APPLICATION NOTE



A COLLABORATIVE TRANSLATIONAL INFORMATICS ENVIRONMENT AT THE NCI

Jason Levine¹, John Otridge², Steve Chen² ¹Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD ²BioFortis Inc., Columbia, MD

ABSTRACT

Translational research programs have been created to provide a better understanding of disease susceptibilities and pathway mechanisms. These programs require integration and analysis of multiple disparate data sets that come from different sources such as the hospital and pathology/molecular laboratories. At the NIH, a subset of the patients treated at the Clinical Center consent into various Institutes' research protocols. Here we describe how NCI investigators utilize a centralized database system that 1.) securely stores, annotates, curates, and tracks information such as patient data, clinical phenotypes, biospecimen & derivatives data, and experimental research data, 2.) communicates clinical and research information with the NCI Clinical Data Registry and BTRIS systems, 3.) enables collaborative exchange and sharing of information amongst research groups, and 4.) provides an intuitive environment for investigators to query and review their collected data with minimal need for direct IT support.

BACKGROUND

When patients are admitted at the NIH, their demographics information, orders, results, participating protocols, and other clinical care information are stored in the Clinical Center's central electronic medical record database, Clinical Research Information System (CRIS). NCI Center for Cancer Research (CCR) maintains its own separate Clinical Data Registry for patients who are enrolled in intramural NCI protocols. However, investigators traditionally have been "on their own" to capture, annotate, and store research information from consented, participating patients. As a result, many investigators managed clinical and research data in disparate information systems, often relying on spreadsheets and homegrown databases, which led to suboptimal retrieval and exchange of information that reduced the efficiency and effectiveness of translational medical research. Beginning in 2006, the Labmatrix software environment was introduced to CCR researchers with the following areas of emphasis and capabilities:

I. ACCESS & SECURITY

CCR's Labmatrix system is a clinical research data management platform that supports intramural research endeavors. Centrally-hosted and maintained on-site behind NIH's firewall, the system utilizes a web-browser thin-client/server model that allows for quick set up and management of users and groups based on laboratories, collaborations, research domains, and IRB protocols. While a set of central data resources are available to all Labmatrix users, a flexible security framework ensures private research data will only be accessible to designated group members and collaborators; furthermore, specific types of transactions against all data records (search, create, view, edit, delete) are configurable to ensure proper level of access. Individual's access to Labmatrix is initially facilitated by the existing user authentication authority at the NIH (Active Directory); upon signing in, the user's Labmatrix permissions are calculated via the individual's user group(s) assignment. To prevent misuse and interception of PHI, PII, and private research data, all user interactions with the system are encrypted while in transmission. Additionally, data values are logged in pre/post-state details for audit & security forensics.

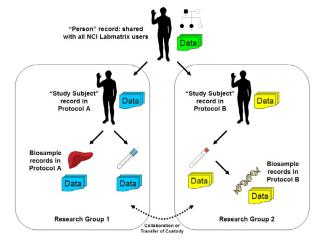
ser Group - * Top-Lew	a Pill Circa	ue years	Incase								Administrator Console Access Groups Barcoding IRB Pro	tocols License Report
neral Permissions												Search Results
lick the "Add" button below fter adding an Access Gro	to select a	n Acces fil des	s Group on r n.by right-c	which I Solding	to grant per a column he	nission ader ar	fil acress	by right-	ddingar	ndule. N	IRB Protocol Subject Attribute 1 Tissue Sample Sample Type S	Search Results
Lab	Search		ties		ORME		Edit		Delete		Name	Access Group(s
Crgonizations Referch	Search		View		Create		Ed:		Delete			
Patients Minimitative	Search		Vievs		Create		Edit		Delete	-	carotid sheath, left	Shared Vocab
Bonaterials		~				×		~		~	carotid sheath, right	Shared Vocab
Communications			_		_	×	-		-	~	cervix	Shared Vocab
Contects	P+C		PIE		94	¥	P-C		P+E	~	calan	Shared Vocab
Custam Data		v		v		¥		-			colon, ascending	Shared Vocab
Family		V		v		×		~		*	colon, descending	Shared Vocab
Genotype Detections		¥				¥						
Genotype Sequence	~ PH	~	100	×	0.040	M	A PHC		- Ped	~	colon, hepatic flexure	Shared Vocab
Cenetype Variation	1 PH 6		1916		1 PH C	Y	1940	×	194	~	colon, sigmoid	Shared Vocab
Imports				¥		¥		¥			colon, spienic flexure	Shared Vocab
Pedgree		~				*		×		~	colon, transverse	Shared Vocab
Query Builder		1		~		¥		×		~	common bie duct	Shared Vocab
Storage						۲		*				
Subjects		~		~		*		*		*	diaphragm	Shared Vocab
Morkfow		~		~		~		~		*	duodenum	Shared Vocab
Personnel Protected Contacts	Search		Vien		Oreste		Edit		Delete		C Records 1-100 of 165 total found.	
Prevenues Contacts	10 and		1.00		- veste		eurs.		144.36		The records a second second second build.	

User Group Permissions Management

len	Delete			9 80 🖓 🕻
IRB I	Protocol - (PO	0) 04-C-0102		
TRA	Protocol			
	Name	(POR) 04-C-0102		
Are		Hematologic Malgrancy Biology		
	Numbers	0+C-0102		
	Description:	To allow sample acquisition for use	in the study of hematologic mai	Ignancies.
		A variaty of laboratory investigation	is designed to support National	institutes
		of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic	
Con	sent Form(s)	of Health translational trials, apply molecular, genetic, and genomic b	iology, probe for new therapeutic	
Non	ne i File	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	stargets,
Nan Orig	ne i Filo Inal	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	: targets, : Signed Consent Forr false
Non Orig Ame	nc File inal endment A	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	signed Consent Forr false
Non Orig Ame Ame	nc Filo inal indment A indment B	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	signed Consent Form false false
Non Orig Ame Ame	nc Filo inal indment A indment B indment C	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	signed Consent Forr falce falce falce falce
Non Orig Ame Ame Ame	nc File inal indment A indment B indment C indment D	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	: targets, : Signed Consent Form false false false false false false false false
Nan Orig Ame Ame Ame Ame	nc Filo inal indment A indment B indment C	of Health translational trials, apply molecular, genetic, and genomic b and develop new treatment approx	iology, probe for new therapeutic iches, will be performed.	signed Consent Forr falce falce falce falce

II. STANDARDIZATION & WORKFLOWS

Several key goals for the CCR Labmatrix system are to establish an informatics environment that facilitates data standardization, reduce data replication and divergence, encourage intramural collaborative participation, and prioritize re-use of existing informatics components. Commonlyused data domains in translational research, such as patient-centric information (e.g. contact details, demographics, phenotypes, genotypes, and familial relationships) and biospecimen-centric information (e.g. physical measurements, creation/derivation/ tracking details, and physical storage assignments) are supported with Labmatrix's built-in data forms, fields, and business logic. For research groups that want to capture information unique to their data workflows, custom data collection forms and relational data models can be created. In order to facilitate interoperability, ease-of-(re)use, and ease-of-maintenance, entities such as vocabulary terminologies, data sources, data forms, and workflows can be either shared or private resources in the system.

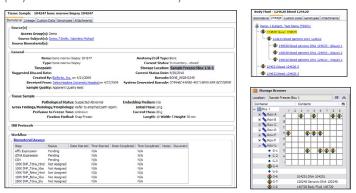


When a patient is enrolled and consented into a research protocol at the NCI, the investigating team can create a single reference "person" record in Labmatrix and auto-populate data from CRIS or from the NCI Clinical Data Repository; data for the person record are then shared with all Labmatrix users. Individual investigators can then assign the person into one or more studies with corresponding subject records. Protocol-specific data are entered into these access-controlled study records. Data access permissions can be configured to support multiple needs, such as collaboration between groups or transfer of data ownership. In addition to managing de-identified human and non-human subjects, the system also securely tracks and manages PHI/PII information from research subjects.

Joject Biomatorials C		unications C	icnotypes At	tachments				Less than 1 time in	Less than half the	half the	Mare than half	Almost Alvzys	Your
	Person: <u>Subie</u> ed Studies: Demo ary Study: Demo		2		Custom ID: T5001	Incomplete emptying Over the past meeth, here effest have you had a senantien of not mustying your bladder completely after you finish reinsning?	0	3	2	3	the time	5	1
	Family: Demo Barcode:				Attribute 2:	Frequency Over the past meeth, how often have yes had to sonate anni les than two hows effer yes finished sonation? Informittency Over the past meeth, how often have yes found you stopped and started again several times when yes menated?	•	1	2 2	3	4	5 5	2 0
System Generate		9068-0302-4	04E-8D08-CA	39 1EE 38 5 30									
Associated Contac	ts				Urgency Over the pair meets, have definesh have you found it to pergeness visuation? Weach Stream Over the pair meets, how offest have you had a weak unitary presam?	0	1	2	3 3	4	5 (5)	3 5	
Phenotypes Focus Phenotype		Confirmed	Start Date	End Date									
adenocarcinoma - esophagus / 9/1/2008 (Current) GHD, acute 4/15/2010 (Current) Panaroolis Cysta 2/21/2010 (Current)						Straining Over the past month, how other have you had to push or usals to begin uninating?	0	1	2	3	4	6	5
IRB Protocols	0011						None	1 time	2 times	3 times	4 times	5 times or more	Yea
Start: 12/1/2005 End:			ie(s) Status	Date	Signed Form(s)	Notcuria Over the past menth, how many times did yes typically get on to winate from the time yes want to had will the time	0	1	2	3	4	(5)	5
			not sig signed	3/12) 3/24)	2007	you get up in the meening?							21
			other	3/12	2006	Total IPSS score							

Standard Subject Data Form (Left). Study-specific Data Form (Right)

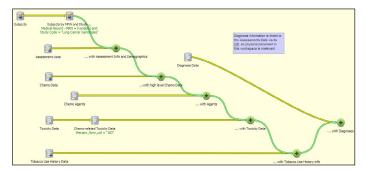
A critical feature of Labmatrix is its ability to manage the data & metadata associated with all stages and processes of a biospecimen's usage life-cycle: textual & image information on the biospecimen itself, its donor subjects/sources, aliquot & derivative lineage, storage locations, chains of custody, process workflow assignments, and experimental results are all captured in a hierarchical relational model - so users can quickly retrieve relevant data, historical information, and accountability logs pertaining to a given biosample. Operationally, Labmatrix is utilized to facilitate common practices in biospecimen management, such as batch create/update/aliguot functions, barcode labeling and storage assignments, inventory status tracking, current volume/mass, and physical transfer of custody.



Biospecimen Form (Left), Lineage (Top Right), and Storage Location (Bottom Right)

III. INFORMATION RETRIEVAL

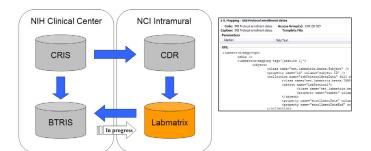
Traditionally, it has been difficult for investigators to query and obtain results for translational research because patient clinical information, biospecimen information, pathology information, molecular profile information, and other experimental data were kept in separate data repositories; even defining the correct linkage between data records from one repository to another can be very time-consuming (or even impossible) for research groups without dedicated IT support. With the deployment of Labmatrix, all types of data described above can be centrally federated/stored and accessed in a "one-stop shop" fashion. Utilizing the system's graphical guery user interface, investigators do not have to be trained in programmatic database languages in order to query across their data sets, hence receiving speedier data reports while freeing up IT resources. Query results can be exported in a variety of formats, while the queries themselves can be saved and shared. Furthermore, query parameters can be saved with static or dynamic values.



Graphical Query Interface

IV. PROGRAMMABILITY & MAINTENANCE

Labmatrix has a default web-browser user interface that presents different domain data in separate logical groups for ease of user consumption. This interface has been honed over multiple years of feedback from translational researchers. There are also alternative interfaces to access Labmatrix resources that adhere to the same system security rules. For example, Labmatrix's Java API allows for programmatic access, so that designated IT personnel can develop and maintain custom projects, features, and user interfaces. Furthermore, with the "user scripting" and "extract-transform-load" (ETL) tools, advanced data workflows and complex data imports are possible. Exchange of information with other systems is supported by multiple connectivity options, working examples include NCI's Clinical Data Repository and NIHCC's BTRIS.



Information Systems Data Flow (Left). Example ETL definitions interface (Right)

CONCLUSION

Labmatrix supports best practices in security, accountability, standardization, collaboration, efficiency, and data consumption for translational research. The implementation of CCR's Labmatrix system provides a centralized, secure, integrated, and flexible informatics platform. NCI investigators utilize Labmatrix in their daily routines to store, review, manage, share, and explore clinical research data.



Copyright 2012 BioFortis Inc. All Rights Reserved. Labmatrix and Qiagram are registered trademarks of BioFortis Inc. All other trademarks are the property of their respective owners. Specifications, terms and pricing subject to change.

General Information: info@biofortis.comMailing Address - Corporate Headquarters: BioFortis, Inc.Tel: +1.443.276.2464User Support: support@biofortis.com10320 Little Patuxent Pkwy, Suite 410, Columbia, MD 21044Web: www.biofortis.com