Evaluating Generalizability of a Biospecimen Informatics Approach: Support for Local Requirements and Best Practices

Cindy Chen, M.A.[1], Regina T. Wulff, M.S.[2], Evan T. Sholle, M.S.[1], Gail J. Roboz, M.D.[2], David A. Kraemer, Jr., B.S.[1], Thomas R. Campion, Jr. Ph.D.[1,3,4,5]

[1]Information Technologies and Services Department, Weill Cornell Medicine, New York, NY; [2]Department of Medicine, Weill Cornell Medicine, New York, NY [3]Department of Healthcare Policy and Research, Weill Cornell Medicine, New York, NY; [4]Clinical and Translational Science Center, Weill Cornell Medicine, New York, NY; [5]Department of Pediatrics, Weill Cornell Medicine, New York, NY;

Abstract

To enable clinical and translational research, academic medical centers increasingly implement biospecimen information management systems. At our institution, one laboratory successfully implemented a multi-system solution that enabled collection and reporting of specimen- and aliquot-level data. The objective of this study was to assess the solution against the laboratory's requirements and with respect to support of best practices for biospecimen information management systems defined by the International Society for Biological and Environmental Repositories (ISBER). The solution supported the laboratory's reporting needs and 90% (n=26) of ISBER best practices. To the best of our knowledge, this is among the first studies to demonstrate the generalizability of a biospecimen informatics approach. Findings suggest that development and evaluation of biospecimen informatics approaches can potentially improve through closer collaboration of informatics and biorepository professional societies.

Introduction

High-quality biospecimens are crucial to biomedical research, and academic medical centers increasingly require biospecimen information management systems (BIMS) capable of tracking availability, location, and metadata of biological materials as well as integrating with clinical data from electronic health record (EHR) systems (1,2). Although several studies have described novel biospecimen information management systems implemented in single institutions (3–9), few investigations have evaluated the generalizability of approaches to other settings. The goal of this study was to test the hypothesis that an existing biospecimen informatics approach (10) could support researchers at our institution. We evaluated the approach's support of local requirements and best practices for biospecimen information management systems as defined by the International Society for Biological and Environmental Repositories (ISBER).

Methods

Setting

The Weill Cornell Medicine (WCM) Physician Organization constitutes a multi-specialty group practice with over 900 physicians serving more than 2 million patients at more than 20 clinics across the New York City area. All WCM physicians have admitting privileges to NewYork-Presbyterian Hospital (NYP), a long-time teaching affiliate. In addition to clinical care, WCM serves medical education and biomedical research missions, the latter of which includes a Clinical and Translational Science Award (CTSA) hub and several core facilities.

The Leukemia Program is a major clinical referral center with expertise in acute myeloid leukemia, myelodysplastic syndrome, and other disorders. To support research in this population, the Leukemia Program Laboratory provides a biorepository and testing facility. A manager and two technicians support daily operations of the laboratory.

For documenting clinical care, WCM physicians use EpicCare Ambulatory in outpatient clinics and Allscripts Sunrise Clinical Manager in the inpatient setting. Separate information technology teams from WCM and NYP oversee the outpatient and inpatient clinical systems. As described elsewhere, the WCM Research Informatics group enables secondary use of data from institutional EHR systems as well as support of research-specific applications (11).

In 2012, the Leukemia Program Laboratory identified the need for a biospecimen information management system to replace its Microsoft Excel-based approach. This legacy approach consisted of manual entry of data elements into spreadsheets describing banked samples, aliquots, laboratory results, demographics, and diagnoses. Use of spreadsheets complicated reporting. Goals of the new system included streamlining biospecimen recordkeeping,

linking to clinical data in local EHR systems, and supporting compliance with the College of American Pathologists (CAP) and Clinical Laboratory Improvements Amendments (CLIA). Because WCM's institutional biorepository core facility was in the early stages of development and lacked an enterprise-wide BIMS, the Leukemia Program Laboratory, with approval from the Office of the Research Dean, elected to fund and implement a tailor-built system to support its needs.

In developing this solution, the Leukemia Program collected requirements from laboratory personnel, clinicians, and researchers. Laboratory personnel provided these requirements in the form of a specific suite of reporting requirements, which drove the development of the approach. Additionally, the group assessed publicly available systems for supporting biospecimen information management, retained the services of a consultant to advise on system selection. After the selection of RURO FreezerPro Elite (referred to in this text as "FreezerPro"), the group conferred with the WCM information technology department about hosting and security for the system. To support system implementation and legacy data migration, the Leukemia Program Laboratory allocated effort of a laboratory manager and technician while the WCM Research Informatics group dedicated a software engineer and project manager.

Approach for biospecimen information management

To support the Leukemia Program Laboratory, we implemented a solution based on the experience of Chery and colleagues (10) that combined REDCap, an electronic data capture system, for specimen collection and parent specimen data (12), FreezerPro for child aliquot data (13), and Microsoft SQL Server for reporting. As described below, laboratory workflow dictated the decision to use multiple applications rather than one for storing and reporting specimen- and aliquot-level data.

Use of REDCap for storing sample collection data

For each specimen collected, laboratory personnel created a new record in a REDCap project. The REDCap project consisted of a single instrument with multiple fields and automatically generated a unique twelve-digit identifier for each record. Each specimen arrived in the laboratory with a paper manifest identifying the study for which the specimen was collected, the type and time of sample collected, and the patient from whom the specimen was collected, which laboratory personnel transcribed to REDCap. Based on entry of a patient's medical record number (MRN), the Dynamic Data Pull (DDP) plugin (14) automatically populated REDCap form fields with patient demographics and diagnoses retrieved from the WCM Epic EHR. After completing specimen registration in REDCap, laboratory personnel copied-and-pasted the specimen's twelve-digit unique identifier from REDCap to FreezerPro. Laboratory personnel accessed REDCap via the institution's Microsoft Active Directory (Active Directory) authentication with usernames and passwords, and REDCap project-specific user groups determined access to system features and data.

Use of FreezerPro for storing aliquot data

To support the needs of each study collection protocol, laboratory personnel created one or more records in FreezerPro to represent the aliquots obtained from each parent sample. For example, if Study Z required creation of five aliquots from a specimen with the twelve-digit identifier 567800000002 assigned by REDCap, then laboratory personnel created five new records in FreezerPro, copied 567800000002 from REDCap, and pasted 567800000002 into each new aliquot record in FreezerPro. Each aliquot record, which had a unique identifier automatically generated by FreezerPro, described laboratory processing and corresponded to a freezer location to enable subsequent retrieval. While REDCap stored data describing each parent sample, FreezerPro stored data describing the child aliquots obtained from each parent sample. This solution was designed to facilitate subsequent identification of available aliquots based on queries that leveraged data recorded in REDCap, including EHR data derived via DDP.

As shown in Table 1, we used FreezerPro standard and user-defined fields to store aliquot data. Of standard fields, FreezerPro captured the location of each aliquot with respect to a storage container, such as a freezer and box within a freezer. We used sample groups to represent studies approved by the WCM Institutional Review Board (IRB) under which collections occurred, sample sources to represent patients as study participants, and sample types to represent the types of tissues and liquids collected. With user-defined fields, we captured details of specimen processing and viability that did not exist in native FreezerPro forms. Both standard and user-defined fields enabled system administrators to configure field types (e.g., free text, dropdown menu) with options for discrete fields.

Table 1. FreezerPro standard fields and user-defined fields.

Standard Fields	User-Defined Fields
Name	Legacy ID

Description	Process method
Freezer location	Process start time
Sample source	Storage media used
Sample group	Total viable cells
Sample type	Method of sample destruction
Sample owner	Nucleic acid concentration by qubit
	Percent viability
	Final destination
	Percent purity by flow

To support Good Laboratory Practice (GLP) for specimen handling and recordkeeping (15), we enabled FreezerPro's GLP mode. GLP mode required laboratory personnel to enter a comment after modifying any existing aliquot record, and prevented users from updating multiple records in a single operation. For example, a user could not simultaneously update the "process start time" field for three aliquots centrifuged in the same batch. Instead, a user needed to perform three separate operations to update the "process start time" field of each aliquot record with the same value.

In addition to aliquot attributes, FreezerPro enabled creation of user groups to define roles and permissions, restricting access to specific freezers, sample types, user-defined fields, and other elements. As with the REDCap project, laboratory personnel accessed FreezerPro with their existing usernames and passwords through institutional Active Directory authentication infrastructure.

Use of Microsoft SQL Server for reporting sample- and aliquot-level data

Although REDCap and FreezerPro enabled users to generate reports containing data from within each system, we used Microsoft SQL Server 2014 to deliver reports integrating data from both systems. To obtain sample-level data, we configured a linked server between REDCap's MySQL database and a Microsoft SQL Server database. Similarly, to obtain aliquot-level data, we developed a Python script to extract records from FreezerPro via its application programming interface (API) to a Microsoft SQL Server database hosted on the same server. We configured both methods to refresh data regularly. With data from both systems in Microsoft SQL Server databases, we joined data using the twelve-digit unique identifier for samples common to both systems. Subsequently, we made reports containing data from both systems available to laboratory personnel using web-based Microsoft SQL Server Reporting Studio, which provides user- and role-based access to configurable reports that ran against the regularly-refreshed source data.

Rationale for separate applications for storing specimen-, sample-, and aliquot-level data

Our development of this two-tiered process, whereby sample data was recorded in REDCap and aliquot data in FreezerPro, was driven by the differing capabilities of the two systems. While laboratory personnel initially planned to store specimen-, sample-, and aliquot-level data in FreezerPro, a mismatch between system functionality and laboratory workflow made this impracticable. To record the various levels of data in FreezerPro with Good Laboratory Practice mode enabled, laboratory personnel needed to re-enter sample data for each new aliquot created in the system. With up to twenty fields per sample, laboratory personnel identified the time required to complete the data entry in FreezerPro as a significant impediment to laboratory workflow. However, laboratory personnel also valued GLP mode for aliquot-level data, especially for its support of CLIA and CAP certification goals. To reconcile the competing priorities of ease of data entry and adherence to industry standards, laboratory and informatics personnel elected to use REDCap and FreezerPro for specimen and sample data and aliquot data storage, respectively.

Data migration and system go-live

Migration of legacy data from Microsoft Excel to REDCap and FreezerPro required several iterations between laboratory personnel and the software engineer to verify data quality and ensure referential integrity within and across systems. This involved close collaboration to develop a common standard to verify the data against and provide

structure for the reporting data elements. The process required two months of ongoing review of legacy data, verification of business rules, and updates to records.

In April 2016, the Leukemia Program started using the novel biospecimen information management solution combining REDCap and FreezerPro for 2,264 samples and 19,726 related aliquots. As of December 2016, the system contained records for nearly twice as many samples (n=4,615) and almost twice as many aliquots (n= 36,481) as at go-live. Increased system usage over time and general positive feedback from laboratory staff suggest that the combination of REDCap and FreezerPro has succeeded as a biospecimen information management solution for the Leukemia Program.

Evaluation

We assessed our solution (hereafter FreezerCap-SQL) with regard to the extent to which it met the Leukemia Program's requirements. To perform this component of the evaluation, laboratory personnel identified a list of nine reporting examples, divided into three principal categories by the user type who might request the report (clinicians, researchers, and laboratory staff). Three additional reporting examples were designed to assess the fitness of the system to respond to particular technical challenges.

We also assessed our solution with respect to ISBER best practices for BIMS. Specifically, the ISBER Informatics Working Group (16) has identified 29 best practices across five areas of features—controlled vocabulary and data integrity; security, privacy, and auditing; subject management; biospecimen management; analytics and reporting; and technical/API interoperability—as documented in the Information System Evaluation Checklist (17). For each of the 29 best practices, the study team (DAK, CC, TRC) assessed if and how FreezerCap-SQL adhered to the standard. We resolved disagreements through discussion to reach a consensus.

Results

The FreezerCap-SQL approach successfully supported the nine reporting requirements specified by the Leukemia Program and described in Table 2. Based on review of Microsoft SQL Server access logs, each day users run 2-6 reports that combine data from REDCap and FreezerPro.

Table 2. Reporting requirements specified by Leukemia Program personnel

User role	Reporting requirement
Clinician	For Patient Y: How many (parent) specimens were submitted to the biobank? How many bone marrow PLASMA sample aliquots were submitted by each collection protocol? List the aliquots and their locations.
Clinician	For Diagnosis X: How many (parent) specimens were submitted to the biobank? How many bone marrow PLASMA sample aliquots were submitted by each collection protocol? List the aliquots and their locations.
Clinician	For Protocol Z: How many (parent) specimens were submitted to the biobank? How many bone marrow PLASMA sample aliquots are stored with LEUKEMIA STUDY IDs? With TRANSPLANT STUDY IDs? List the aliquots and their locations.
Researcher	How many biobanked samples fit the following description: 50-70 year old female (age frozen in time), CD15+ WBC Cell Pellet stored, aliquot >95% Pure by flow (regardless of diagnosis)? List the aliquots and their locations.
Laboratory personnel	How many biobanked samples are from >90 year old females? (age is frozen in time)? How many by sample type (PLASMA, MNC CELL PELLET, MNC CRYOPRESERVED, etc.)? List the aliquots and their location.
Laboratory personnel	What is the total average turn-around-time from specimen collection to specimen storage? By collection protocol? By process method? By technician?
Laboratory personnel	From today's date, how many MNC CRYOPRESERVED aliquots have been stored for >10 years? What are the "same day" CBC characteristics for those aliquots? List the aliquots and their location.
Any	A single report of all available data in FreezerPro given 20+ unique Specimen ID numbers? *(This exceeds the 10 search fields available through FreezerPro's search function)*
Any	Identify all peripheral blood plasmas less than 6 hours old

Of the 29 ISBER best practices evaluated, the FreezerCap-SQL solution supported 90% (n=26). FreezerCap-SQL supported the feature areas of controlled vocabulary and data integrity; security, privacy, and auditing; and technical/API interoperability. Feature areas where FreezerCap-SQL failed to support best practice included subject management; biospecimen management; and analytics and reporting. Specifically, for subject management features, FreezerCap-SQL did not "…[provide] ability to track and manage events associated with a particular subject and date (e.g., visit, donation, examination)." Similarly, for analytics and reporting features, the solution failed to "[provide] the user with the ability to define ad hoc queries/searches and custom reports in common terms, without requiring knowledge of proprietary code."

Table 3. Evaluation of FreezerCap-SQL with respect to ISBER best practices reproduced from (16).

ISBER Information System Evaluation Checklist - Best Practices	FreezerCap-SQL
Controlled Vocabulary and Data Integrity Features	
System provides ability to maintain controlled vocabularies (ontologies) to enforce data standardization and control.	Yes
System provides users with intuitive on-demand access to data and the ability to represent the data (for view, file export or print) in a variety of formats without knowledge of proprietary code.	Yes
Security, Privacy and Auditing Features	
System is capable of maintaining user profiles and credentialing users for different levels of access and functionality.	Yes
System provides audit trail that minimally tracks userid, date, and content of change at the field level	Yes
Subject Management Features	
Data elements pertaining to each subject are sufficiently extensive or extensible (e.g. the ability to maintain demographic attributes including race, ethnicity, date of birth, gender)?	Yes
System provides ability to track and manage events associated with a particular subject and date (e.g. visit, donation, examination, etc.)	No
Provides ability to associate a subject with a study	Yes
Provides ability to assign system-generated, unique subject identifier	Yes
Provides ability to manage subject de-identification.	Yes
Biospecimen Management Features	
System provides ability to generate/assign a unique accession to each sample.	Yes
System provides ability to define an unlimited number of samples: whole-blood, tissue, cellular lysates, DNA, RNA, proteins, etc.	Yes
System provides ability to assign sample component with physical location	Yes
System provides ability to associate the sample and component to the participant.	Yes
System provides ability to annotate sample with sample attributes, like method of specimen preparation and environmental conditions under which specimen is stored: (e.g. type of sample, volume, container size, description, date drawn, source of sample, person storing, temperature.)	Yes
System provides ability to a user with special permissions to define hierarchical storage configurations.	Yes
System provides ability to track specimens with barcoded IDs printed on labels	Yes

Provides ability to query/search inventory of specimen and specimen components.	Yes
System maintains sample genealogy on aliquots, derivative, and pooled samples (e.g. DNA derived from PBMC derived from whole blood)	No
System provides logistics management and chain-of-custody tracking (e.g. shipping and receiving)	Yes
System provides ability to configure rules or restrictions to manage access to specimen information.	Yes
System provides mechanism for maintaining specimen lifecycle and disposition (e.g., system tracks amount used and decrements from available amount).	Yes
System provides ability to create and maintain complex queries/searches using associated subject attributes or experiment data as search criteria	Yes
Analytics and Reporting Features	
Provides the user with a defined process for creating queries/searches on data in the system.	Yes
Provides ability to create and maintain reports as standard reports that can be selected by a user from a list	Yes
Provide the user with the ability to define ad hoc queries/searches and custom reports in common terms, without requiring knowledge of proprietary code.	No
Provide the user with the ability to save queries/searches for future reuse.	Yes
System provides ability to export data in delimited formats: .csv, or .xls and XML	Yes
Technical Features	
System has an integrated database. Describe the database management system (DBMS) that is most commonly used with the System.	Yes
System is capable of interfacing with third party data sources, applications and services.	Yes

Discussion

In a laboratory at our institution, we implemented an existing biospecimen informatics approach (10), which met the needs of laboratory personnel and supported 90% of best practices. Our evaluation indicated that the approach can meet needs beyond the institution where it was developed. To the best of our knowledge, this is among the first studies to demonstrate the generalizability of a biospecimen informatics approach. Development and evaluation of biospecimen informatics approaches can potentially improve through closer collaboration of informatics and biorepository professional societies.

While the list of ISBER best practices and Leukemia Program reporting requirements by which we evaluated the FreezerCap-SQL approach may not constitute an exhaustive list of requirements, it nonetheless offers one metric by which to assess the validity of our novel solution for biospecimen management. Other metrics, including those proposed by Prokosch and colleagues (18), may have provided alternate methods to determine relative performance, but they appeared difficult to operationalize compared to the checklist format of the ISBER best practices. The results of a study conducted at Duke University (19), which demonstrated the value of collaboration between informatics and biorepository experts, and our investigation suggest that formal collaboration between ISBER and the American Medical Informatics Association may improve implementation and evaluation efforts for biospecimen information management systems.

One principal limitation of this analysis is its failure to compare the success of the FreezerCap-SQL approach to other techniques, including both open-source solutions, such as OpenSpecimen (20), and vendor-offered full-service BIMS. While a simulation of these solutions could have provided an indirect assessment of the relative performance of the FreezerCap-SQL approach, it would be incapable of assessing the real-world performance of the comparison solution,

as it would be impracticable to have laboratory personnel fully utilize both systems on a day-to-day basis. FreezerCap-SQL still requires laboratory personnel to manually align identifiers between aliquots and their parent specimens, which other solutions, such as OpenSpecimen or commercial products, may obviate. Future work may address the suitability of the FreezerCap-SQL approach to supporting the NCI Common Biorepository Model (21) as well as the potential of addressing the ISBER best practice of patient event capture using REDCap.

A particular strength of the FreezerCap-SQL approach is the relative ease by which informatics personnel can enable additional reporting based on data from other systems, including research electronic data capture systems and the EHR. Utilizing the REDCap DDP plugin allows for queries that return samples only for patients who meet specific criteria determined by EHR data, a unique strength of this approach that may have gone undiscovered had we adopted an alternate approach. Additionally, the modular nature of the SQL reporting structure allows for the easy addition of filters in SQL Server Reporting Studio that derive from other systems that track the REDCap specimen number, including research electronic data capture systems. Laboratory personnel have identified this as a particular strength of this approach, as investigators often require samples for particular patient cohorts as defined by data elements that exist only in the EHR or in research electronic data capture (EDC) systems. However, the FreezerCap-SQL approach still requires close collaboration between laboratory and informatics personnel to generate new reports, as a comprehensive business intelligence tool allowing users to query any and all data elements without coding expertise does not exist, to the best of our knowledge.

The fact that informatics personnel did not engage in requirements gathering until relatively late in the selection process may have had some impact on the development of the methodology. However, it is difficult to say whether earlier engagement from informatics professionals would have identified the key unit-of-analysis distinction within FreezerPro that necessitated the addition of REDCap for sample-level data tracking.

A crucial element of the development of the FreezerCap-SQL approach was the centrality of collaboration between informatics and laboratory personnel. The top-down imposition of a vendor-provided biospecimen information management system would have proved unable to easily address the requirement of a reporting system capable of integrating query elements from external systems, including the EHR and various research EDC systems in use by the Leukemia Program. By working together with laboratory personnel, the informatics team was able to identify this multi-system solution that both adheres to industry best practices and directly addresses the identified requirements of the system's users.

As with many institutions, our approach relied on ad hoc adjustment and integration of disparate systems rather than an out-of-the-box approach. In hindsight, a full-fledged biospecimen information management system capable of tracking biospecimen data at specimen, sample, and aliquot levels in compliance with GLP standards may have provided a preferable solution; however, the FreezerCap-SQL methodology has proved to be a workable approach that meets investigator and laboratory personnel needs while adhering to many industry best practices. Close collaboration between laboratory and informatics personnel is critical for successful implementation of a biospecimen information management system.

Acknowledgments

References

1.   Litton J-E. Biobank informatics: connecting genotypes and phenotypes. Methods Mol Biol [Internet]. 2011 Jan [cited 2015 Apr 28];675:343–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20949402
2.   Tenenbaum JD, Avillach P, Benham-Hutchins M, Breitenstein MK, Crowgey EL, Hoffman MA, et al. An informatics research agenda to support precision medicine: seven key areas. J Am Med Informatics Assoc [Internet]. 2016 Jul 1 [cited 2017 Aug 28];23(4):791–5. Available from: https://academic.oup.com/jamia/article-lookup/doi/10.1093/jamia/ocv213
3.   Dangl A, Demiroglu SY, Gaedcke J, Helbing K, Jo P, Rakebrandt F, et al. The IT-infrastructure of a biobank for an academic medical center. In: Studies in Health Technology and Informatics. 2010.
4.   Prokosch H-U, Mate S, Christoph J, Beck A, Köpcke F, Stephan S, et al. Designing and implementing a biobanking IT framework for multiple research scenarios. Stud Health Technol Inform [Internet]. 2012 [cited 2017 Aug 26];180:559–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22874253
5.   Segagni D, Tibollo V, Dagliati A, Perinati L, Zambelli A, Priori S, et al. The ONCO-I2b2 project:

integrating biobank information and clinical data to support translational research in oncology. Stud Health Technol Inform [Internet]. 2011 [cited 2017 Aug 26];169:887–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21893874

6. Hofer P, Fiegl H, Angerer J, Mueller-Holzner E, Chamson M, Klocker H, et al. A concept of a MIABIS based register of biosample collections at the Medical University of Innsbruck. Stud Health Technol Inform [Internet]. 2014 [cited 2017 Aug 26];205:293–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25160193

7. Boutin N, Holzbach A, Mahanta L, Aldama J, Cerretani X, Embree K, et al. The Information Technology Infrastructure for the Translational Genomics Core and the Partners Biobank at Partners Personalized Medicine. J Pers Med [Internet]. 2016 Jan 21 [cited 2017 Feb 21];6(1):6. Available from: http://www.mdpi.com/2075-4426/6/1/6

8. Lasalle B, Varner M, Botkin J, Jackson M, Stark L, Cessna M, et al. Biobanking informatics infrastructure to support clinical and translational research. AMIA Jt Summits Transl Sci proceedings AMIA Jt Summits Transl Sci [Internet]. 2013 [cited 2017 Aug 26];2013:132–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24303252

9. Anderson N, Ponko S, Black T, Prosser J, Stein B, Tarczy-Hornoch P, et al. Biotrust: a comprehensive system for acquiring and distributing biospecimens. AMIA Jt Summits Transl Sci proceedings AMIA Jt Summits Transl Sci [Internet]. 2013 [cited 2017 Sep 25];2013:4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24303285

10. Chery L, Maki JN, Mascarenhas A, Walke JT, Gawas P, Almeida A, et al. Demographic and clinical profiles of Plasmodium falciparum and Plasmodium vivax patients at a tertiary care centre in southwestern India. Malar J [Internet]. 2016 Dec 25 [cited 2017 Sep 27];15(1):569. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27884146

11. Sholle E, Kabariti J, Johnson S, Leonard J, Pathak J, Varughese V, et al. Secondary use of patients' electronic records (SUPER): an approach for meeting specific data needs of clinical and translational researchers. AMIA Annu Symp Proc. 2017;In press.

12. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform [Internet]. 2009 Apr [cited 2013 Sep 20];42(2):377–81. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2700030&tool=pmcentrez&rendertype=abstract

13. FreezerPro® – RURO, Inc [Internet]. [cited 2017 Sep 25]. Available from: https://ruro.com/products/freezerpro

14. Campion TR, Sholle ET, Davila MA. Generalizable Middleware to Support Use of REDCap Dynamic Data Pull for Integrating Clinical and Research Data. AMIA Jt Summits Transl Sci proceedings AMIA Jt Summits Transl Sci [Internet]. 2017 [cited 2017 Sep 25];2017:76–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28815111

15. eCFR — Code of Federal Regulations [Internet]. [cited 2017 Sep 29]. Available from: https://www.ecfr.gov/cgi-bin/text-idx?SID=83c258e564365061c7fdb7f6302fbab3&mc=true&node=pt21.1.58&rgn=div5

16. Informatics Working Group - ISBER [Internet]. [cited 2017 Aug 28]. Available from: http://www.isber.org/page/INF

17. ISBER Information System Evaluation Checklist [Internet]. [cited 2017 Sep 25]. Available from: https://www.isber.org/resource/resmgr/documents/ISBERInformationSystemEvalu.xlsx

18. Prokosch HU, Beck A, Ganslandt T, Hummel M, Kiehntopf M, Sax U, et al. IT Infrastructure Components for Biobanking. Appl Clin Inform [Internet]. 2010 Jan [cited 2015 Jun 11];1(4):419–29. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3633318&tool=pmcentrez&rendertype=abstract

19. Ellis H, Joshi M-B, Lynn AJ, Walden A. Consensus-Driven Development of a Terminology for Biobanking, the Duke Experience. Biopreserv Biobank [Internet]. 2017 Apr [cited 2017 Sep 25];15(2):126–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28338350

20. McIntosh LD, Sharma MK, Mulvihill D, Gupta S, Juehne A, George B, et al. caTissue Suite to OpenSpecimen: Developing an extensible, open source, web-based biobanking management system. J Biomed Inform [Internet]. 2015 Oct [cited 2017 Sep 25];57:456–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26325296

21. No Title [Internet]. Available from: https://github.com/NCIP/common-biorepository-model