Translational Research Enhancement Core

(Patient Registry, Clinical Information Systems, and Tissue Repository)

Jim Potter –Core PI/PD

Hopkins Conte Digestive Diseases Basic and Translational Research Core Center

Mark Donowitz, MD – Center PI

The Division of Gastroenterology and Hepatology
The Johns Hopkins University School of Medicine

Translational Research Enhancement Core

SOM Informatics Support

It is impossible to thank everyone who has contributed to the development and testing of informatics for clinical research. I have highlighted only a few in the following slide, but we would not be successful without the efforts and contributions of many others. I am in awe at the dedication of effort and intellect by so many to enable clinical research.

Translational Research Enhancement Core

SOM Informatics Support

Dr Dan Ford - The Vice Dean for Clinical Research
Diana Gumas - IT Director, SOM Welch Health Sci Informatics
Joseph DiMaggio - IT Project Manager caTissue/CRMS
Dorothy Damron - Dir Clinical Research Admin CRMS
Chris Shafer - Dir of Center for Clinical Data Analysis
Dave Thiemann - CCDA
Sam Meiselman – SQL DBA, CCDA
Matthew Marcetich – TREC caTissue superuser

- CISSCI Committee (Committee on Information Systems to Support Clinical Investigation)
- Research Data Sub Council (**Stuart Ray, Jennifer Kulynych**; to enable high quality, innovative data research, while establishing a uniform process and standards to protect security and privacy)
- Biospecimen Committee
- CRMS SOM Development Committee:
- JHMCIS, Integration Team – Medical Informatics
- Center for Clinical Data Analysis - I2b2/SHRINE Pilot Group

Change Management

Anticipate and Embrace Change

"The only change people don't fear is the change that jingles in their pockets"

Informatics in Service of Clinical Research

Ask “WHAT”, and then “HOW”

“Vision Drives Change”

"To wrest from nature the secrets which have perplexed philosophers in all ages, to track to their sources the causes of disease, to correlate the vast stores of knowledge, that they are quickly available for the prevention and cure of disease – these are our ambitions."

- Sir William Osler, 1906

Hopkins NIH/NIDDK Conte Digestive Diseases Basic & Translational Research Core Center

Mission

The goal of this Center is to advance basic and translational GI research at Hopkins and to encourage further interactions among the members of our research base.

Mark Donowitz, MD
Director

Ann Hubbard, PhD
Associate Director

Steve Leach, MD, PhD
Associate Director

Bob Cole, PhD
Director
Proteomics Core

Olga Kovbasnjuk, PhD
Ann Hubbard, PhD
Associate Directors
Imaging Core

Nick Zachos, PhD
Director
Mouse Physiology Core

Jim Potter
Director
Translational Research Enhancement Core

The Hopkins Digestive Diseases Basic and Translational Research Core Center is a NIDDK Silvio Conte Core Center which has as its goal to advance basic science and translational digestive diseases research at the Johns Hopkins University School of Medicine

The major areas being studied include:

1. Epithelial Transporter Function and Regulation, including trafficking, protein-protein interactions and bioinformatics.
2. Inflammation, fibrosis and cancer development, including infectious diarrheal diseases, inflammatory bowel diseases, fatty liver, alcoholic and non-alcoholic steatohepatitis, obesity, and pancreatitis.
3. Neurogastroenterology
4. Pancreatic Development Biology.
5. Epigenetic Aspects of GI Cancer.

Cores

Mouse Physiology Core

The goal of the Mouse Physiology Core is to advance the understanding of digestive diseases by providing services to increase the ease and efficiency of studying the physiology and pathophysiology of mouse models, including those knockout and transgenic models in which colitis and small intestinal disease occur.

Imaging Core

The goal of the Fluorescence Imaging Core is to provide state-of-art light microscopy technology to the members of the Hopkins Basic Research Digestive Disease Development Center and to the whole Hopkins scientific community.

Proteomics Core

The Proteomics Core uses mass spectrometry coupled to one (1D) and two (2D) dimensional separations by column chromatography or gel electrophoresis to identify, quantify or characterize proteins and their post-translational modifications, that are expressed in well characterized protein fractions from the small intestine, colon, kidney, liver and pancreas.

Translational Research Enhancement Core

The overarching goal of the Translational Research Enhancement Core (TREC) is to provide the expertise and infrastructure for cost effective tissue and clinical data collection that meets Best Practices through the establishment of uniform Standard Operating Procedures (SOPs) to promote scientific advances by, and scientific interactions among Core users, collaborative investigators and young investigators, without each investigator having to invest in the equipment/resources or to individually work out the specific SOPs.

Translational Research Enhancement Core (TREC)

Goal of the TREC:

- To provide a strong scientific foundation for clinical and translational research
- To promote scientific advances by, and scientific interactions among Core users, collaborative investigators and young investigators.
- To provide the expertise and infrastructure for cost effective tissue and clinical data collection
- To follow Best Practices through the establishment of uniform Standard Operating Procedures (SOPs)

Advantages of the TREC

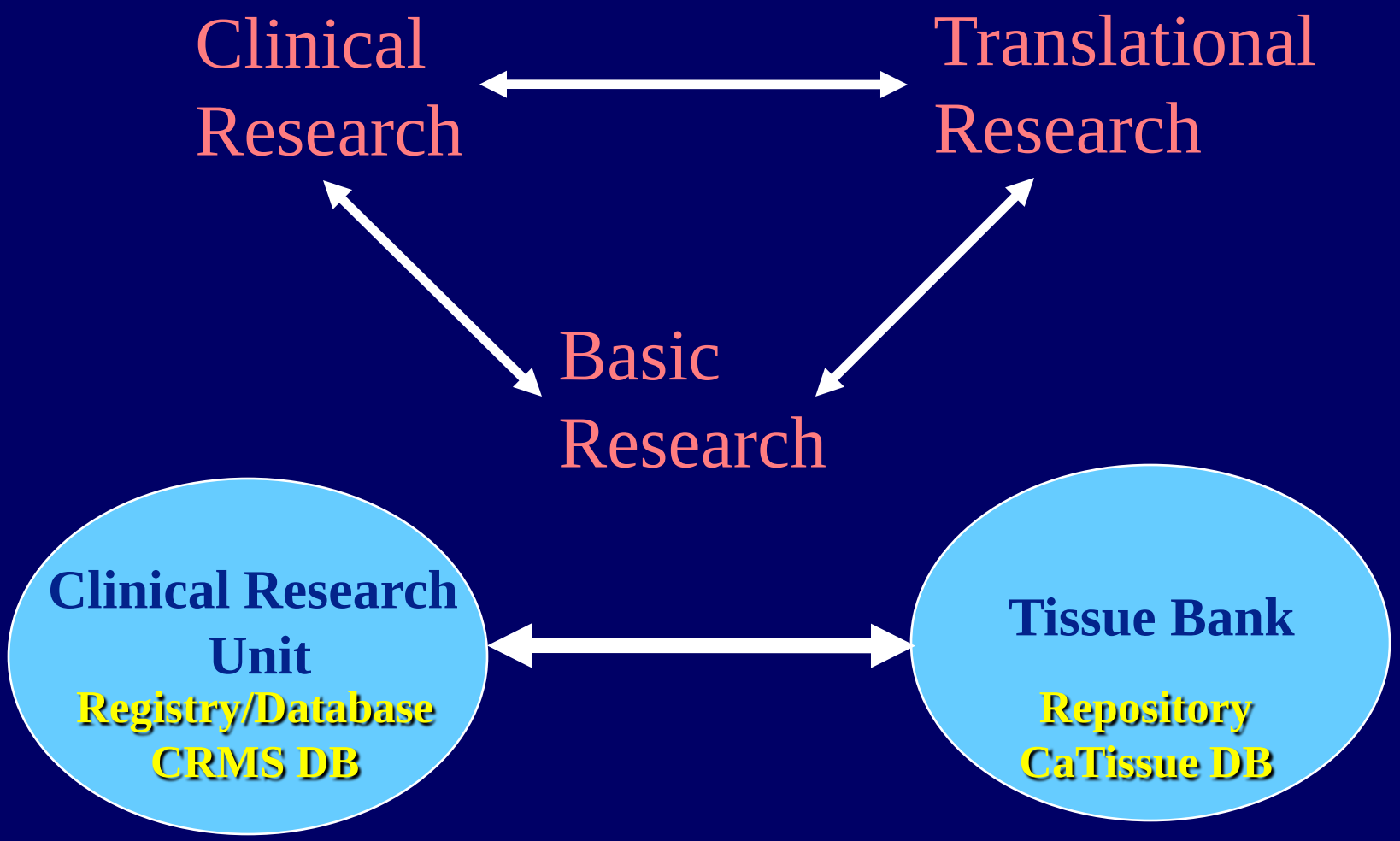
- Harvest of Biospecimens with associated clinical data thus allowing specialized study of disease processes
- Data Security and Integrity
- Facilitate Hypothesis Driven Research
- Facilitate cohort identification, IRB approval, Patient Screening / Registration / Consent
- Best Practices/SOP

Specimen Collection:
Factors to be considered:

- Purpose of the Biorepository
- Best Practices / SOPs
  - The Scientific Foundation for the Research
- Consent
- Governance / Oversight
- Funding Support
- Hardware
- Software/Audit/Security
- Enterprise Bioinformatics Systems
- Personnel (FTE)

Translational Research Enhancement Core

GI Division Sponsored Units



Translational Research Enhancement Core

Provides infrastructure for specimen (GI-TxB) and clinical data collection (CRU) following established Best Practices and Standard Operating Procedures as defined by an Oversight Committee

CRU Roles:

- patient identification
- patient recruitment
- facilitation of patient trials
- patient registry / cohort discovery

Tissue Bank Roles:

- Collection and transport
- Annotation and DB entry
- Processing
- Storage
- Management
- Quality assurance
- Distribution (shipment)
- Ethical use
- Oversight

Governance - TREC Oversight Committee

Division Director

Anthony Kalloo

TREC Board

TREC PI/PD-Jim Potter, Center PI – Mark Donowitz,

Investigators

- One PI from each of "8 study sections"
- At least two rotating **independent non-GI**, JHU investigators

External Scientific Advisory Committee

Governance/Oversight - Roles

- Provide ethical oversight of all research involving repository specimens
- Assure the best possible biospecimen quality
  - ✓ BP are followed
  - ✓SOP Manual reviewed, updated and followed
  - ✓External Scientific Review

Governance/Oversight - Roles

- Provide consistent documentation
- Establish guidelines for collection, processing, storage and retrieval of specimens
- Assessment of new technologies and the incorporation of new technologies into the repository
- Assure the fair and impartial utilization of specimen resources

Translational Research Enhancement Core

Highlights – First Year

- 21, 000 specimens by 7-30-12
- Completed implementation of Enterprise Data Systems
- Organoid Sample Bank
- 4 significant grant proposals by young investigators including R21 and ADA
- Collaborations
- Directed collections for pancreas, liver and lumen
- $2 Million in additional support
- 20 publications in first year

Translational Research Enhancement Core Enterprise Systems

What is in Place?

- CRMS
- EPR 2020 (Amalga) –i2b2 Pilot /SHRINE,
  -Center for Clinical Data Analysis
- EPIC
  - ✓ Clinical Research Team
  - ✓ Consents
- Surgical Path
- Clinical Registries Workshop
- CaTissue
  - ✓ Consultant-Trainer
  - ✓ Coordinator-Programmer

Translational Research Enhancement Core Enterprise Systems

Clinical Research Management System

Hypothesis → eIRB → Screening → Scheduling → Billing → Formbuilder Clinical Data → Analysis

Epic (EPR2020)
I2b2 - pilot
ICTR CCDA

→ Clinical Registry Development

Consents

Surgical Path

↓

caTissue – Tissue Repository Software

(Support for multiple repositories within a single caTissue instance)

(Clinical Annotation with Specimens)

caTissue Suite:

- Core Functions / Modules
  - Administration: Create and edit users, protocols, and storage systems associated with a biospecimen inventory. Manage multiple repositories under one roof.
  - Biospecimen: Create and edit data concerning participants and their corresponding biospecimens.
  - Query: Identify biospecimens and their data based upon one or more selection criteria.

Key Functions and Benefits of caTissue Suite

- **Search**

  Allows biospecimen resource staff and scientists alike to search for biospecimens based on a combination of properties

- **Biospecimen Tracking**

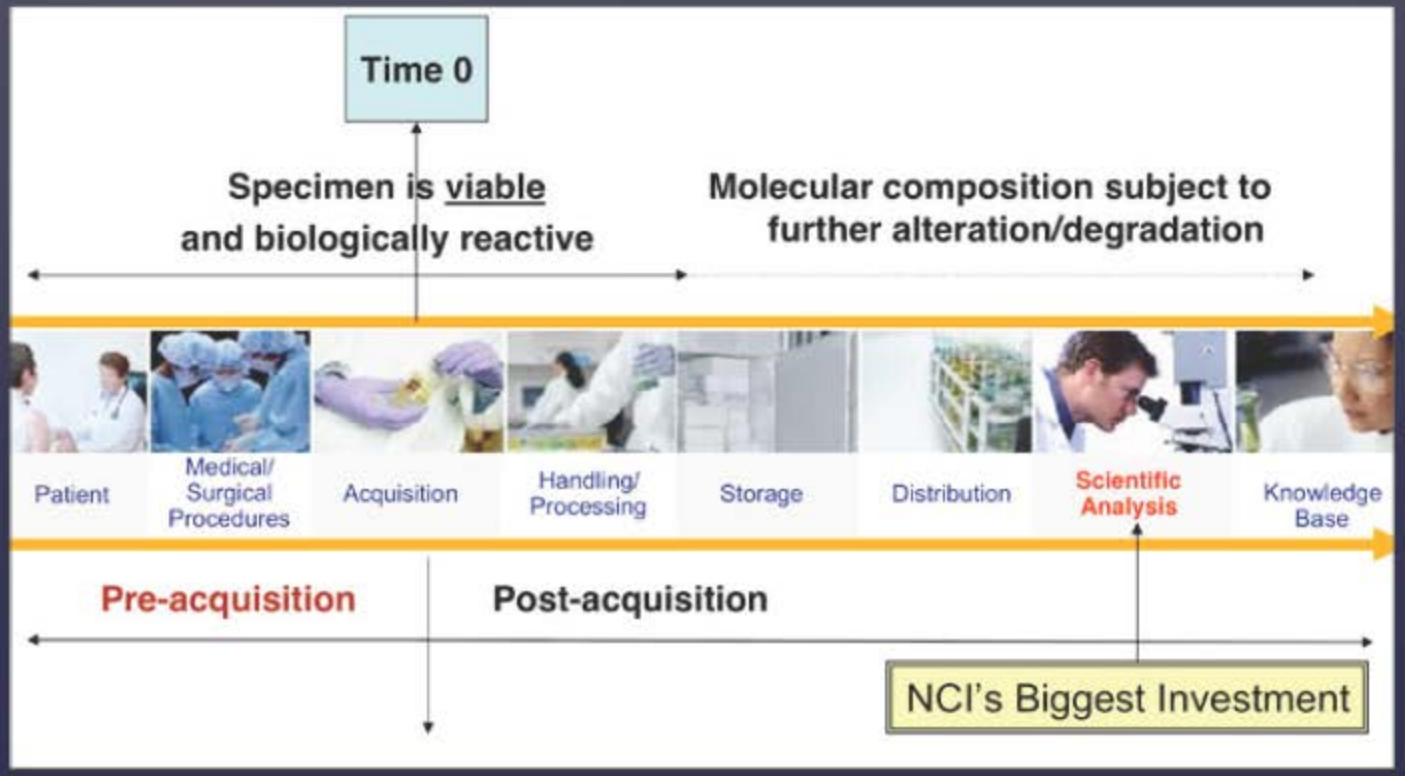
  Allows biospecimen resource staff to record events related to biospecimen processing and quality assurance and to annotate biospecimens with pathological, clinical, and custom data

- **Biospecimen Inventory Management**

  Allows biospecimen resource staff to record data about biospecimen collection, storage, requests, and distribution.

Most Biospecimen Research

- Focus on Analysis not collection or processing



Today's Research & Path Non- CLIA Lab

- No Standardized SOP's for collection, processing - few SOP's for storage & analysis of biospecimens

Collection (variables not validated)

Processing (variables not validated)

Storage

Analysis

=

Improve Pt. Outcomes

caTissue - Benefits

For University

- Centralized application
- HIPAA compliance easy to monitor
- Data Security and Integrity
- Multiple Repositories within one caTissue instance
- For PIs:
- No need for custom databases
- Easier data sharing
- Improved data organization
- Study Team can enter participants
- Powerful Query Capability

caTissue - Needs

- Ongoing Development of the Existing System
  - Krishagni offshore support
  - Users Group
- Integration with CRMS and Other Systems

TREC Informatics Needs

1. An enterprise informatics solution for multicenter consortiums- The BETRNet Virtual Biorepository DB

BETRNet –Barrett's Esophagus Translational Research Network

U Penn	Case Western	UNC
Columbia	Michigan	U of Washington
Wash U	Hopkins	NCI
Mayo Clinic	NIH	Fred Hutchinson CRC

2. Clinical Registry Development Capabilities / Cohort Discovery
3. Link to external datasets
    - BETRNet Biorepository Database
    - CAPS Study (Ca Pancreatic Screening) (Multinational
4. Tiered Consents, Consents linked to specimens

TREC and Gastroenterology Solutions

1. An enterprise informatics solution for multicenter consortiums- The BETRNet Virtual Biorepository DB

   caTissue

2. Clinical Registry Development Capabilities / Cohort Discovery

   i2b2 (Informatics from the Bench to the Bedside)

3. Link to external datasets (BETRNet Biorepository Database, CAPS Study)

   I2b2 SHRINE (Shared Health Research Information Network)

4. Tiered Consents, Consents linked to specimens

   EPIC - CRMS

Hopkins Conte Digestive Diseases Basic and Translational Research Core Center

	Barrett's Cryo-Preserved Specimens by Pathology Diagnosis and Event							
	Normal	Barrett's-No Dysplasia	Barrett's-Indeterminate	Barrett's Low Grade Dysplasia	Barrett's High Grade Dysplasia	Barrett's AdenoCA	Barrett's-Squamous-CA	Total
Baseline	754	286	30	53	44	10	5	1182
Followup-1	236	80	9	8	6	0	0	339
Followup-2	98	32	5	1	2	0	0	138
Followup-3	45	22	8	3	1	0	0	79
Followup-4	19	7	4	0	2	0	0	32
Followup-5	4	1	1	0	0	0	0	6
Total	1156	428	57	65	55	10	5	1776

	1.0ml Serum Aliquots of Barrett's Patients by Visit Diagnosis and Event							
	Normal	Barrett's-No Dysplasia	Barrett's-Indeterminate	Barrett's Low Grade Dysplasia	Barrett's High Grade Dysplasia	Barrett's AdenoCA	Barrett's-Squamous-CA	Total
Baseline	333	377	68	99	180	16	8	1081
Followup-1	121	100	11	30	28	0	0	290
Followup-2	44	43	5	14	11	0	0	117
Followup-3	35	21	8	11	11	0	0	86
Followup-4	9	16	0	3	8	0	0	36
Followup-5	4	7	0	4	0	0	0	15
Followup-6	0	4	0	0	0	0	0	4
Total	546	568	92	161	238	16	8	1629

Hopkins Conte Digestive Diseases Basic and Translational Research Core Center

Dr Meltzer's Repository Containing Matching Tissues, Serum, Plasma, and WBC Samples						
Barrett's No Dysplasia	Barrett's Low Grade Dysplasia	Barrett's High Grade Dysplasia	Barrett's AdenoCA	Total Barrett's Specimens	Non-Barrett's, Non AdenoCA Control	Total Specimens
850	200	90	250	1390	1100	2490

I2b2 Query of EPR 2020

		Pathology Report Findings													
		Adeno CA		BE-HGD		BE-IND		BE-LGD		BE-ND		NORMAL		SQ-CA	
Diag Group	Pt Count	Pts	Events	Pts	Events	Pts	Events	Pts	Events	Pts	Events	Pts	Events	Pts	Events
All Diagnoses	7752	2503	2802	722	1323	573	1070	618	1230	223	250	3515	3442	591	869
Barrett esophagus (530.85)	4558	1114	1575	512	1087	514	1004	514	1109	188	220	2580	2929	105	180
Barrett in EPR Problem List	1833	770	1201	237	514	139	387	181	443	83	92	876	1633	172	283
Malignant neoplasm of abdominal esophagus (150.2)	52	17	22	4	8	0	0	3	4	0	0	15	44	23	41
Malignant neoplasm of esophagus (150)	2863	1452	1990	419	812	120	299	238	532	56	79	949	1551	476	701
Malignant neoplasm of esophagus unspecified site (150.9)	2348	1247	1694	348	688	104	277	202	471	49	70	830	1418	372	557
Malignant neoplasm of lower third of esophagus (150.5)	782	535	1000	197	415	61	155	109	274	22	31	300	626	97	147
Malignant neoplasm of middle third of esophagus (150.4)	189	53	97	15	33	4	33	16	47	3	4	62	101	90	133
Malignant neoplasm of other specified part of esophagus (150.8)	719	425	803	134	283	36	123	76	205	15	19	310	541	147	242
Malignant neoplasm of thoracic esophagus (150.1)	76	32	67	17	34	3	3	12	15	0	0	30	36	30	46
Malignant neoplasm of upper third of esophagus (150.3)	146	64	129	18	29	3	12	12	27	2	2	56	98	54	98

Sam Meiselman, CCDA

I2b2 Query of EPR 2020

		CA Tissue Findings													
		Adeno Carc		BE-HGD		BE-IND		BE-LGD		BE-ND		CA-IM		CA	
Diag Group	Pt Count	Pts	Events	Pts	Events	Pts	Events	Pts	Events	Pts	Events	Pts	Events	Pts	Events
All Diagnoses	7752	3	4	33	41	27	44	45	50	165	180	1	1	5	5
Barrett esophagus (530.85)	4558	1	1	31	38	26	43	45	50	162	178	0	0	4	4
Barrett in EPR Problem List	1833	0	0	12	14	10	16	12	15	42	56	1	1	2	2
Malignant neoplasm of abdominal esophagus (150.2)	52	0	0	0	0	0	0	0	0	1	1	0	0	0	0
Malignant neoplasm of esophagus (150)	2863	3	4	20	22	7	9	16	17	40	46	1	1	5	5
Malignant neoplasm of esophagus unspecified site (150.9)	2348	3	4	18	20	7	9	16	17	38	44	1	1	5	5
Malignant neoplasm of lower third of esophagus (150.5)	782	2	3	9	10	3	3	4	4	23	26	1	1	4	4
Malignant neoplasm of middle third of esophagus (150.4)	189	0	0	0	0	0	0	0	0	1	1	0	0	0	0
Malignant neoplasm of other specified part of esophagus (150.8)	719	2	3	6	6	2	3	2	3	12	15	1	1	2	2
Malignant neoplasm of thoracic esophagus (150.1)	76	0	0	1	1	0	0	0	0	2	2	0	0	0	0
Malignant neoplasm of upper third of esophagus (150.3)	146	0	0	0	0	0	0	0	0	2	2	0	0	0	0

Sam Meiselman, CCDA

WD patient clinical data summary

Patients	Mutation 1st allele	Mutation 2nd allele	Age of onset	Ceruoplasmin mg/dL	Cu Serum ug/dL	Cu Urine ug/24 hr	KF ring (%)
Normal	none	none		**25-45**	70-140	0-50	-
Heterozygous	**H1069Q**	none		**31.8 +/- 6.7**	100.4 +/- 24	18.5	-
Homozygous	**H1069Q** median [IQR]	**H1069Q**	27.9 +/- 9.9	**13.9 [9.2]**	61.2 [34]	163 [255]	+ (28%)
M	**H1069Q**	**S653Y**	-	**9.2**	35	135	-
G	**H1069Q**	**S653Y**	22	**nd**	22	30	-
J	**H1069Q**	**S653Y**	21	**3.5**	27	1050	+
BP	**H1069Q**	**S653Y**	29	**46.4**	131	162	+
BP prior to LT	**H1069Q**	**S653Y**	32	**14**	250	3335	+

Lelita Braiterman

TREC and Gastroenterology Needs

3. Clinical Registry Development Capabilities / Cohort Discovery

--Self Service

The Division of Gastroenterology is pursuing clinical registries in all subspecialties of GI, beginning with already funded initiatives, including pharmacological testing with organoids, lineage/family studies with FAP and IBD, GWAS in IBD, HCC, Cholagio Ca, Pancreatic Ca, diarrheal dis., Gi Motility, Gastric Neuroendocrine disorders, and liver disease.

There is a great need to share data.

I2b2/SHRINE

Pilot Project at JHMI
(The Center for Clinical Data Analysis)

Scenario: Combined Demographics, Diagnosis and Time Constrained Lab Results Query

JH i2b2 Pilot – Phase II

Summary Task	Status	Data Set Size	Estimated Completion	Notes
Patient Mapping and Dimension	Done	2.5 Million	11/23/2012	Patients and visits based on EPR2020 identity backbone including multi-institution and multi-mrn references
Visit Mapping and Dimension	Done	42 Million	11/23/2012	
Provider Dimension	Done	35073	11/26/2012	contents of EPR/EPR2020 Provider dictionary (CDTProvider)
Demographics	Done	2.5 million	11/30/2012	Demographics to include: Age (i2b2 shifts age in older patients where patient could be identified),Zipcode,race,gender
Diagnosis	In Progress		12/6/2012	Diagnosis based on ICD9 codes from IDX via EPR2020.
Procedures	In Progress		12/6/2012	Procedures based on CPT and ICD9 coded from IDX via EPR2020
Lab Data	Not Started			Finalized Lab result data from PDS via EPR2020, Phase II will load PDS exsiting ontology as is defined in EPR2020
Med Data	Not Started			Dispensed Meds from SCM , Phase II will use rxNorm mapping where possible, otherwise Multim ontology as is in SCM.

Chris Shafer, CCDA

Sample Registry Query

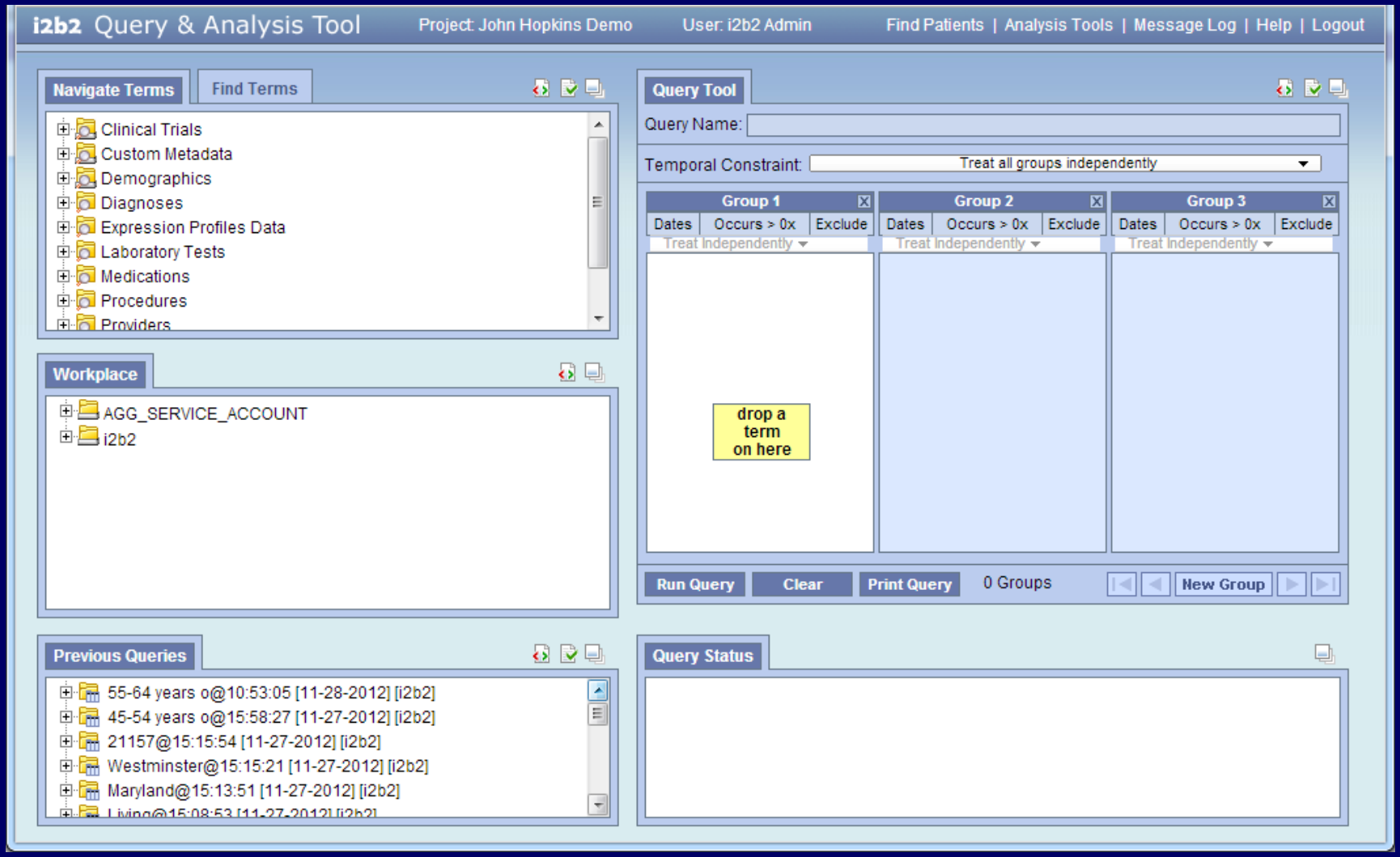
Category	Criteria
Demographic	• Adults Aged 21-64
Diagnosis	• Type 2 Diabetes Mellitus • Obesity
Lab Results	• HbA1c ≤ 8.5% within 8 months

Chris Shafer, CCDA

Query of Dr Jeanne Clark

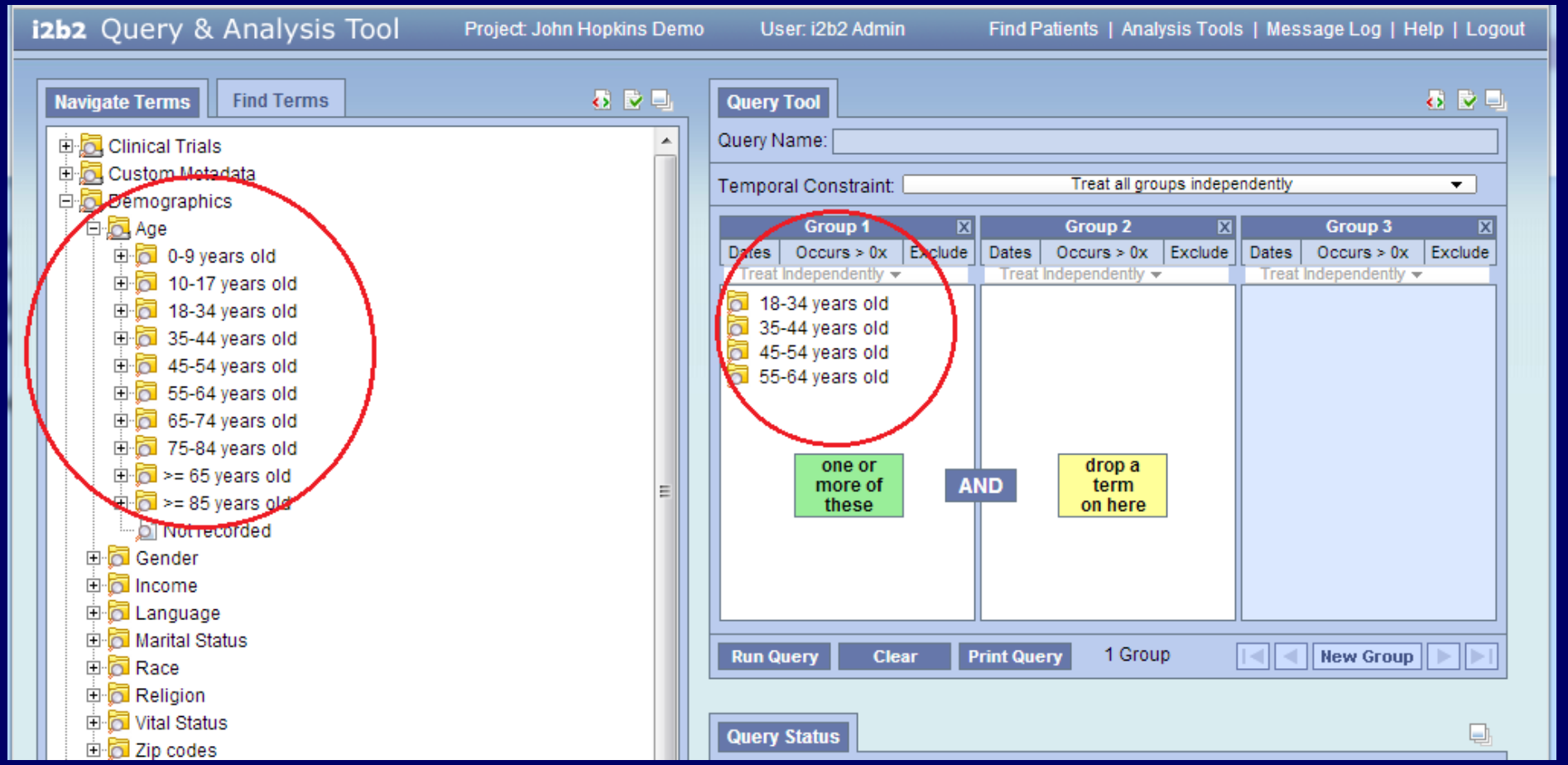
I2b2 at JHMI

- The Find Patients Query Engine



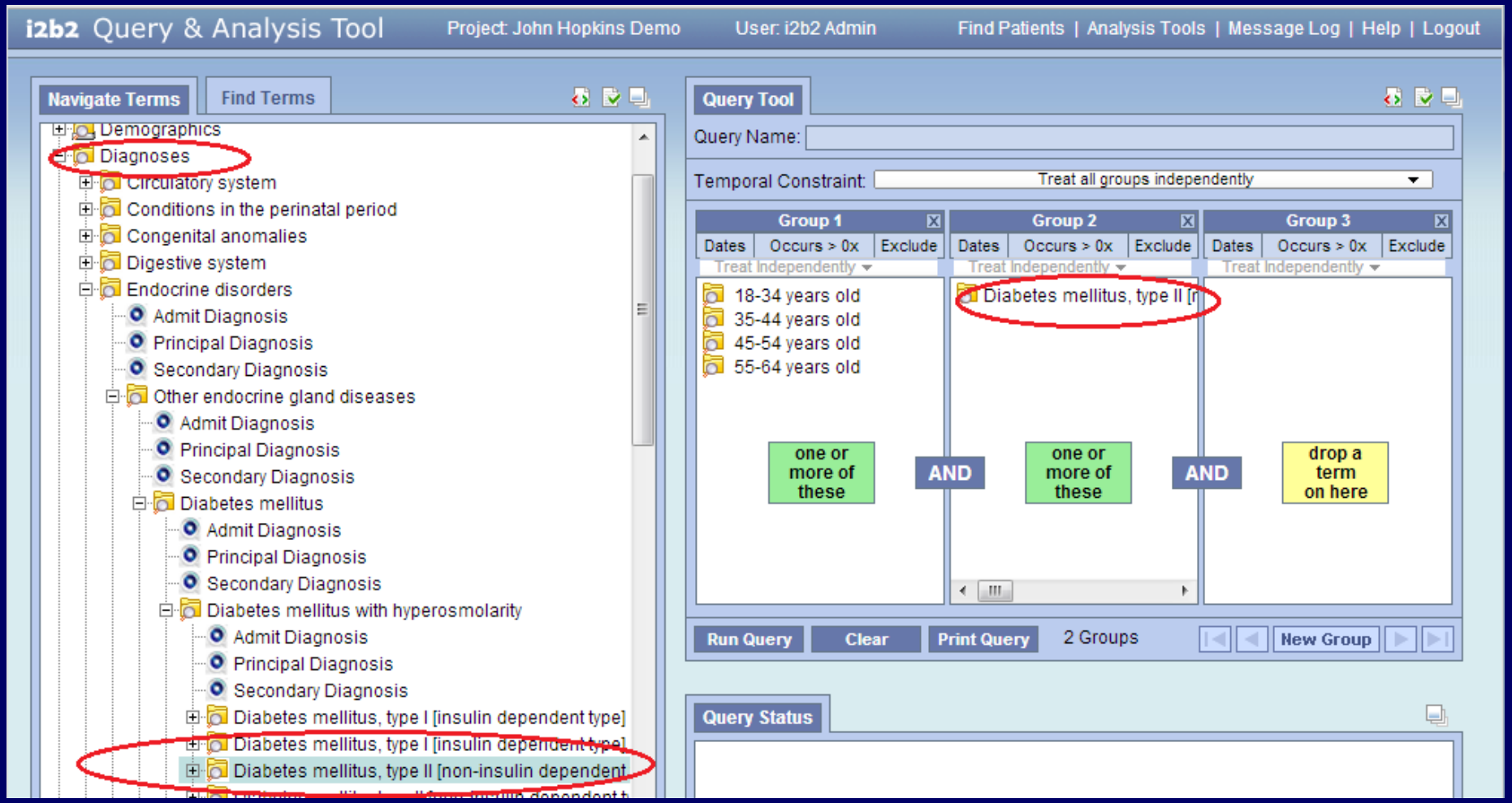
Chris Shafer, CCDA

Age Range Selection



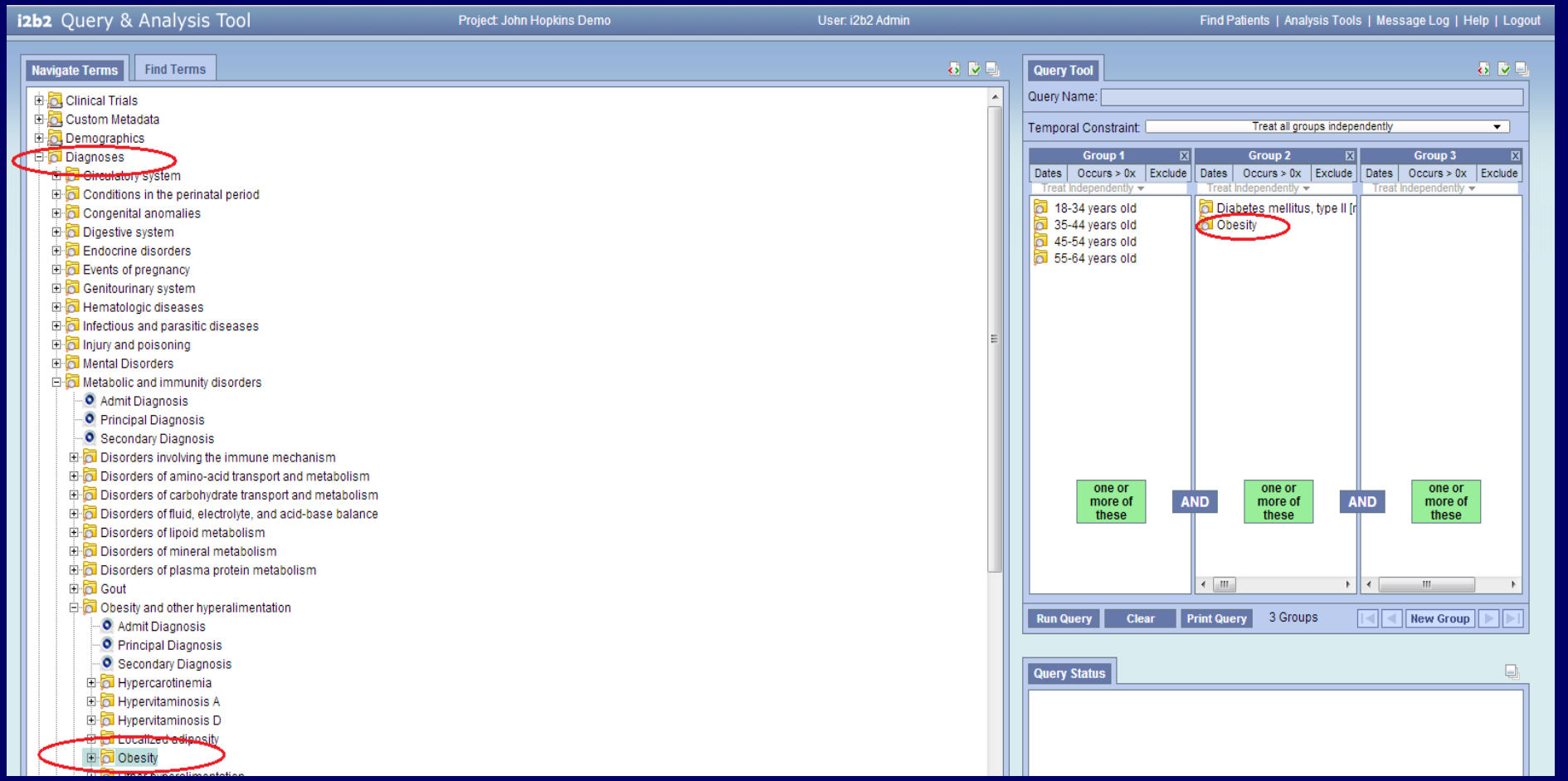
Chris Shafer, CCDA

Diagnosis Selection 1 of 2



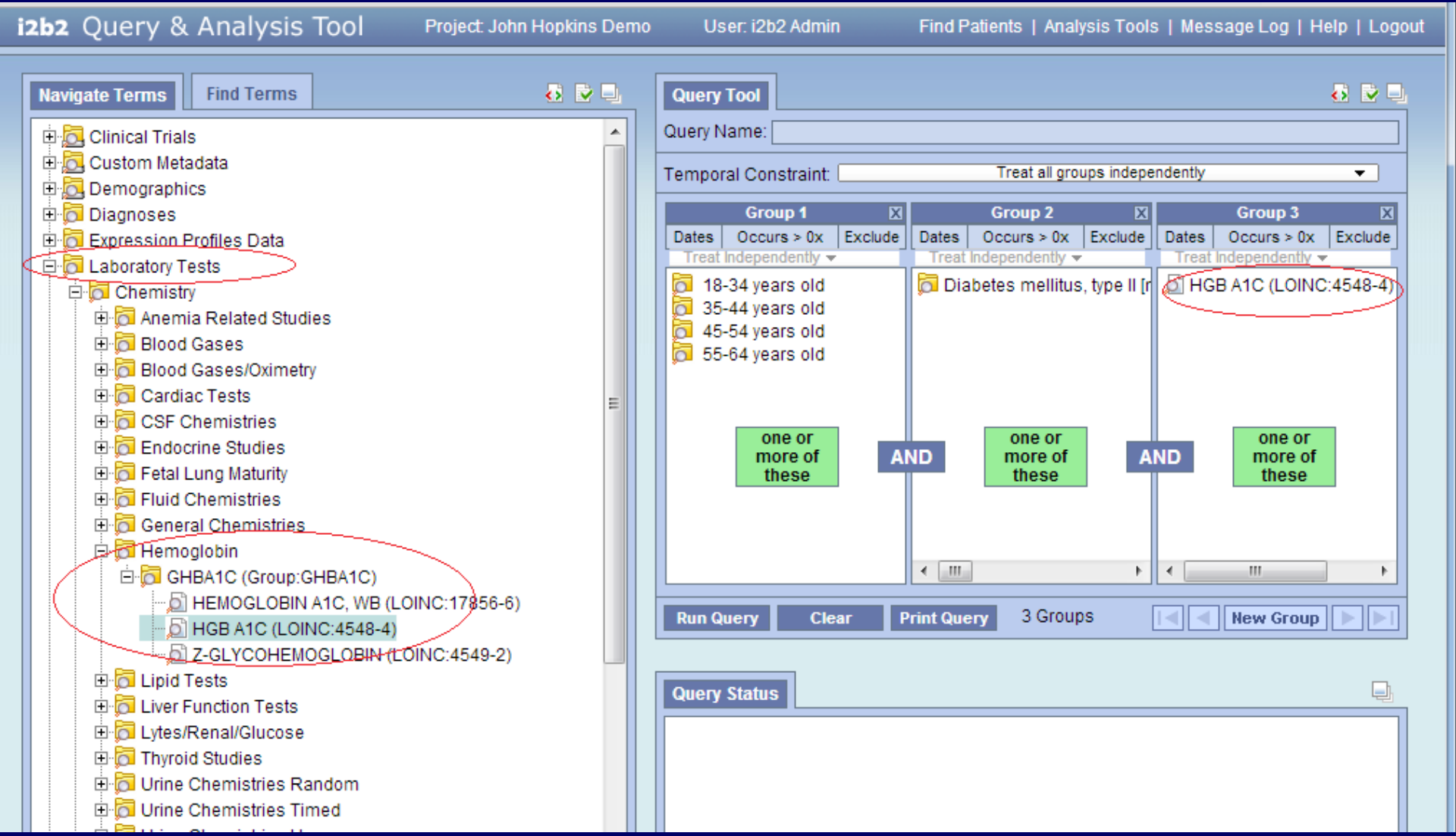
Chris Shafer, CCDA

Diagnosis Selection 2 of 2



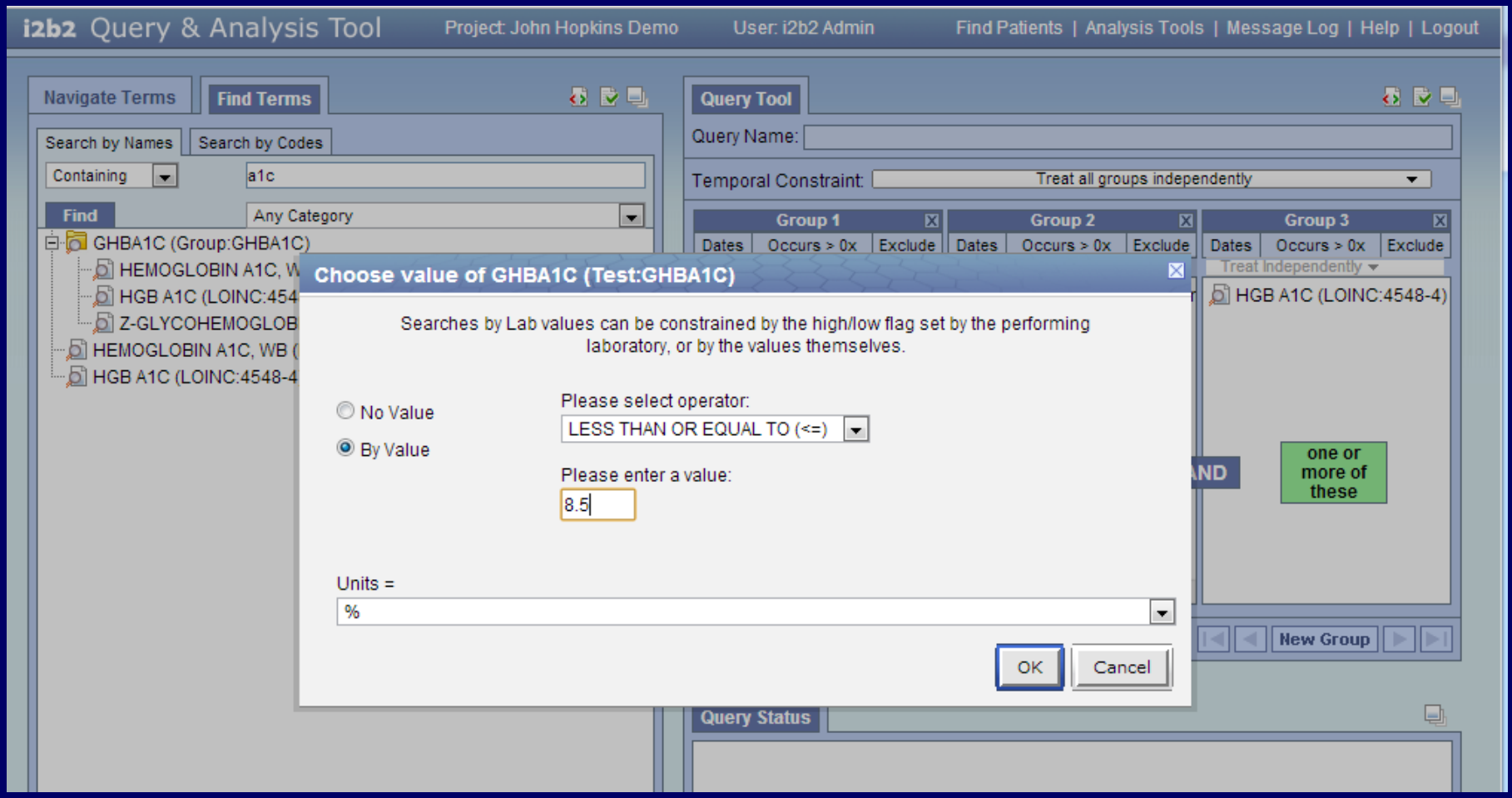
Chris Shafer, CCDA

Lab Result Selection



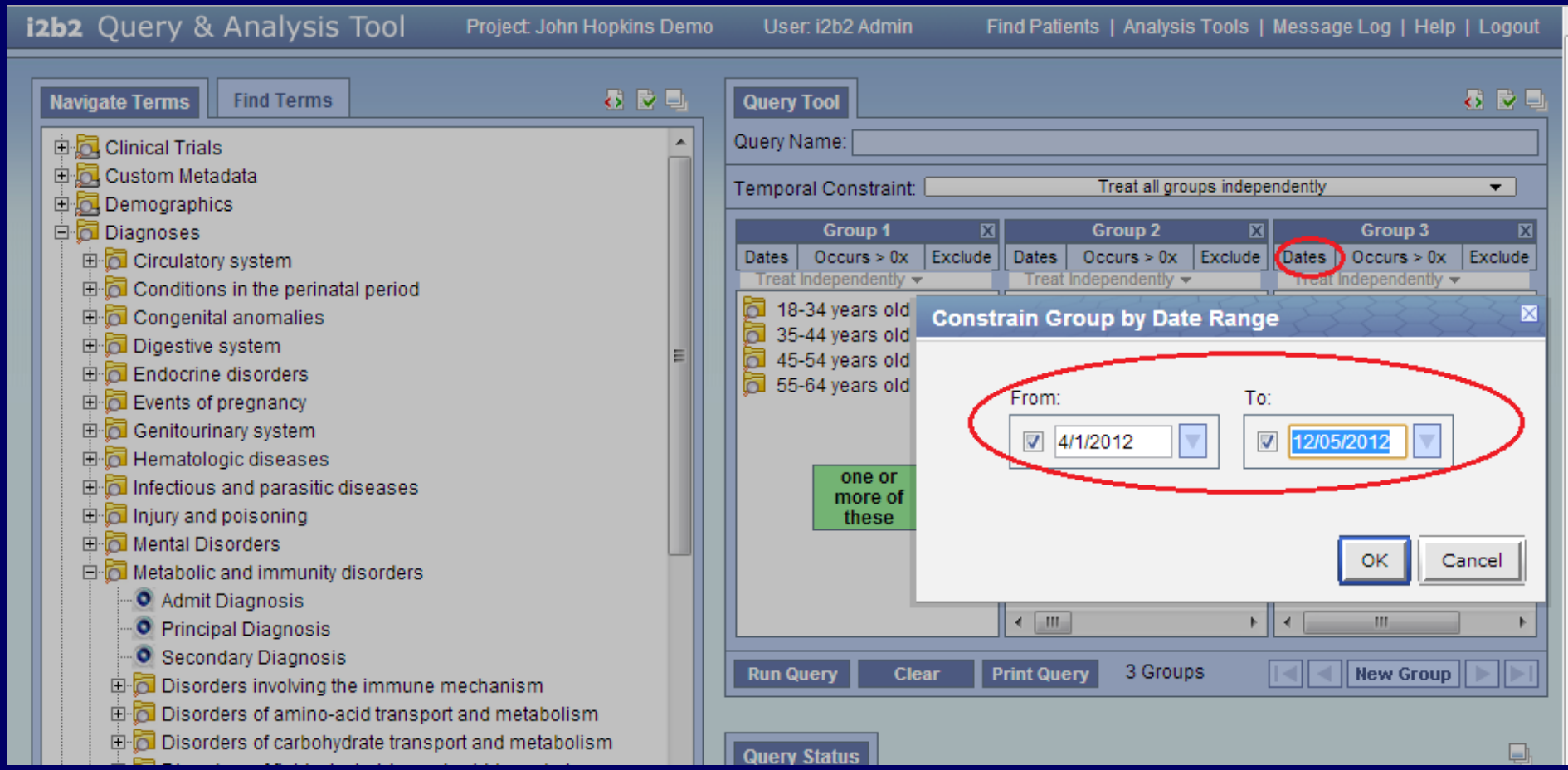
Chris Shafer, CCDA

Lab Result Range Selection



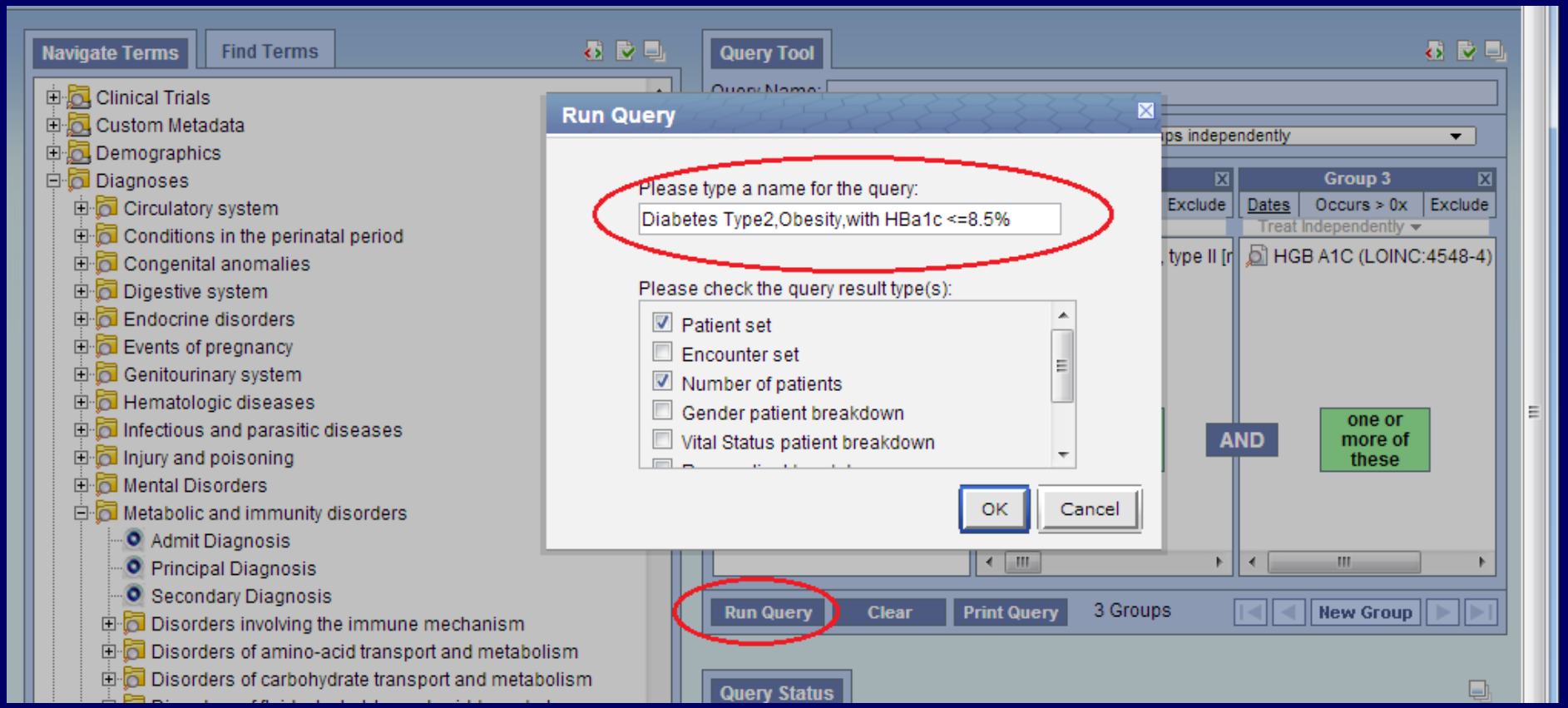
Chris Shafer, CCDA

Lab Result Date Range



Chris Shafer, CCDA

Running and Naming a Query



Chris Shafer, CCDA

Printing a Query

Query Name: Diabetes Type2,Obesity,with HBa1c <=8.5%
Temporal Constraint: Treat all groups independently

Group 1

Date From: none **Date To:** none **Excluded?** false **Occurs X times:** > 0 **Relevance %:** 100 **Temporal Constraint:** Treat Independently

Path	Concept/Term	Other Information
Demographics \ Age \ 18-34 years old	18-34 years old	
Demographics \ Age \ 35-44 years old	35-44 years old	
Demographics \ Age \ 45-54 years old	45-54 years old	
Demographics \ Age \ 55-64 years old	55-64 years old	

Group 2

Date From: none **Date To:** none **Excluded?** false **Occurs X times:** > 0 **Relevance %:** 100 **Temporal Constraint:** Treat Independently

Path	Concept/Term	Other Information
Diagnoses \ ... \ Diabetes mellitus, type II [non-insulin dependent {..}] or unspecified type with hyperosmolarity, not stated as uncontrolled	Diabetes mellitus, type II [non-insulin dependent type] [NIDDM type] [adult-onset type] or unspecified type with hyperosmolarity, not stated as uncontrolled	
Diagnoses \ Metabolic and immunity disorders \ Obesity and other hyperalimentation \ Obesity	Obesity	

Group 3

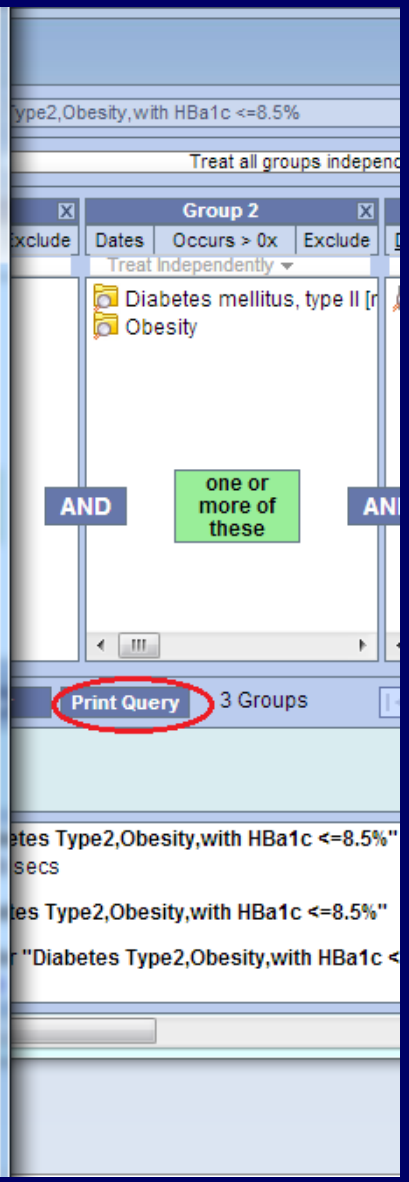
Date From: 04/01/2012 **Date To:** 12/5/2012 **Excluded?** false **Occurs X times:** > 0 **Relevance %:** 100 **Temporal Constraint:** Treat Independently

Path	Concept/Term	Other Information
Labtests \ Chemistry \ Hemoglobin \ GHBA1C (Group:GHBA1C) \ Hgb a1c (LOINC:4548-4)	HGB A1C (LOINC:4548-4)	LE : 8.5 %

Finished Query: "Diabetes Type2,Obesity,with HBa1c <=8.5%" [20.6 secs]
Compute Time: 18 secs

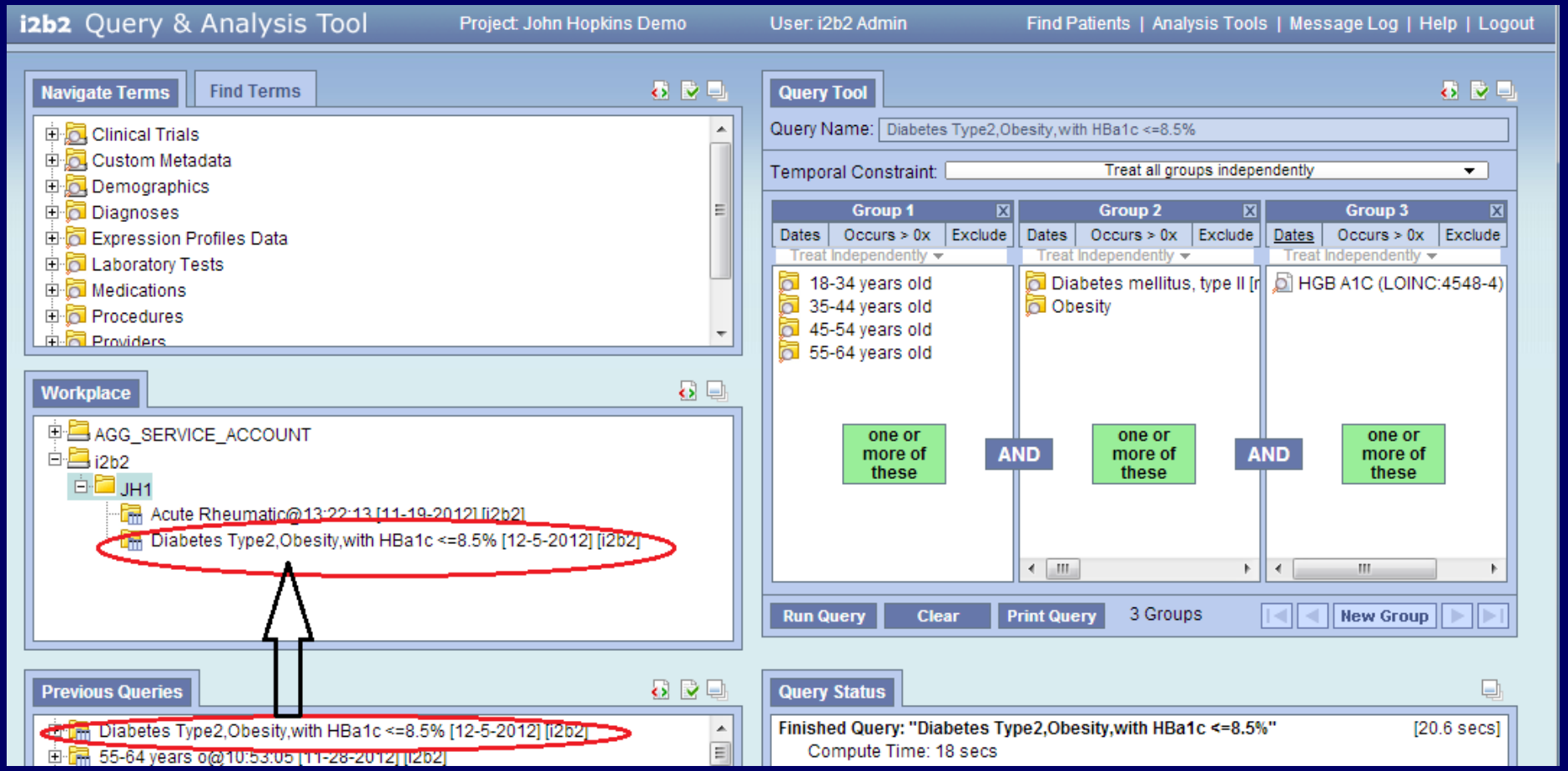
Patient Set for "Diabetes Type2,Obesity,with HBa1c <=8.5%"

Number of patients for "Diabetes Type2,Obesity,with HBa1c <=8.5%"
patient_count: 189



Chris Shafer, CCDA

Saving a Query



Chris Shafer, CCDA

Quality (ICD9 Codes)

Fast is fine, but accuracy is everything.

(Wyatt Earp)

Caution – Data Validity Must Be Monitored

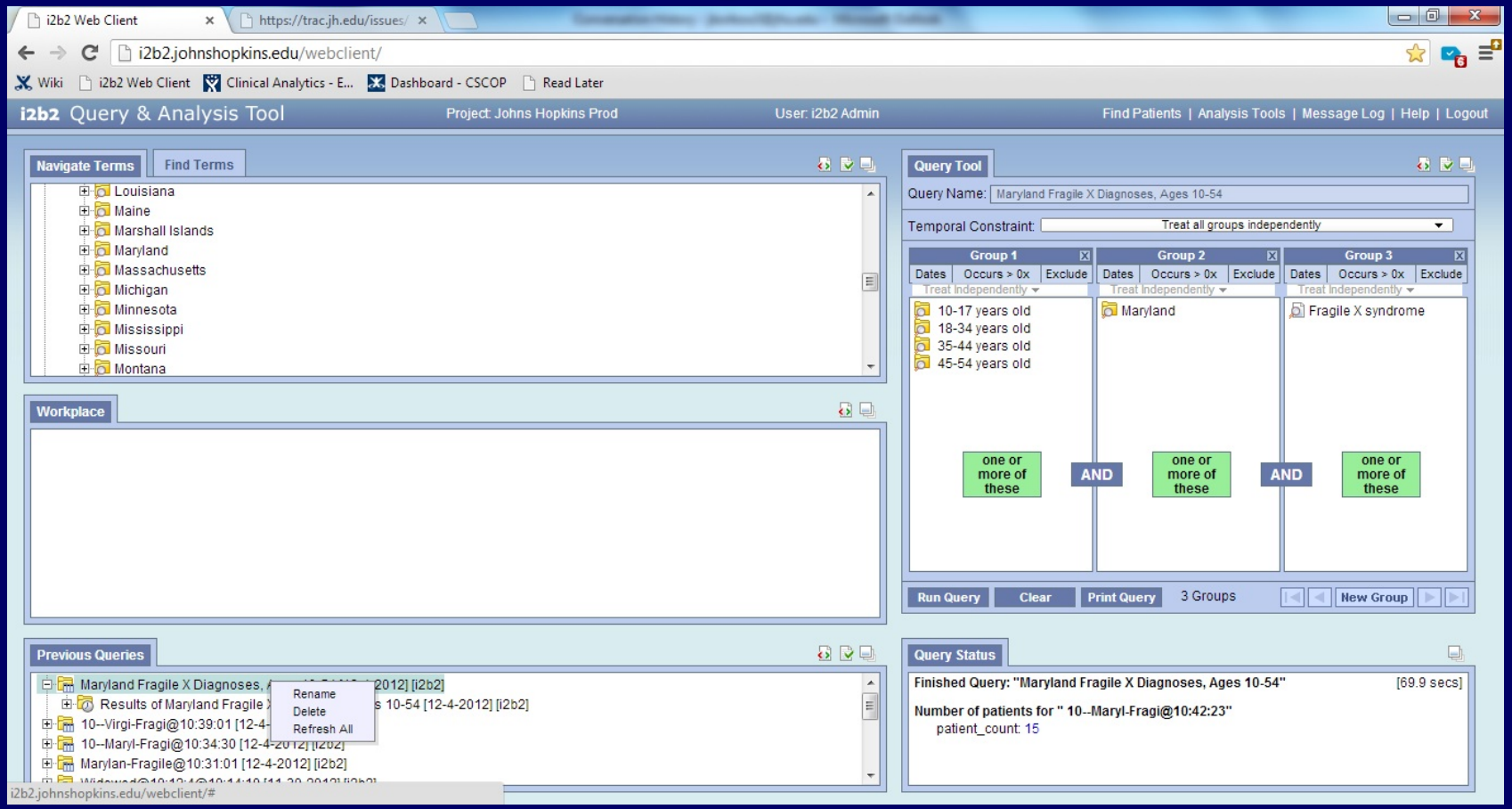
Future State

- Single platform – Epic
- Enterprise Governance
- Research Architecture
- Center for Clinical Data Analysis
- Self – Service
- Open Source Tools
- Questions

I2b2 Demo – JH Instance

- Clinical Analytics Team and CCDA
- Positioning for Enterprise Launch
  - SQL and Linux Servers from ESM/Midrange/Cloud Services
  - Pending integration with Enterprise Active Directory
  - Data sources to include EPR2020 and SCM
- Project Phases
  - Demo Phase
  - Starting Data Set Phase (EPR2020)
  - Cross System phase (Blended EPR2020 / SCM)
  - Full Data Set Phase

Fragile X Query



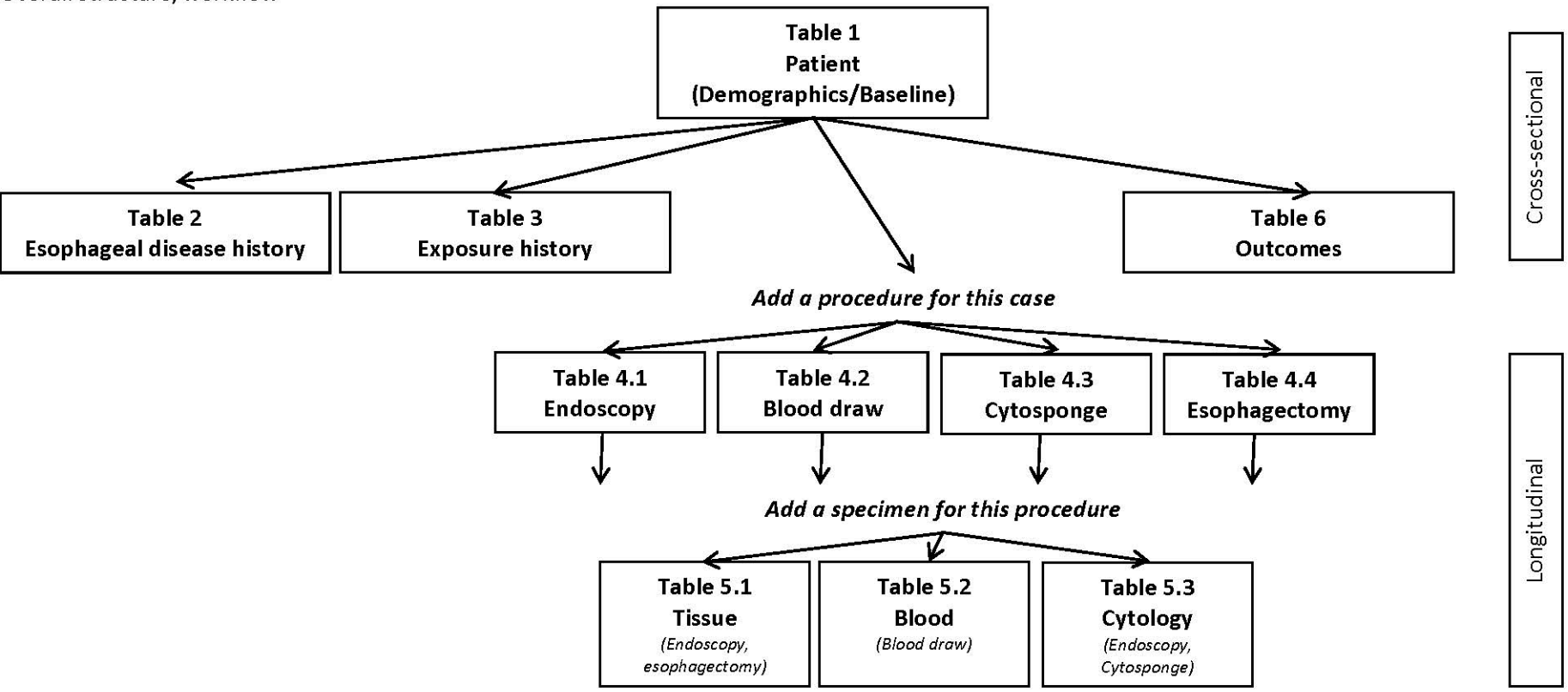
Use Case: Search Q of patients with Fragile X diagnosis in Maryland. Output is sample cohort for IRB request to contact patients with diagnosis for invitation to participate in JHM study

Chris Shafer, CCDA

Example of Data Accessible Through i2b2/SHRINE

BETRNet Virtual Biorepository Database

Overall structure/workflow



{General note on skip logic: Child fields follow any skip-logic imposed on parent field (as well as any skip logic specific to the child); in other words, if the parent field is skipped due to the value entered for the grandparent field, the child field also will be skipped.}

Table 1
Demographics/baseline
Repeated measures: No
Child of: N/A

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
		One-time measures			
Demographics / baseline	01	Patient ID	Comment box	Character string	Database case ID assigned upon case submission. All measures for a given case a keyed to this ID. Research sites are respons for maintaining a concordance log to matc database case ID to local study ID, medical record number, or other unique local ident
Demographics / baseline	02	Consent status	Dropdown menu	Agree Disagree Withdraw consent	Patient consent for participation in biomed research. Data accepted only for cases wit consent status indicated as 'Agree'.
		(a) Date of consent	Calendar	YYYY-MM-DD	Date on which consent form is signed by patient
{skip if consent status !=Agree}		(b) Type of consent	Checkboxes	Primary research Secondary research Broad research Associated with medical information Re-contact patient	For patients who agree to participate in research, record type of consent. **Primary** → Patient consents to use of specimen/data only in the study specificall defined and described in the consent docu **Secondary** → Patient consents to use of specimen/data in additional studies beyon scope of the specific study defined in the consent document **Broad** → Patient consents to use of specimen/data in any biomedical study **Associated with medical information** → Patient consents to association of specime with medical record and history **Re-contact patient** → Patient consents to future contact about potential use of specimen/data in future studies All categories that apply should be indicate For patients enrolled in BETRNet studies, '**research consent**', allowing specimens to b '**associated with medical information**' will maximize accessibility of specimens and da across the network.
Demographics / baseline	03	Date of baseline data collection	Calendar	YYYY-MM-DD	Date on which patient baseline data is colle

Table 3
Exposure history
Repeated measures: No
Child of: Demographics/baseline

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
Exposure history	01	What is the patient's smoking history?	Dropdown menu	Ever-smoker Never-smoker Unknown	Patient history of smoking tobacco produc **Ever-smoker**→ Patient has smoked at leas cigarettes in his or her lifetime **Never-smoker** → Patient has smoked fewe than 100 cigarettes in his or her lifetime
{skip if smoker != Ever}		(a) Ever-smoker: current or former?	Dropdown menu	Current smoker Former smoker Unknown	If patient is ever-smoker, record whether patient still smokes at time of baseline dat collection (**current smoker**) or has quit sm (**former smoker**)
Exposure history	02	What is the patient's alcohol consumption history?	Dropdown menu	Ever-drinker Never-drinker Unknown	Patient history of consuming alcoholic beverages **Ever-drinker**→ Patient has had at least 12 alcoholic drinks in his or her lifetime **Never-drinker** → Patient has had fewer th alcoholic drinks in his or her lifetime
{skip if drinker != Ever}		(a) Ever-drinker: current or former	Dropdown menu	Current drinker Former drinker Unknown	If patient is ever-drinker, record whether patient still drinks alcohol at time of baseli data collection (**current drinker**) or no lon drinks alcohol (**former drinker**)
Exposure history	03	Does the patient have previous or current diagnosis of H. pylori infection?	Dropdown menu	Yes No Unknown	Indicator of whether patient has been diagnosed with past and/or current infecti with H. pylori bacterium
Exposure history	04	What is the patient's history of PPI usage?	Dropdown menu	Ever-user Never-user Unknown	Patient use of proton pump inhibitor (PPI) medications **Ever-user** → Patient has taken PPIs at som time, past or present **Never-user** → Patient has never taken PPI
{skip if PPI !=Ever}		(a) PPI ever-user: current or former?	Dropdown menu	Current user Former user Unknown	If patient is ever-user of PPIs, record whet patient still uses PPIs at time of baseline da collection (**current user**) or no longer takes (**former user**)
{skip if PPI ever !=Current}		(i) Single or double-dose equivalent?	Dropdown menu	Single-dose Double-dose Unknown	If patient is current user of PPIs, record wh patient's PPI dosage is a single- or double- equivalent
Exposure history	05	What is the patient's history of aspirin and/or NSAID usage?	Dropdown menu	Daily user Occasional user Never user Unknown	Patient use of aspirin or NSAID medication **Daily user** → Patient takes aspirin or NSAI every day **Occasional user** → Patient takes aspirin or NSAIDs sometimes, but less often than eve day **Never user** → Patient never takes aspirin o NSAIDs

Table 4
Procedures
Repeated measures: Yes
Child of: Demographics/baseline

To add a procedure for a patient:
Endoscopy: table 4.1
Blood draw: table 4.2
Cytosponge: table 4.3
Esophagectomy: table 4.4

Table 4.1
Procedure: Endoscopy

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
Procedure: Endoscopy	01	Procedure type	Auto-filled	Endoscopy	This table records procedure type → **endoscopy**
Procedure: Endoscopy	02	Date of procedure	Calendar	YYYY-MM-DD	Date of endoscopy procedure
Procedure: Endoscopy	03	Was BE (columnar epithelium) visible on endoscopy?	Dropdown menu	Yes No Unknown	Indicator of whether Barrett's esophagus v visible on endoscopy
{skip if endoscopic BE != Yes}		(a) C (circumferential) value (cm)	Numerical	String (decimal)	If BE visible, record C (circumferential) valu extent of BE, in units of centimeters. Enter for unknown.
		(b) M (maximal) value (cm)	Numerical	String (decimal)	If BE visible, record M (maximal) value for extent of BE, in units of centimeters. Enter for unknown.
{skip if C+M values are known}		(c) Length of BE segment (cm)	Numerical	String (decimal)	If BE visible, and C+M values not known, re length of BE segment, in units of centimete Enter 999 for unknown.

Table 5
Specimens
Repeated measures: Yes
Child of: Procedure

To add a specimen for a procedure:
Tissue: table 5.1 (available for procedures: endoscopy, esophagectomy)
Blood: table 5.2 (available for procedures: blood draw)
Cytology: table 5.3 (available for procedures: endoscopy, cytosponge)

Table 5.1
Specimen: Tissue
(Allowed procedures: endoscopy, esophagectomy)

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
Specimen: Tissue	01	Specimen type collected	Auto-filled	Tissue	This table records specimen type → **tissue**
Specimen: Tissue	02	Specimen collection procedure	Dropdown menu	Forcep biopsy Endoscopic mucosal resection (EMR) Esophagectomy	Method of obtaining tissue specimen For esophagectomy specimen, 'Esophagec must be selected. This is not a valid value f endoscopic specimens.
Specimen: Tissue	03	Specimen type(s) stored	Checkboxes	Preserved tissue (frozen or fixed) Derived DNA Derived RNA	Record all types of specimen and derivative material stored from tissue specimen colle
Specimen: Tissue	04	Specimen tissue of origin	Dropdown menu	Barrett's esophagus Barrett's esophagus, recurrent post-ablation Squamous esophagus Neosquamous esophagus Gastro-esophageal junction Gastric cardia Duodenum Various (esophagectomy) Other	Anatomical region/organ and/or tissue typ from which specimen was obtained For esophagectomy specimen, 'Various (esophagectomy)' must be selected. This is a valid value for endoscopy specimens. If various samples are taken on endoscopy, a separate specimen record should be entere for each separate specimen.
{skip if tissue of origin != BE, recurrent}		(a) Recurrent BE: Number of follow-up examinations with negative findings, between final ablative session and finding of recurrent BE	Numeric	Integer	If specimen tissue of origin is recurrent BE, record number of follow-up examinations negative findings between final ablative se and finding of recurrent BE. Enter 999 for unknown. **Exam with negative findings** → exam in w a biopsy sample is taken, and biopsy shows metaplasia (or higher) in esophagus, and n dysplasia in gastric cardia

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
{skip if tissue of origin != BE}*		(b) (1) BE: distance from incisors (cm)	Numeric	String (decimal)	If specimen tissue of origin is BE or recurre BE, record distance from incisors (in units c centimeters) of the BE lesion. Enter 999 fo unknown.
		(b) (2) BE: focal lesion?	Dropdown menu	Yes No Unknown	If specimen tissue of origin is BE or recurre BE, indicator of whether BE is a focal lesion
{skip if focal BE != Yes}		(i) Focal BE lesion: clock location	Numeric	1-12	If BE lesion is focal, record lesion's 'clock location' around circumference of esophag Enter 999 for unknown.
{skip if tissue of origin != squam or neosquam}		(c) Squamous or neosquamous esophagus: distance from incisors (cm)	Numeric	String (decimal)	If specimen tissue of origin is squamous or neosquamous, record distance from inciso units of centimeters) of esophageal tissue sampled. Enter 999 for unknown.
{skip if tissue of origin != gastric cardia}		(d) (1) Gastric cardia: curvature	Dropdown menu	Greater Lesser Unknown	If specimen tissue of origin is gastric cardia record whether tissue was sampled from greater or lesser curvature of stomach
		(d) (2) Gastric cardia: A/P	Dropdown menu	Anterior Posterior Unknown	If specimen tissue of origin is gastric cardia record anterior/posterior location of tissue sampled
		(d) (3) Gastric cardia: laterality	Dropdown menu	Right Left Unknown	If specimen tissue of origin is gastric cardia record left/right location of tissue sampled
{skip if tissue of origin != duodenum}		(e) Duodenum: portion	Dropdown	Duodenal bulb Second portion Unknown	If specimen tissue of origin is duodenum, r portion of duodenum from which tissue wa sampled
Specimen: Tissue	05	Highest pathology associated with procedure in which specimen was obtained	Dropdown menu	Normal squamous and/ or normal gastric cardia Non-erosive esophagitis Erosive esophagitis Specialized intestinal metaplasia (SIM) Indefinite Low-grade dysplasia (LGD) High-grade dysplasia (HGD) Adenocarcinoma (EAC) Unknown	Highest pathology seen on endoscopic procedure or esophagectomy in which specimen obtained For esophagectomy, value selected must b 'Adenocarcinoma' Hierarchy of pathologies: normal < non-erc esophagitis < erosive esophagitis < SIM < indefinite < LGD < HGD < EAC

Table 5.2
Specimen: blood
Allowable procedure: blood draw

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
Specimen: Blood	01	Specimen type collected	Auto-filled	Blood	This table records specimen type → **blood**
Specimen: Blood	02	Specimen type(s) stored	Checkboxes	Whole blood Serum Plasma Buffy coat Derived DNA Derived RNA	Record all types of specimen and derivative material stored from blood specimen colle
Specimen: Blood	03	Highest pathology associated with endoscopic examination performed contemporaneous with obtaining specimen	Dropdown menu	Normal squamous and/ or normal gastric cardia Non-erosive esophagitis Erosive esophagitis Specialized intestinal metaplasia (SIM) Indefinite Low-grade dysplasia (LGD) High-grade dysplasia (HGD) Adenocarcinoma (EAC) Unknown Endoscopy not performed at time of specimen collection	Highest pathology seen on endoscopic procedure performed at same time as bloc draw Hierarchy of pathologies: normal < non-ero esophagitis < erosive esophagitis < SIM < indefinite < LGD < HGD < EAC
Specimen: Blood	04	Specimen available for use?	Dropdown menu	Yes No	Indicator of whether specimen is available use by other investigators
{skip if specimen available = No}		(a) Storage location and contact name	Dropdown menu	Case Western – Amitabh Chak Columbia – Julian Abrams Fred Hutch – Bill Grady Johns Hopkins – Mimi Canto Mayo – Ken Wang UMichigan – David Beer UNC – Nick Shaheen UPenn – John Lynch UWashington – Eric Seibel WUSTL – Jean Wang	If specimen available for use, record institu and contact name for specimen
		(b) Local specimen identifier code	Comment box	Character string	If specimen available for use, record local specimen identifier code
		(c) Specimen depleted, lost, or otherwise no longer available for use?	Dropdown menu	Yes No	If specimen available for use at one time is lost, depleted, or otherwise no longer avai for use, change value of this variables from to 'Yes'

Table 5.3
Specimen: Cytology
Allowable procedures: endoscopy, cytosponge

Table {skip logic}	#	Variable name	GUI format	Permissible values	Definition, disambiguation, or other data collection note
Specimen: Cytology	01	Specimen type collected	Auto-filled	Cytology	This table records specimen type → **cytolo**
Specimen: Cytology	02	Specimen collection procedure	Dropdown menu	Brush biopsy Cytosponge Fine needle aspiration (FNA)	Method of obtaining cytology specimen For cytosponge specimen, 'Cytosponge' m be selected. This is not a valid value for endoscopic specimens.
Specimen: Cytology	03	Specimen type(s) stored	Checkboxes	Slide Pellet Derived DNA Derived RNA	Record all types of specimen and derivative material stored from cytology specimen collected
Specimen: Cytology	04	Specimen tissue of origin	Dropdown menu	Barrett's esophagus Barrett's esophagus, recurrent post-ablation Squamous esophagus Neosquamous esophagus Gastro-esophageal junction Gastric cardia Duodenum Lymph node Various (cytosponge scraping) Other	Anatomical region/organ and/or tissue typ from which specimen was obtained For cytosponge specimen, 'Various (cytosp scraping)' must be selected. This is not a va value for brush biopsy or FNA.
{skip if specimen tissue of origin != BE, recurrent}		(a) Recurrent BE: Number of follow-up examinations with negative findings, between final ablative session and finding of recurrent BE	Numeric	Integer	If specimen tissue of origin is recurrent BE, record number of follow-up examinations negative findings between final ablative se and finding of recurrent BE. **Exam with negative findings** → Exam in w a biopsy sample is taken, and biopsy shows metaplasia (or higher) in esophagus, and n dysplasia in gastric cardia

Table 6
Outcomes
Repeated measures: No
Child of: Demographics/baseline

Table {skip logic}	#	Variable name	GUI format	Permissible values	Validation (hard or soft stop) or oth development note
Outcomes	01	Progression to cancer?	Dropdown menu	Yes No Unknown (patient is lost to follow-up)	Indicator of patient progression to cancer. Change value to 'Yes' if/when patient progresses to cancer.
{Skip if cancer != Yes}		(a) Date of cancer diagnosis	Calendar	YYYY-MM-DD	If patient progresses to cancer, record date cancer diagnosis
		(b) Final path report available	Dropdown menu	Yes No Unknown	If patient progresses to cancer, indicator o whether patient's final path report is availa
Outcomes	02	Deceased?	Dropdown menu	Yes (patient has died) No (patient's last known vital status is alive) Unknown (patient is lost to follow-up)	Indicator of patient death during follow-up Change value to 'Yes' if/when patient dies during follow-up.
{Skip if deceased != Yes}		(a) Date of death	Calendar	YYYY-MM-DD	If patient has died, record date of death
		(b) Cause of death (relative to esophageal malignancy)	Dropdown menu	Directly related Indirectly related Not related Unknown	If patient has died, record whether cause o death is known to be directly, indirectly, or related to esophageal malignancy

Where Are We? (i2b2)

(Slide courtesy of Chris Shafer)

- 2.6M deidentified encounters from various parts of JHM – no JHCP data currently
- No Epic data yet – we are working through the planning and approval gateways
- The last five years of numerical lab data (text and complex labs eventually)
- Billing diagnosis and procedure codes – with all associated caveats
- Demographics – age is bucketed to 10 year brackets for high level search

caTissue i2b2 Integration

Jefferson Kimmel Cancer Center
NCI-designated

THOMAS JEFF

PATIENT CARE | RESEARCH | EDUCATION | ABOUT THE KCC | HOW TO HELP |

HOME » RESEARCH

INFORMATICS SHARED RESOURCE

DIRECTOR:

Jack London, PhD
Tel: (215) 503-4599
Email: jack.london@mail.jci.tju.edu

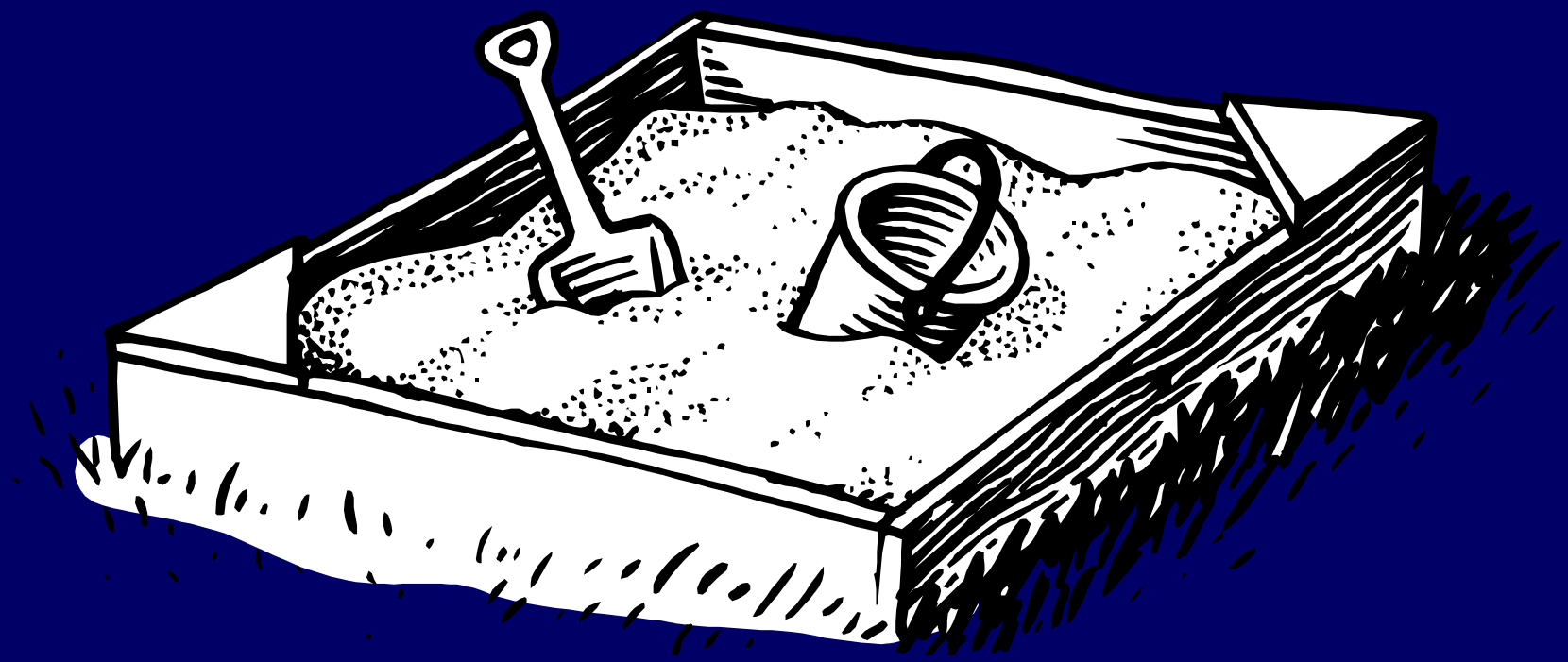
LOCATION:

812 BLSB
233 S. 10th Street
Philadelphia, PA 19107

CONTACT:

Joseph White
Tel: (215) 503-4606
Email: ISRHelp@mail.jci.tju.edu

The Informatics Shared Resource (ISR) serves as a biomedical research informatics resource for KCC's basic, clinical, and translational investigators. This support includes infrastructural provisioning, software development or acquisition, and consultative collaboration. ISR services are available directly to peerreviewed funded cancer investigators, or indirectly to other cancer center core facilities, thereby providing services to all program members. The ISR also provides informatics support for the clinical research activities of the Jefferson Kimmel Cancer Center Network of community hospitals.



Thank You!

- Questions?

References:

- Best Practices for Repositories I: collection, storage and retrieval of human biological materials for research, ISBER. Cell Preservation Technology; Volume3, Number1, 2005, Update 2008.
- National Cancer Institute, Best Practices for Biospecimen Resources 2007
- Case Studies of Human Tissue Repositories; Best Practices for a Biospecimen Resource for the Genomic and Proteomic Era. RAND Scientific Corporation, prepared for the National Cancer Institute. http://biospecimens.cancer.gov/nbn/rand.asp

References:

- Cancer Bioinformatics Grid, caBIG: https://cabig.nci.nih.gov/.
- First Generation Guidelines for NCI Supported Biorepositories:http://biospecimens.cancer.gov/biorepositories/NCI-Supported_Biorepositories.pdf
- Federal Register / Vol. 71, No. 82 / Friday, April 28, 2006. First-Generation Guidelines for NCI Supported Biorepositories

Save the Bay

Save the Bay