f. Detailed demographics (smoking history and marital status, etc.) |

informaticists and other internal biobanking stakeholders, and after reviewing the literature and previous work done by standards organizations, including other authoritative sources such as the NCI’s Cancer Data Repository and Registry (caDSR).²⁵ The existing data elements used in each legacy system were also carefully considered to ensure that all terms would be represented in the data elements and associated terminology. With an eye to improving biospecimen annotation, data elements regarding key preanalytical variables related to biospecimen quality and “fit for purpose” were identified in publications to improve biobanking practices, even if they were not currently in use in any participating biobanks.^{26,27} All authoritative sources are described in Table 3.

The Oversight Committee provided the WGs with an initial list of data elements to help clarify scope and promote discussion, and the activities of each WG were integrated to facilitate progress. Each WG compiled a draft of their internally approved candidate data elements for review by the informatics leads. Based upon the informa-

tion provided by the WGs and the review by the informatics leads, the Oversight Committee harmonized terms, concepts, and definitions provided by the individual WGs. This was done to ensure consistency and resolve conflicts. Once the work of the WGs concluded, the terms were put into a domain model, which helped to identify gaps. All terms were then distributed to the institutional biobanking community and other affiliated researchers for a 30-day internal comment period. When the commercial BIMS was purchased, out-of-the-box terms in the commercial BIMS were identified, and engagement with each bank during detailed functional requirements specifications and data migration resulted in additional terms. The Oversight Committee managed the final terms and definitions and data types (Fig. 3).

Results

In total, 361 common data elements were established and approved by the Oversight Committee during the

TABLE 3. AUTHORITATIVE SOURCES USED FOR DATA ELEMENTS

| <i>Name</i> | <i>Primary function</i> | <i>Rationale for inclusion/exclusion</i> |
|--|---|--|
| The NCI Thesaurus ^{28,29} | A collection of curated terms, definitions, and synonyms of primarily cancer-related biomedical concepts that are used by NCI projects, researchers, and collaborators to promote semantic interoperability | While many terms and definitions were adopted from the NCI Thesaurus, it was deemed too cancer centric for wholesale adoption when considering the needs of Duke's noncancer researchers. |
| The NCIs caDSR ²⁵ | An ISO 11179 metadata repository for common data elements used in clinical research. Researchers can query the caDSR for common data elements to help build case report forms that would be consistent and comparable with previous research | Most of the terms and definitions overlapped with the NCI Thesaurus; so this resource was not used heavily. It also was cancer centric, but it was useful for defining permissible values for a limited number of terms. |
| The NCIs CBM ³⁰ | A data model to help facilitate sharing of biospecimen resources. The CBM focuses on metadata about biospecimen resources related to the samples and participants, and contains yes/no indicators about sample annotation and sample availability | This data model was already in use at Duke in a software tool designed to "advertise" biospecimen resources; hence, it was critical that these data elements and definitions were incorporated into the terminology. |
| Commercial BIMS Software | The out-of-the-box terms that came with the inherent functionality of the BIMS | The Biobanking Data Element Standardization Project was well under way when a commercial BIMS was identified and purchased, after which the product's out-of-the-box terms were incorporated. |
| Legacy inventory systems | Data elements in use in institutional legacy inventory systems. Legacy systems included in-house developed databases and other commercial inventory systems | Each biobank that planned to use the BIMS also participated in the terminology effort and provided a list of data elements from their existing systems. Definitions were established together since they were not necessarily readily available. |
| ISBERs Best Practices for Repositories ¹⁵ | A glossary provided ISBER related to their published biobanking best practices | Provided some basic terms and definitions related to the foundations of biobanking. |
| The NCI Best Practices for Biospecimen Resources ¹⁶ | A glossary provided by both the NCI and ISBER related to their published biobanking best practices | Provided some basic terms and definitions related to the foundations of biobanking. |
| IRB website | The IRB website serves as a resource for Duke researchers regarding policies and procedures | Provided terms and concepts specifically related to research approvals, policies, and informed consent requirements. |
| BRISQ ²⁷ | A list of data elements that represent factors believed to influence biospecimen quality and should be considered for reporting | Very relevant and specific data elements related to biobanking science and sample quality |
| Important preanalytical variables defined by CAP ²⁶ | Variables that may affect the quality and/or value of a biospecimen from the time of consenting until the biospecimen is used banked or used for testing | Very relevant and variable related to biobanking science and sample quality |
| MIABIS ^{19,31} | Set of 52 attributes defined as the minimum data set for biobanks and studies using human biospecimens that describe a biobank's content | Relevant data elements related to meta data and information needed for sharing samples |

BIMS, biobanking information management system; BRISQ, Biospecimen Reporting for Improved Study Quality; caDSR, Cancer Data Standards Registry and Repository; CAP, College of American Pathologists; CBM, Common Biorepository Model; IRB, Institutional Review Board; MIABIS, Minimum Information About Biobank data Sharing; NCI, National Cancer Institute.

Biobanking Data Element Standardization Project (Supplementary Appendix 1; Supplementary Data are available online at www.liebertpub.com/bio). Each data element was assigned to a general category, a definition, and data type. Allowable values are to be defined in a later

stage as the biobanks are implemented in the BIMS. The completed data elements are not represented in a data model, and therefore relationships are not specified between them. The data elements are grouped into seven general categories described in Table 4.

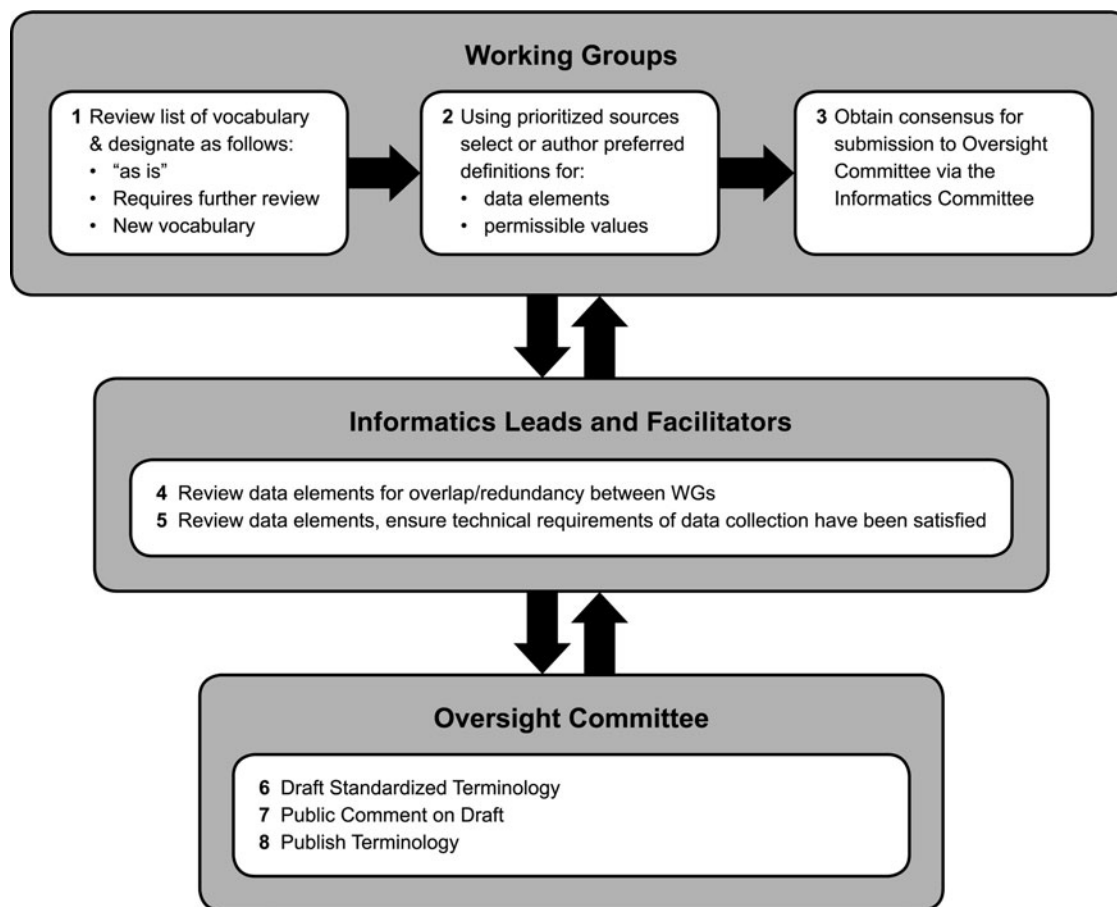


FIG. 3. Data element development process.

Discussion

Challenges

One goal of creating an enterprise-wide biobank is to leverage multiple biospecimen resources. In the context of many independent diverse biobanks, this necessitated the use of common data elements and terminology. The development of standard data elements and associated terminology demands training and commitment and is a meticulous and multistage process. Subject matter experts as well as informatics experts must be committed to ensuring their mutual understanding of each element so that an appropriate consensus can be reached. The scope can be difficult to define when topics overlap, impeding progress. These considerations, among others, presented both challenges and opportunities.

A common terminology reduces the time-consuming activity of mapping one term to another, which in turn reduces loss of information. Similar terms may not have the same meaning, leading to incorrect use of data and information. Indeed, conversations around the capacity and size of our institution’s many biobanks clearly illustrated the challenge of using a nonstandard language. The seemingly simple question of “how many samples are in your bank?” could not easily be answered without further description of what was meant by “samples.” We found that at our institution, we had no less than five definitions of the word “sample.” At the completion of the terminology effort, these five definitions became five distinct data elements (Table 5). In addition, basic terms such as “sample” needed to be defined by one WG before another WG could develop and define terms within their purview.

TABLE 4. CATEGORIES OF DATA ELEMENTS

| | <i>Category</i> | <i>Description</i> | <i>No. of terms</i> |
|----|----------------------|---|---------------------|
| 1. | Clinical annotation | Clinical data and information related to the participant that is important for selection of a sample for downstream use | 56 |
| 2. | Informed consent | Data elements related to the process of informed consent | 13 |
| 3. | Study administration | Data elements related to management of a biobanking study | 60 |
| 4. | Package | Data elements related to shipping and distribution of samples | 24 |
| 5. | Participant | Data elements related to a consented individual who is participating in a research study | 55 |
| 6. | Samples | Data elements related to biological material | 113 |
| 7. | Storage | Data elements related to storage of samples in a biobank | 40 |

TABLE 5. EXAMPLE OF MULTIPLE DEFINITIONS FOR SINGLE-TERM SAMPLE

| <i>Institutional legacy term</i> | <i>Definition</i> | <i>New term</i> |
|----------------------------------|--|-----------------|
| Sample | A single unit of biological material (noun) | Sample |
| Sample | Several units of biological material collected at the same time from one participant (noun) | Sample set |
| Sample | A set of different units of biological material that reflect parent/child relationships (noun) | Sample family |
| Sample | The participant from whom the biological material was collected (noun) | Participant |
| Sample | To collect biological material from a participant (verb) | Collect |

Agreement on critical aspects such as format and structure is time-consuming and should be considered when planning. In fact, the entire process took almost twice as long as the 9 months that were predicted. Change is difficult; therefore, it is essential to engage stakeholders and obtain their input, buy-in, and support. A delicate balance was required to ensure essential stakeholders were included, while also keeping the groups small enough and adequately engaged to complete activities. Groups that were comfortable with structured data elements with specified valid values will likely have an easier time adapting to a new data element set. On the other hand, those whose terminology evolved over time without any thought to consistency may welcome structure and standardization, having experienced frustrations with inability to query, use, or combine data or samples across studies.

The scope for WG 5 included clinical data elements that could potentially affect the suitability of a sample to be selected for downstream analysis. There were strongly voiced differing opinions as to the scope, and therefore the activity was postponed until after nearly all data elements from the other WGs were complete. The diverse research areas of the WG members (e.g., cancer, Alzheimer's disease, and population studies) was a major factor in this discussion; collecting all of the relevant clinical data elements for these different diseases would have drastically affected the WGs' scope. In addition, there was considerable disagreement regarding to what extent the BIMS should manage clinical data; some felt that all clinical data should be stored in the BIMS and others felt the exact opposite. Furthermore, at our institution, the existing Enterprise Data Warehouse^{32,33} holds the complete clinical data that are linkable to samples, and duplicating those data did not make sense or add value to the initiative. In the end, the 56 clinical data elements that were identified and defined will likely not be managed in the BIMS, but tracked in other database systems.

Successes

Having a collaboratively established standardized data element set and associated terminology helped tremendously during the data migration process of incorporating each biobank into the BIMS, since the preexisting data elements were considered during the standardization process and the biobank members were also members of each WG.

During the project, several articles were published about important biobanking preanalytical variables. The Oversight Committee took them into consideration to identify gaps and help prioritize decisions, as previously described. For example, Robb et al.²⁶ identified 170 biobanking preanalytical variables with a priority score and an indication as to the scientific impact if not recorded. The Oversight Committee reviewed the article and identified many terms that were already on our list, terms to include, and many that we were not

feasible to include due to logistics (e.g., third party clinical service providers) or the unavailability of information. Hence, the common data element initiative was informed by availability and accessibility of data in the real world.

Future work

Our resulting biobanking data elements are being integrated into the Ontology for Biobanking (OBIB³⁴) and the biobanking Informed Consent Ontology (ICO), which are both available on GitHub.^{35,36} Both OBIB and ICO are being developed by a cross-institutional, multidisciplinary collaboration. Ontologies contain formal naming and definition of the types, properties, and their interrelationships terms. Linking standardized data elements to rich ontological knowledge sources allows those querying or using the data to answer additional questions beyond those answerable from the data alone by reasoning over the relationships encoded in the ontology.

Conclusions

Biospecimens have become a highly valued resource and detailed annotation of these samples in a standardized manner is becoming increasingly important for academic organizations. It is clear that the development of the standardized biobanking data elements and associated terminology added value to the BIMS initiative at Duke. Development and use of common and standard data elements and associated terminology are also increasing elsewhere across healthcare and research domains as the need for quick and scalable information retrieval rises. Organizations must make the decision to adopt a common or standard terminology to facilitate the exchange of data or map to an ontology that is outside the information system. Both approaches have merit; similarly, they both take leadership, resources, significant time, and effort.

Acknowledgments

Special thanks to Blair Chesnut for creating the terminology management tool. Also, special thanks to the working group members: Diane L. Satterfield, David Layfield, Lawrence Whitley, Mike Leonard, Tom Burke, Dawn E. Bowles, Eric Lipp, Lori Hudson, Michelle Smerek, Gary Archer, Seth Fehrs, Jessie Tenenbaum, Carol Hill, Pankaj Agarwal, and Paul Debien. Research reported in this publication was supported by the National Center For Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR001117. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Disclosure Statement

No conflicting financial interests exist.

References

1. Hewitt R, Watson P. Defining biobank. *Biopreserv Biobank* 2013;11:309–315.
2. International Society for Biological and Environmental Repositories (ISBER). Connecting Repositories Globally through Best Practices. ISBER website. Available at: www.isber.org/?page=About Accessed September 14, 2015.
3. Leaders in Biobanking Congress 2016. Available at: www.healthtech.com/biobanking Accessed September 15, 2016.
4. Biorepositories and Biospecimen Research Branch. Biospecimen Resource Network. Available at: <http://biospecimens.cancer.gov/researchnetwork> Accessed September 15, 2016.
5. Biobanking and Biomolecular Resources Research Infrastructure website. Available at: www.bbmri.eu Accessed September 15, 2016.
6. International Society for Biological and Environmental Repositories (ISBER). ISBER Affiliate and Associate Partners. Available at: www.isber.org/?page=Affiliations Accessed September 15, 2016.
7. Eiseman E, Haga SB. *Handbook of Human Tissue Sources: A National Resource of Human Tissue Samples*. Santa Monica, CA: RAND Corporation; 2000.
8. Henderson GE, Cadigan RJ, Edwards TP, et al. Characterizing biobank organizations in the U.S.: Results from a national survey. *Genome Med* 2013;5:3.
9. Boyer GJ, Whipple W, Cadigan RJ, et al. Biobanks in the United States: How to identify an undefined and rapidly evolving population. *Biopreserv Biobank* 2012;10:511–517.
10. Shaw DM, Elger BS, Colledge F. What is a biobank? Differing definitions among biobank stakeholders. *Clin Genet* 2014;85:223–227.
11. Compton C. The Revisioning of caHUB[®] as a Center for Biospecimen Science and Standards Development. Available at: <http://biospecimens.cancer.gov/meeting/brnsymposium/2011/docs/Compton-508.pdf> Accessed September 14, 2015.
12. Dvinge H, Ries RE, Ilagan JO, et al. Sample processing obscures cancer-specific alterations in leukemic transcriptomes. *Natl Acad Sci USA* 2014;111:16802–16807.
13. Carlson RH. Poor Quality of Biospecimens Called #1 Limiting Factor in Translational Research. *Oncology Times*. November 10, 2010. Available at: http://journals.lww.com/oncology-times/Fulltext/2010/11100/Poor_Quality_of_Biospecimens_Called__1_Limiting.9.aspx Accessed September 14, 2015.
14. Ellervik C, Vaught J. Preanalytical variables affecting the integrity of human biospecimens in biobanking. *Clin Chem* 2015;61:914–934.
15. Campbell LD, Betsou F, Garcia DL, et al. Development of ISBER best practices for repositories: Collection, storage, retrieval, and distribution of biological materials for research. *Biopreserv Biobank* 2012;10:232–233.
16. Biorepositories and Biospecimen Research Branch. NCI Best Practices. Available at: <http://biospecimens.cancer.gov/practices> Updated July 28, 2014. Accessed September 14, 2015.
17. College of American Pathologists. Biorepository Accreditation Program. Available at: www.cap.org/web/home/lab/accreditation/biorepository-accreditation-program?_afLoop=317441661996566#!%40%40%3F_afLoop%3D317441661996566%26_adf.ctrl-state%3D08ilh64_30 Accessed September 20, 2016.
18. Holub P, Swertz M, Reihls R, et al. BBMRI-ERIC Directory: 515 biobanks with over 60 million biological samples. *Biopreserv Biobank* 2016;14:559–562.
19. Norlin L, Fransson MN, Eriksson M, et al. A minimum data set for sharing biobank samples, information, and data: MIABIS. *Biopreserv Biobank* 2012;10:343–348.
20. U.S. Food and Drug Administration. Therapeutic Area (Disease/Domain) Data Standards Prioritization. Available at: www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM297093.pdf. Updated July 17, 2015. Accessed September 14, 2015.
21. Health Level Seven[®] (HL7) International. HL7 standards—section 3: clinical and administrative domains. Health Level Seven[®] International web site. Available at: www.hl7.org/implement/standards/product_section.cfm?section=3&ref=nav. Accessed September 14, 2015.
22. Nahm M, Walden A, McCourt B, et al. Standardizing clinical data elements. *Int J Funct Inform Personal Med* 2010; 3:314–341.
23. International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC). ISO/IEC 11179, information technology—metadata registries (MDR). Available at: <http://metadata-standards.org/11179/>. Accessed September 14, 2015.
24. Biospecimen Research Network. Lifecycle of Biospecimens. Available at: <http://biospecimens.cancer.gov/researchnetwork/lifecycle.asp>. Accessed September 15, 2016.
25. National Cancer Institute. caDSR Content. Available at: <https://wiki.nci.nih.gov/display/caDSR/caDSR+Content>. Accessed September 15, 2016.
26. Robb JA, Bry L, Sluss PM, et al. A call to standardize preanalytic data elements for biospecimens, part II. *Arch Pathol Lab Med* 2015;139:1125–1128.
27. Moore HM, Kelly AB, Jewell SD, et al. Biospecimen reporting for improved study quality (BRISQ). *J Proteome Res* 2011;10:3429–3438.
28. Sioutos N, de Coronado S, Haber MW, et al. NCI Thesaurus: A semantic model integrating cancer-related clinical and molecular information. *J Biomed Inform* 2007;40:30–43.
29. de Coronado S, Haber MW, Sioutos N, et al. NCI Thesaurus: Using science-based terminology to integrate cancer research results. *Stud Health Technol Inform* 2004;107:33–37.
30. Common biorepository Model (CBM). Available at: <https://github.com/NCIP/common-biorepository-model>. Accessed September 15, 2016.
31. Brochhausen M, Fransson MN, Kanaskar NV, et al. Developing a semantically rich ontology for the biobank-administration domain. *J Biomed Semantics* 2013;4:23.
32. Horvath MM, Winfield S, Evans S, et al. The DEDUCE Guided Query tool: Providing simplified access to clinical data for research and quality improvement. *J Biomed Inform* 2011;44:266–276.
33. Horvath MM, Rusincovitch SA, Brinson S, et al. Modular design, application architecture, and usage of a self-service model for enterprise data delivery: The Duke Enterprise Data Unified Content Explorer (DEDUCE). *J Biomed Inform* 2014;52:231–242.
34. Brochhausen M, Zheng J, Birtwell D, et al. OBIB—a novel ontology for biobanking. *J Biomed Semantics* 2016;7:23.
35. GitHub website. Biobanking. Available at: <https://github.com/biobanking> Accessed September 15, 2016.
36. GitHub website. Informed Consent Ontology. Available at: <https://github.com/ICO-ontology/ICO> Accessed September 15, 2016.

Address correspondence to:
Helena Ellis, BS
Biobanking Without Borders, LLC
4812 Stafford Drive
Durham, NC 27705

E-mail: helena.ellis@biobankingwithoutborders.com