BIOFORTIS
a Q² Solutions Company

Turning Hope
Into Help™

A collaborative translational informatics environment at the NCI

Jason Levine,[1] John Otridge,[2] Steve Chen[2]
[1]Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD
[2]BioFortis 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 National Institutes of Health (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 and derivatives data, and experimental research data, 2.) communicates clinical and research information with the NCI Clinical Data Registry and the Biomedical Translational Research Information System (BTRIS), 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.

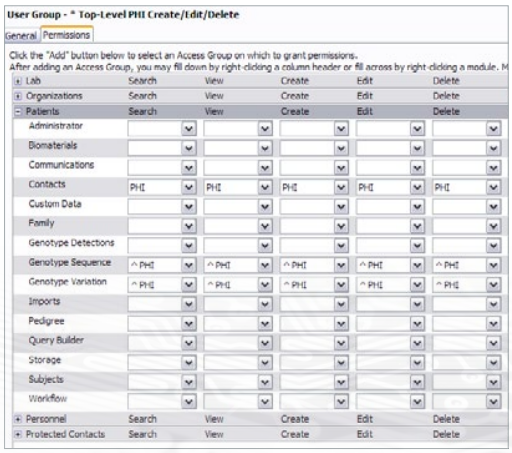
Background

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:

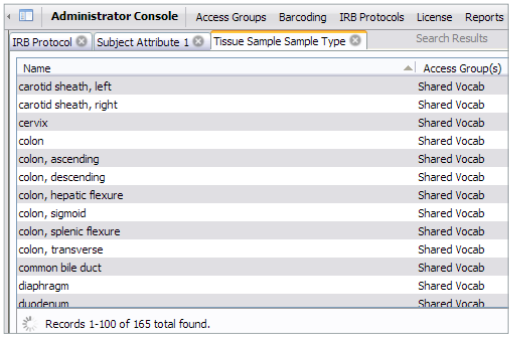
1 Access and security

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 Institutional Review Board (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 protected health information (PHI), personal identifiable information (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 and security forensics.

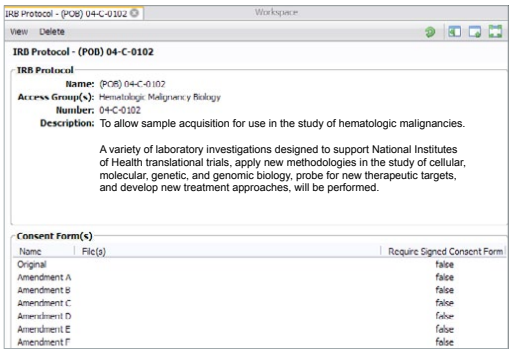
User group permissions management



Shared vocabularies management

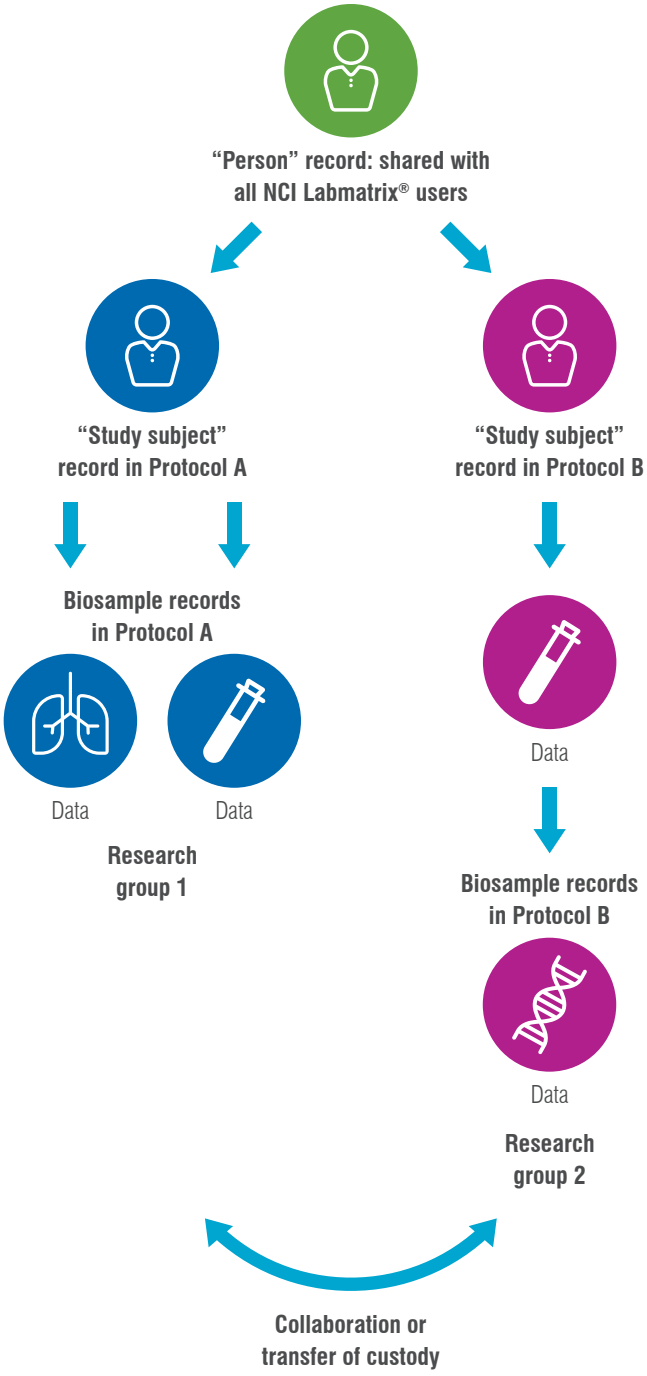


IRB protocols management



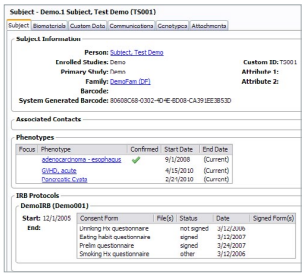
2 Standardizations and workflows

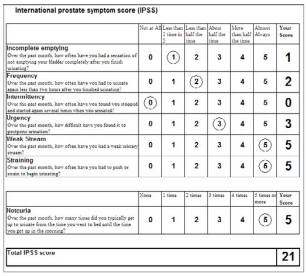
Several key goals for the CCR Labmatrix® system are to establish an informatics environment that facilitates data standardization, reduces data replication and divergence, encourages intramural collaborative participation, and prioritizes re-use of existing informatics components. Commonly used data 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.

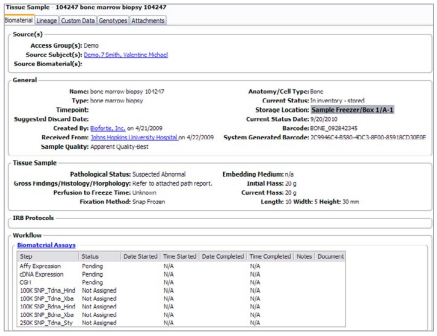
Standard subject data form (left) study-specific data form (right)

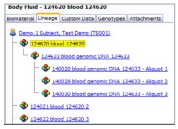


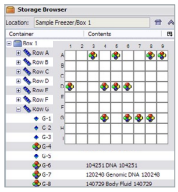


A critical feature of Labmatrix® is its ability to manage the data and metadata associated with all stages and processes of a biospecimen's usage life-cycle: textual and image information on the biospecimen itself, its donor subjects/sources, aliquot and derivative lineage, storage locations, chains 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/aliquot 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)



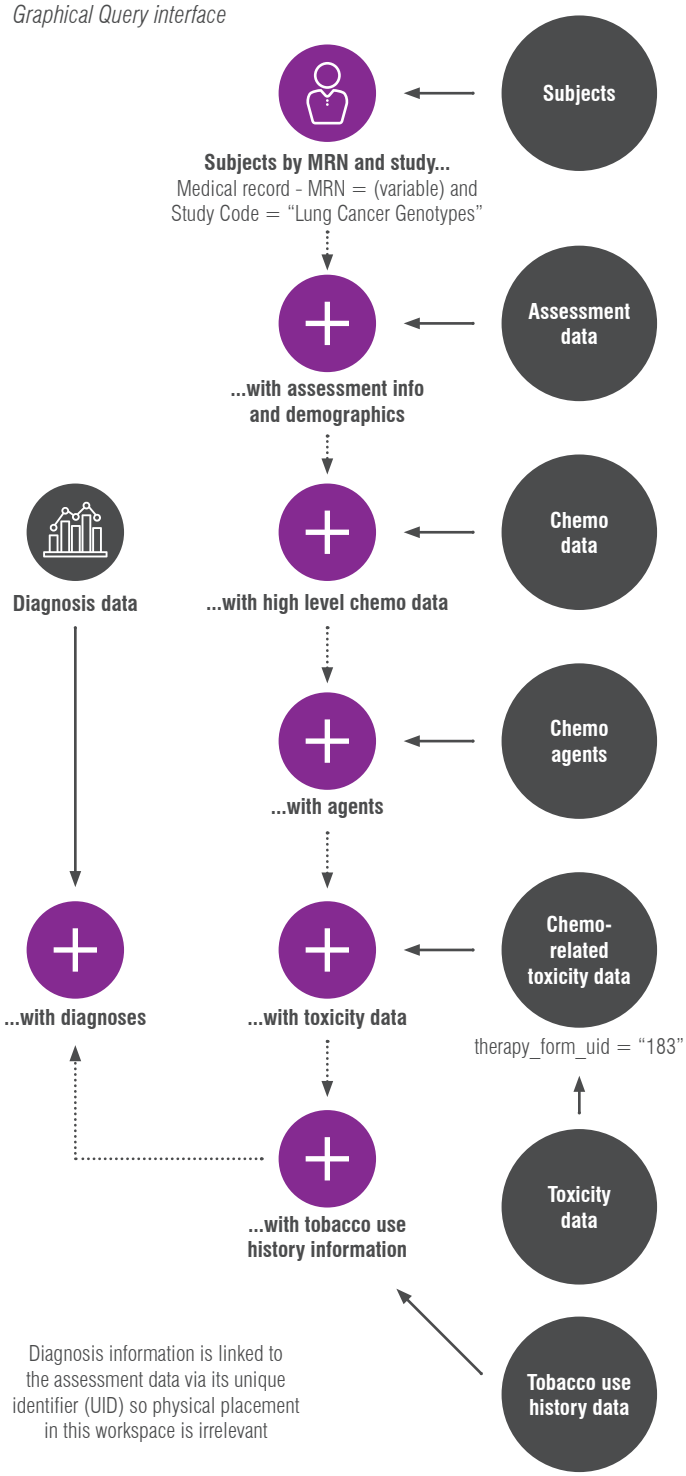




3 Information retrieval

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 query 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.

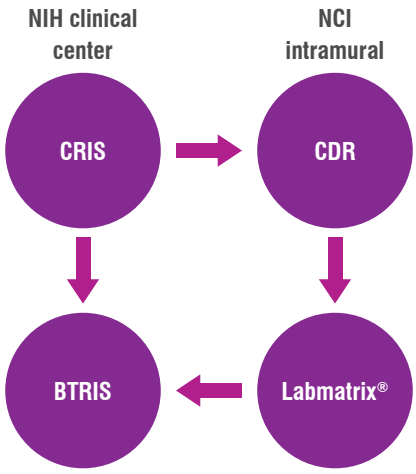
Graphical Query interface

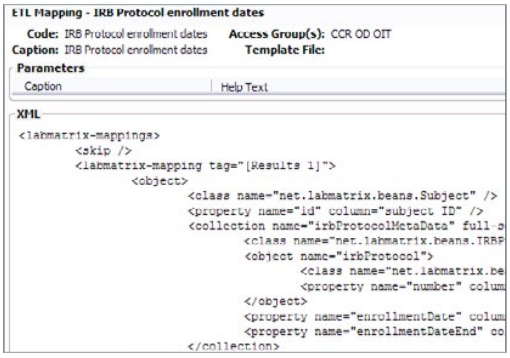


4 Programability and maintenance

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 application programming interface (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 (top). Example ETL definitions interface (bottom)





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.

If you have questions or comments, we can be contacted at info@biofortis.com

Website: q2labsolutions.com/biofortis
Inquiries: info@q2labsolutions.com
Tel: +1 (443) 276-2464
Mail: BioFortis, 10320 Little Patuxent Pkwy, Suite 410, Columbia, MD 21044, USA

BIOFORTIS
a Q² Solutions Company