C B M I
Informatics to facilitate discovery and translation

BIOBYTES

Need Medical Record?

CIDER – A Comprehensive, HIPAA compliant Patient Data Warehouse

Any investigator who has tried to gather data for research using manual chart reviews knows how daunting it can be. Data can be queried only one patient at a time and there is no easy way to identify a cohort of patients that meet specific criteria. Furthermore, chart reviews pose significant privacy risks to patients. CBMI has come up with a solution to this problem — a software application called "**CIDER**" short for Clinical Investigation Data Exploration Repository. **CIDER** has a built in query interface that allows access to a HIPAA compliant, patient data repository containing comprehensive electronic medical records from both inpatient and outpatient encounters at Washington University and Barnes Jewish practice sites.

What data does CIDER have and what can it do?

CIDER contains almost 5 million patients, 24 million encounters, 70 million lab results, 17 million medication orders, and 22 million vitals.

CIDER database contains discrete data such as patient demographics, laboratory test/results, medication orders, billing diagnoses and procedures, allergies, and vitals. In addition, text documents (e.g., radiology reports, surgical pathology reports, and admit/discharge summaries etc.) are available.

Users can directly query for discrete variables while text documents may be searched using a 'Google-like' interface.

CIDER application has a built-in vocabulary interface which provides its users with the flexibility to search clinical data using multiple standard vocabularies such as SNOMED CT

(Systematized Nomenclature of Medicine--Clinical Terms), LOINC (Logical Observation Identifiers Names and Codes), ICD-9, CPT4, etc.

CIDER

- ➢ **HIPAA compliant**
- ➢ **IRB approved**
- ➢ **Standard Vocabularies**
- ➢ **Easy access via web application**
- ➢ **Discrete and non-discrete data**
- ➢ **Advanced query options**
- ➢ **Help Desk Support**

Sign up for training

The first public release of **CIDER** was completed on March 1, 2010. CBMI offers **CIDER** training at no charge twice within a month. If you would like to start using **CIDER**, please sign-up for training here.

For the training calendar go to the CBMI Home page http://cbmi.wustl.edu/ and click on the *CIDER training details* link in the left hand panel.

Other useful **CIDER** resources like Frequently Asked Question and user manual are also available. (for FAQ click here, for user manual click here).

Need a CIDER contact person?

For more information please contact Bijoy George at bgeorge@path.wustl.edu. For any questions or issues with **CIDER**, please contact CBMI Support Services at 362–8853 or help@bmi.wustl.edu

Got Biospecimens?

caTissue Suite – A Biorepository Management Tool

Medical Researchers need to have access to high quality, well annotated biospecimens to validate their research and correlate their findings with the associated pathology and clinical annotations. **caTissue Suite**, a caBIG[TM] tool, is a one stop shop for biospecimen inventory, tracking and annotation. It is a web based application that permits users to track the collection, storage, quality assurance, and distribution of specimens as well as the derivation and aliquoting of new specimens from existing ones (e.g. for DNA analysis). **caTissue Suite** is suitable for biospecimen resource facilities and is independent of the type of biospecimen involved in the transaction or the nature of those transactions.

A high performance, web-based application, **caTissue Suite** offers high throughput data entry from any networked computer, search functionality to find and request specimens, and capability to capture annotation data related to the participant and the biospecimens. Biospecimen and study participant annotation includes informed consent status, clinical data (e.g., treatment, follow-up events), pathology data (based on CAP organ site checklists), and textual pathology reports (using caTIES-like functionality). Currently, we have nine groups at Washington University that are actively using **caTissue** to manage their biospecimen inventories.
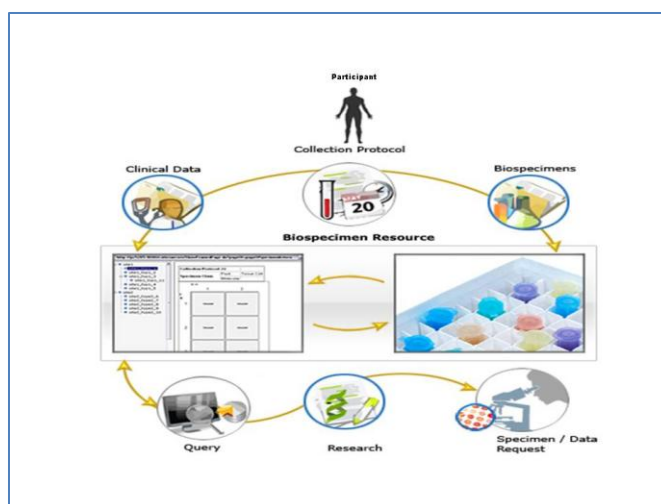
Managing Biospecimens in caTissue

Biospecimen data in **caTissue** is organized within a Collection Protocol template that allows representation of pre-defined specimen processing schemes and multiple study arms at multiple time points for high throughput data entry.

Users can enter data for multiple specimen types over multiple collection time points for a single study participant.

caTissue allows controlled access to different data types and houses separate functions for administrators, biorepository staff, and research investigators.

Need more information?

The Knowledge Center is an NCI-supported entity led by the Siteman Cancer Center, Washington University at St. Louis. It provides a centralized, authoritative repository of knowledge, information and web-based support to facilitate the deployment, ongoing development standards, and infrastructure in the tissue/biospecimen management domain. For more details please click here.

For informational videos and links to the caTissue Public Demo site please click here.

caTissue Training

caTissue offers regular "New User Demonstrations" and "Hands On" training sessions. Our monthly demonstration session is targeted at researchers who have not used **caTissue** and would like a high level overview of the tool and its functionality. Hands On training is appropriate for new and current users of **caTissue**. To sign up for training please click here.

The training calendar can be found here. An overview of **caTissue** is available here.

Need a caTissue contact person?

For more information please contact Dave Mulvihill dmulvihill@path.wustl.edu

For any questions or issues with **caTissue**, please call CBMI Support Services at 362–8853 or help@bmi.wustl.edu.

Got Clinical Data?

ClinPortal – A Clinical Studies Data Management Tool

ClinPortal is a secure, web-based data management system that allows researchers to store clinical study data in a fully customizable, online database. **ClinPortal** is web-based and authorized users can enter data directly into their electronic study forms, eliminating the need for paper case report forms. In addition to formal Clinical Studies, **ClinPortal** may also be used to manage clinical research data derived from disease oriented groups or disease-specific patient cohorts. PHI data is stored in strict compliance with HIPAA and IRB regulations.

A built-in query tool allows **ClinPortal** users to access data among **ClinPortal** studies and across companion CBMI applications such as **caTissue** (biospecimen repository); rendering search results in identified or de-identified fashion based on the authorization of the user.

Since October 2008, 29 research teams have begun using **ClinPortal** to collect, manage, query and store their data. We are currently assisting researchers in managing 29 studies, and seven new studies are planned for release this month.

How to start using ClinPortal?

There are two pathways for a research team to consider when developing their study in **ClinPortal**; Self Service or Full Service.

Self Service

ClinPortal Self-Service study development allows the research team to develop their custom forms and study calendar using the Clinical Study Designer (CSD) tool. The CSD tool is a *free* and easy to use web application offered to Washington University researchers for designing, editing, re-using, and sharing case report forms through an online portal. A study built with CSD can use **ClinPortal** to manage their data collection with no charge to the researcher.

Full Service

With **ClinPortal** Full Service study development, an Adoption Manager will (a) guide the research team through the steps in accordance with best practices in form design and development, (b) build the study calendar, perform quality assurance tests on the forms and (c) create saved and parameterized queries for the users.

Where to get training for ClinPortal?

ClinPortal offers three types of training opportunities (see below for link to schedule).

1) **ClinPortal** "open office" sessions occur approximately three times per month, and are appropriate for nonusers as well as current **ClinPortal** users. Nonusers or potential users can request a high level overview of the application and its functionality, and consult with a **ClinPortal** representative about their data management needs. Current users can consult with CBMI support staff regarding the inclusion of additional studies, the creation of data queries or other data extraction activities, or discuss other **ClinPortal**-related activities.

ClinPortal

- ➢ **HIPAA compliant Clinical Study Data Management**
- ➢ **Web based data entry**
- ➢ **Electronic case report forms**
- ➢ **Query functionality**
- ➢ **Standard vocabulary**
- ➢ **Full integration with caTissue**
- ➢ **Training sessions available**
- ➢ **Help Desk Support**

---

2) Hands-on **ClinPortal** training is a two-hour session with the first hour dedicated to data entry use, and the second, with instructions for users on how to extract data out of **ClinPortal** using our "data by form" query, canned queries and parameterized queries. Advance query training is available by request; call CBMI Support Services at 362–8853 or help@bmi.wustl.edu to schedule one.

3) Hands-on CSD training is a two-hour session covering the creation of electronic forms and study calendar design. Users will learn how to build their own forms, using various controls to facilitate good data entry, and will be able to build their own study following completion of the training.

ClinPortal Resources

Please click here for a training schedule. To sign up for training click here. An overall review of then application is available here.

Need a ClinPortal contact person?

For more information please contact Dianne Oliver doliver@path.wustl.edu.

For any questions or issues with **ClinPortal**, please call CBMI Support Services at 362–8853 or help@bmi.wustl.edu.

Collaborator's Corner

Interview with Virginia Buckles, Executive Director of the Alzheimer's Disease Research Center at Washington University School of Medicine and Mary Coats, Research Associate Professor in the Department of Neurology, Washington University School of Medicine.

By

Rekha Meyer

Q. Could you elaborate on the background and mission of the Alzheimer's Disease Research Center?

A. The Knight Alzheimer's Disease Research Center (ADRC) at Washington University is about 26 years old. It is one of 29 federally designated Alzheimer's Disease Centers that promote innovative research in Alzheimer's disease (AD). Our first, large program project was funded in 1984 and has been continuously renewed since that initial funding. Our Center grant has been funded since 1985. So those two grants together have funded efforts to determine the earliest possible signs to detect or diagnose the disease. The emphasis has been to distinguish between what is normal and what is not normal in ageing. The ability to detect the disease clinically before someone dies is very good, about 93% accurate. So when we say someone has Alzheimer's, if they come to autopsy, they have it. More recently, the emphasis has been not so much on diagnosing and detecting but potentially predicting who will develop the disease with the use of biomarkers.

Q. So the overall goal is to find a way to prevent it?

A. Absolutely, that's where we want to go. Once the brain damage is established, when people are symptomatic, it is looking like effective therapeutic options are not very good. We hypothesize that treatment should be targeted at the very beginning when the damage in the neuron starts, or, at the point of preventing that damage from ever starting, that's where the most hope for therapy may be.

Q. How common is Alzheimer's?

The majority of AD patients are over age 65 but it does occur in younger people. Five to ten percent of people over the age of 65 have AD, this figure approaches 40-50% after age 85. That's pretty big if you think about it. People over 85 and older are the fastest growing segment of our population and over 50% of them will develop Alzheimer's.

Q. What are the risk factors for AD?

A. It is not known what starts the disease process, our current understanding is that the cause of AD is likely to include genetic, environmental, and lifestyle factors. It's like cardiovascular disease in that there is not a single cause; it is a complex set of interactions of a number of things.

There is a rare form of AD that is caused by a gene mutation that causes people in their 20's and 30's to develop Alzheimer's. One of our studies, Dominantly Inherited Alzheimer Network (DIAN), dedicated to understanding this rare form of the disease, is an international research partnership (http://www.dian-info.org) funded by a research grant from the National Institute on Aging.

There is another genetic risk factor for AD which is not a mutation and is called Apolipoprotein E (APOE). This is a type of lipoprotein, for which there are 3 alleles - one allele (ε4) increases your risk of developing AD and also increases your risk of developing it earlier. Another allele (ε2) actually seems to protect against AD and then the most common allele neither imparts risk nor protects against AD, it's a mixed bag.

Q. How do you collect your research data at the Center?

The Alzheimer's Disease Research Center has a broad scope that includes the study of non-Alzheimer's dementias. Due to the complexity of the disease we approach it from many different directions. Our clinical core enrolls research participants for an annual, two hour clinical interview and assessment (including an informant portion). The informant is a study partner who we call a collateral source, somebody who knows the participant well. Any changes experienced by a participant in a variety of areas are reported by the informant to the interviewer. There are six cognitive domains that participants are assessed on; for example, judgment, personal care, memory. This interview is followed by a paper and pencil or computer aided, memory, attention, psychometric testing. That's our basic battery. We draw blood at that time for the Genetics core where their APOE is determined.

If possible, we follow our participants until they die; then we seek autopsy permission to examine their brain in our Neuropathology Core to make a neuropathologic diagnosis of Alzheimer's disease and to collect quantitative data. The tissue is then banked for sharing with other interested researchers. Because our study is longitudinal with participants returning every year, it is possible to observe the first signs of cognitive decline in people who were cognitively normal initially. Then we try to figure out how to detect that decline earlier. A cognitively normal individual who will develop Alzheimer's disease at a future date has been the focus of our Biomarker program.

Q. I understand the Center is in the process of adopting ClinPortal, a Clinical Data Management System to manage your data. Before ClinPortal, how were you managing your data?

A. The Biostatistics and Data Management Core which is housed in the Division of Biostatistics is the current data repository for our data. They created our data entry screens, and currently oversee data acquisition and data entry. The physician or clinician who is doing the interview records 60 pages of responses which are handed to a data entry person who then enters it once and then enters it again a month later as an error catching, error correction mechanism. Once the data is deemed "clean" it is finalized and is then available for analysis and sharing.

Q. How will ClinPortal change the way you collect data?

A. ClinPortal will be our primary storage and data acquisition system. What we envision is that the data will no longer be put on paper and entered by hand. It will be entered live by the clinicians as they do the interview. The data will be saved and stored in ClinPortal, after which it should be available to qualified people to approve and view. Biostatistics will still be involved in managing, checking, cleaning and will also make available de-identified data sets for sharing

with people outside the University and to the National Alzheimer's Coordinating Center (data repository for all 29 Alzheimer's Disease Centers).

Q. caTissue is a CBMI application for managing biospecimens. Could you talk about how the Center has been using caTissue?

The Neuropath, Genetics and Biomarker cores are using caTissue to handle their biospecimens. The CBMI team worked with those cores first to get caTissue up and running and then they brought in the Clinical core. Participants, who come for clinical evaluation, also provide tissue. The Center takes advantage of caTissue in two ways: First, with tissue tracking, which is wonderful and probably the main point of using caTissue to manage and track tissue samples and second, with entering and storing data like values and scores. For example, we measure Aβ values in cerebrospinal fluid which would then be directly available for analyses.

Q. How will ClinPortal be useful for managing your data?

A. ClinPortal is going to capture the data; it will also be the storage, perusal, sharing, data looking system. What we foresee is that a combination of ClinPortal and caTissue will make our data more accessible to people who have questions. For example, questions like "How many apoE4 positive, left-handed males do we have?" Right now we have one or two people who know the data well enough to access it; our data manager has been here for ~25 years and provides invaluable information about the nuances of the data. As our Center grows, producing datasets through one or two individuals becomes a rate-limiting condition. We're hoping that this bottleneck will be diminished so investigators can go directly and safely and securely to ask questions. Instead of one or two people trying to answer all these data requests for a number of projects, we can use ClinPortal to make the data more accessible and therefore more productive. We feel that if this new system (ClinPortal) can handle a complex, longitudinal, natural-history, biomarker, behavioral, clinical, tissue, imaging, messy study like the ADRC, it can probably handle anything an academic medical center can produce.

Q. How do you anticipate using data query in ClinPortal?

A. I think we all will query the data. Once the caTissue data is merged with clinical data, we can ask questions like "How many participants with apoE4 or apoE2 have a CSF value lower than a certain criteria?" Furthermore, we will use the caTissue-ClinPortal integrated systems to generate reports for review by the external advisory committees for each of our big grants. We have to generate data on questions like, "How many participants came to autopsy versus those who died and did not have autopsy". We will be using the data for obtaining this kind of information.

Did you know?

Other specialized services offered by CBMI

➔ Genome annotation and promoter analysis tools
➔ Centralized storage, management and access of Affymetrix microarray expression and SNP genotyping data
➔ Access to microarray analysis applications: Function Express, Spotfire
➔ Consultation on microarray expression, SNP analysis, Copy number and LOH analysis
➔ Centralized storage and management of medical re-sequencing or mutation profiling data

➔ Automated analysis of medical re-sequencing or mutation profiling data

For Questions and Analysis Support

If you need consultation for microarray analysis, copy number analysis or any other CBMI applications, please call or email CBMI Support Services at help@bmi.wustl.edu or call 362.8853.

Featured CBMI staff

Bijoy George, MBA, PMP

Our featured staff in this issue is CBMI Program Manager, Bijoy George. George has been with CBMI for two years. He has a total of nineteen years of experience in information technology, with the last twelve years in Project/Program Management. He is a PMI (www.pmi.org) certified project manager. He has a Bachelor's degree in Aeronautical Engineering and has a Masters in Business Administration (MBA) from Washington University, St. Louis. George is currently responsible for overseeing and managing **CIDER** and **caTissue**. George is a resident of St. Louis since 1994 and likes to play tennis and golf in his leisure time.

In Upcoming issues

VIVO, a networking application for scientific researchers.