

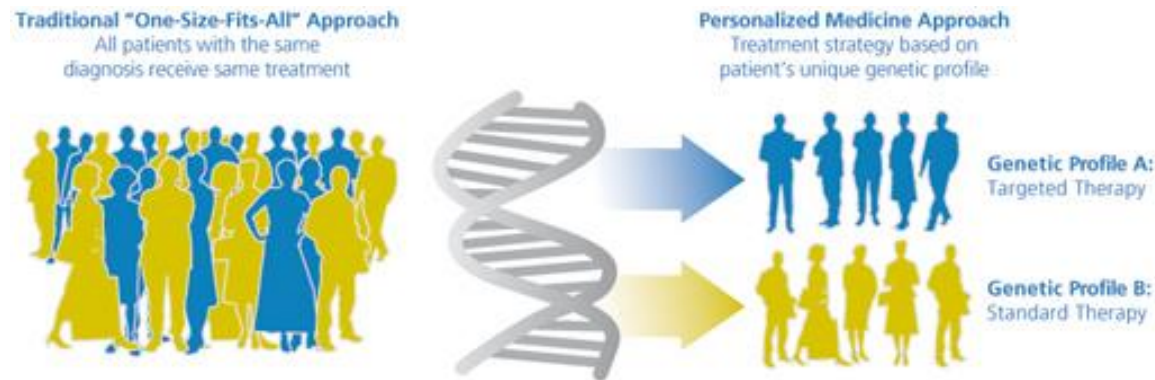
# **Integrating Genomic, Imaging, and Clinical Data for Precision Medicine**

*Shawn Murphy MD, Ph.D.*

*Partners Healthcare and Harvard Medical School*

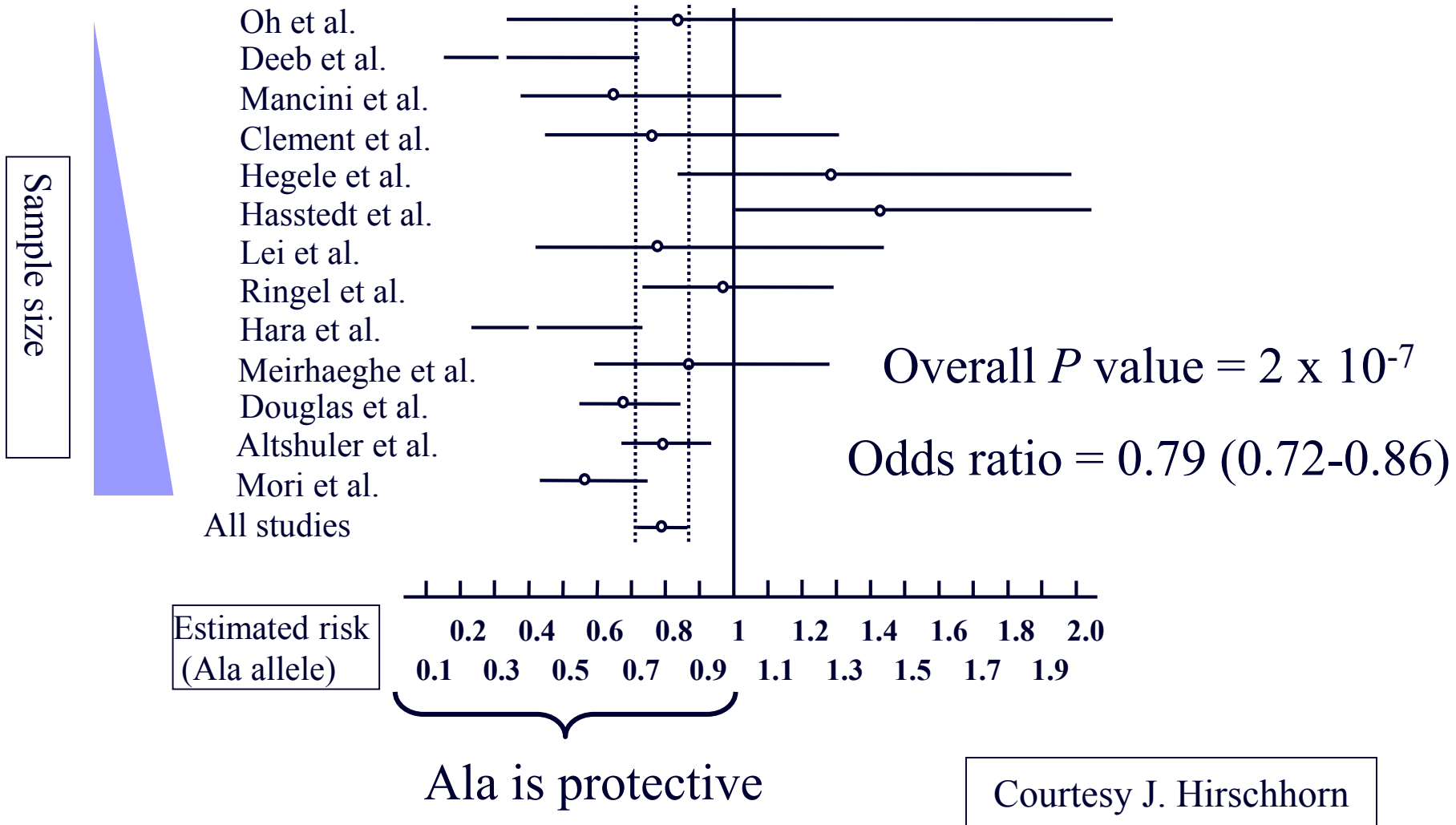
*October 4<sup>th</sup>, 2017*

# Personalized Medicine and Genomic technology are critical to managing populations



- Managing a population involves improving health outcomes of the group as a whole by identifying, monitoring and addressing health needs of individuals through:
  - Subpopulation stratification
  - Targeted, evidence-based treatment protocols
  - Predictive analytics

# Example: PPAR $\gamma$ Pro12Ala and Diabetes



# High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion
- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data
- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing
- Personalized medicine is delivered into clinical care

# High Throughput Methods for supporting Translational Research

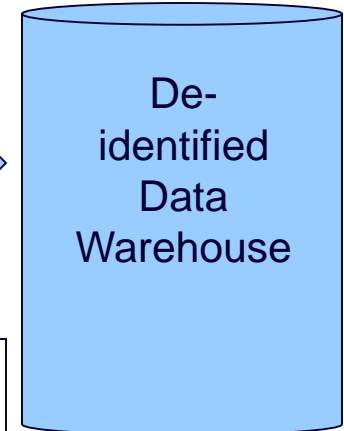
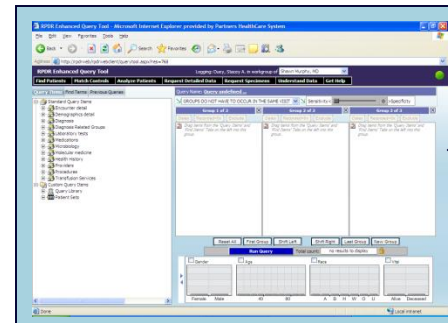
- Set of patients is selected from medical record data in a high throughput fashion
- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data
- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing
- Personalized medicine is delivered into clinical care

# Research Patient Data Registry exists at Partners Healthcare to find patient cohorts for clinical research

## 1) Queries for aggregate patient numbers

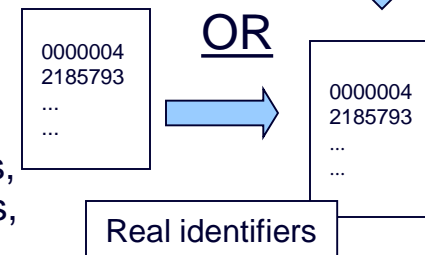
- Warehouse of in & outpatient clinical data
- 6.7 million Partners Healthcare patients
- 2.5 billion diagnoses, medications, procedures, laboratories, & physical findings coupled to demographic & visit data
- Authorized use by faculty status
- Clinicians can construct complex queries
- Queries cannot identify individuals, internally can produce identifiers for (2)

### Query construction in web tool



## 2) Returns identified patient data

- Start with list of specific patients, usually from (1)
- Authorized use by IRB Protocol
- Returns contact and PCP information, demographics, providers, visits, diagnoses, medications, procedures, laboratories, microbiology, reports (discharge, LMR, operative, radiology, pathology, cardiology, pulmonary, endoscopy), and images into a Microsoft Access database and text files.



Test_ID	Test_Description	Result	Result_Text	Abnormal_Flag	Reference	Unit	Reference_Range
SPOPT	APPT	21.8				sec	22.1-26.1
SPOPT	APPT	21.8				sec	22.1-26.1
SPOPT	APPT	46.4		H		sec	22.1-26.1
SPOPT	APPT	43.1				sec	22.1-26.1
SPOPT	APPT	26.7	MODERATELY H			sec	22.1-26.1
SPOPT	APPT	23.7				sec	22.1-26.1
SPOPT	APPT	24.4				sec	22.1-26.1
SPOPT	APPT	24.7				sec	22.1-26.1
SPOPT	APPT	24.0				sec	22.1-26.1
SPOPT	APPT	24.7				sec	22.1-26.1
SPOPT	Survival APPT	31.3				sec	22.1-26.1
SPOPT	APPT	34.5				sec	22.1-26.1
SPOPT	APPT	40.0		H		sec	22.1-26.1
SPOPT	APPT	46.0		H		sec	22.1-26.1
SPOPT	Survival APPT	50.2		H	Note: None in H	sec	22.1-26.1
SPOPT	APPT	51.6				sec	22.1-26.1
SPOPT	Survival APPT	34.3				sec	22.1-26.1
SPOPT	APPT	37.9		H		sec	22.1-26.1
SPOPT	APPT	22.6				sec	22.1-26.1
SPOPT	APPT	37.4		H		sec	22.1-26.1
SPOPT	APPT	37.2				sec	22.1-26.1
SPOPT	APPT	38.4		H		sec	22.1-26.1
SPOPT	APPT	38.4	MODERATE HE H			sec	22.1-26.1

# FINDING PATIENTS

Query items

Person who is using tool

The screenshot displays the RPDR Enhanced Query Tool interface within a Microsoft Internet Explorer browser. The browser title bar reads "RPDR Enhanced Query Tool - Microsoft Internet Explorer provided by Partners HealthCare System". The address bar shows the URL "http://rpdweb/rpdwebclient/querytool.aspx?res=768". The tool's header includes a navigation menu with buttons for "Find Patient", "Match Controls", "Analyze Patients", "Request Detailed Data", "Request Specimens", "Understand Data", and "Get Help". The user is logged in as "Duey, Stacey A. in workgroup of Shawn Murphy, MD".

The main interface is divided into several sections:

- Query Items:** A tree view on the left lists categories such as "Standard Query Items" (including Encounter detail, Demographics detail, Diagnosis, etc.) and "Custom Query Items" (including Query Library and Patient Sets).
- Query Construction:** The central area shows a "Query Name: Query undefined ..." and a "GROUPS DO NOT HAVE TO OCCUR IN THE SAME PLACE" option. It features three group containers (Group 1 of 3, Group 2 of 3, Group 3 of 3) with fields for "Dates", "Recorded>0x", and "Exclude". Each group contains a placeholder instruction: "Drag items from the 'Query Items' and 'Find Items' Tabs on the left into this group." Below the groups are buttons for "Reset All", "Clear", "Limit Left", "Limit Right", "Test Group", and "New Group".
- Run Query:** A prominent blue button labeled "Run Query" is present, along with a "Total count: no results to display" indicator.
- Results:** At the bottom, there are four summary tables for "Gender", "Age", "Race", and "Vital". The "Gender" table shows "Female" and "Male" columns. The "Age" table shows "40" and "80" columns. The "Race" table shows "A", "B", "H", "W", "O", "U" columns. The "Vital" table shows "Alive" and "Deceased" columns.

Query construction

Results - broken down by number distinct of patients

Query Items Find Terms Previous Queries

Search For: Containing

All Categories

Search Items

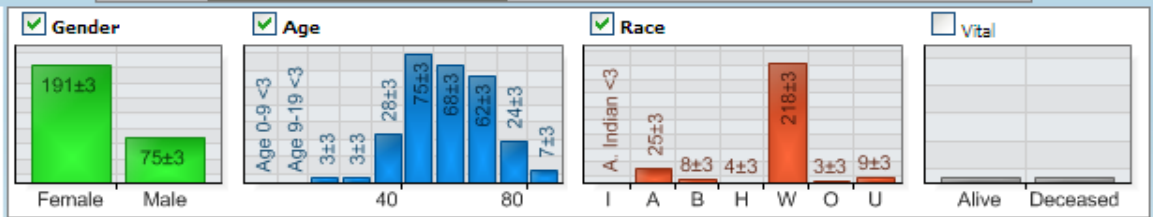
- EGFR
  - eGFR (Test:bc1-1384)
  - eGFR (Test:fc500.1750)
  - eGFR (Test:fc500.1800)
  - eGFR (Test:fc500.1850)
  - eGFR (Test:mcsq-egfr)
  - eGFR (Test:mcsq-egfr1)
  - eGFR (Test:mcsq-pegfr)
  - eGFR (Test:ncgfrnaa)
- EGFR Gene Mutations (Group:EGFR)
  - EGFR Sequencing (Test:mcsq-egfrs)

Query Name: EGFR, Respiratory and... on 01/24/2011 #3

GROUPS DO NOT HAVE TO OCCUR IN THE SAME VISIT  >Specificity

Group 1 of 3	Group 2 of 3	Group 3 of 3
Dates Recorded>0x Exclude	Dates Recorded>0x Exclude	Dates Recorded>0x Exclude
<p>One or more items recorded</p> <ul style="list-style-type: none"> <li>EGFR                             <ul style="list-style-type: none"> <li>[2236_2252del; 2258delC Responsive)</li> <li>2235_2249del (Responsive)</li> <li>2236_2252del 2257delC Responsive)</li> <li>2261A&gt;G (Presumed Res)</li> <li>2264C&gt;A (Presumed Res)</li> <li>2314_2319dup (Unknown Significance)</li> <li>2317_2319dupCAC (Unknown Significance)</li> <li>c.2065G&gt;C (Unknown Sig)</li> <li>c.2093C&gt;T (Unknown Sig)</li> <li>c.2117T&gt;C (Unknown Sig)</li> <li>c.2125G&gt;A (Responsive)</li> <li>c.2126A&gt;T (Unknown Sig)</li> </ul> </li> </ul>	<p>One or more items recorded</p> <ul style="list-style-type: none"> <li>Respiratory and intrathoracic organs                             <ul style="list-style-type: none"> <li>Malignant neoplasm of larynx</li> <li>Malignant neoplasm of nasal cavities, middle ear and accessory sinuses</li> <li>Malignant neoplasm of other and ill-defined sites within the respiratory system and intrathoracic organs</li> <li>Malignant neoplasm of pleura</li> <li>Malignant neoplasm of thymus, heart, and mediastinum</li> <li>Malignant neoplasm of</li> </ul> </li> </ul>	<p>Drag items from the 'Query Items' and 'Find Items' Tabs on the left into this group.</p>

Total count: 269±3 patient(s)





RPDR Detailed Data Request Wizard -- Web Page Dialog



Using Partners IRB#2002P000381 (Research Patient Data Registry (RPDR)) to obtain data from the RPDR

You are logged in as Duey, Stacey A. in workgroup Shawn Murphy, MD

**Please enter your IRB protocol.**

Partners IRB (required):	2002P000381
	Title: Research Patient Data Registry (RPDR)
	Status: Active - Ongoing
Newton Wellesley Hospital IRB:	
Spaulding Rehabilitation Hospital IRB:	
North Shore Medical Center IRB:	NSM 2008-786 demo
	Title:
	Status:

Options for returned set of patients:

- Exclude Partners Healthcare employees
- Create a static set of patients from this query that can be used in other RPDR queries
- Rerun the base query shown above to obtain a fresh set of patients

Help

< Back

**Step 3**

Next >

Cancel

# Detailed data is gathered

Process IRB files ...

Environment Record Options Help

File: SNM0\_022502164303648842.XML

File Type: Control File Current Production Database: RPDR\_12 RPDR\_12\_5241

Update Status  Start Process After 9:00:00 PM

IRB Information

IRB Number: 2000P000000 Ends:

Date from: 01/01/1900 Date to: 01/01/1900

Primary User: snm0

Files to MGH Users: Partners\snm0,Partners\zpp,Partners\kcs3

BWH Users: Partners\kra1,Partners\snm0,Partners\kcs3

Data Requested

Demographics  Medical Record Numbers  Chemistry

Encounters  Contact Information  Radiology

Hematology  PCP  Pathology

Discharge Summaries  Immunology  LMR Notes

Medications  Operative Notes  LMR Problems

LMR Allergies  LMR Medications  Build Access Database

Groups: BUN

Run Close Clear

Output files placed in special directory

Data is gathered from RPDR and other Partners sources

Microsoft Access - [MGH\_Labs : Table]

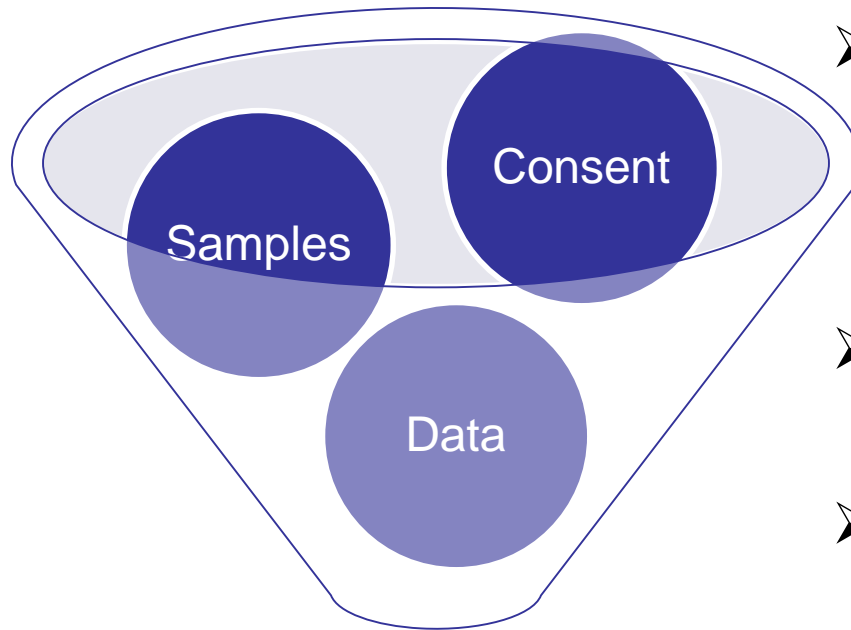
Test Id	Test Description	Result	Result Text	Abnormal Flag	Reference Uni	Reference Range
SQ-XPTT	Superstat APTT	29.8			sec	22.1-35.1
SQ-PTT	APTT	32.8			sec	22.1-35.1
SQ-PTT	APTT	37.8		H	sec	22.1-35.1
SQ-PTT	APTT	46.4		H	sec	22.1-35.1
SQ-PTT	APTT	43.1	MODERATELY	H	sec	22.1-35.1
SQ-PTT	APTT	25.7			sec	22.1-35.1
SQ-PTT	APTT	23.7			sec	22.1-35.1
SQ-PTT	APTT	25.4			sec	22.1-35.1
SQ-PTT	APTT	24.7			sec	22.1-35.1
SQ-PTT	APTT	24.0			sec	22.1-35.1
SQ-PTT	APTT	24.7			sec	22.1-35.1
SQ-XPTT	Superstat APTT	31.3			sec	22.1-35.1
SQ-PTT	APTT	34.5			sec	22.1-35.1
SQ-PTT	APTT	40.0		H	sec	22.1-35.1
SQ-PTT	APTT	45.0		H	sec	22.1-35.1
SQ-XPTT	Superstat APTT	55.2	*** Note: New n	H	sec	22.1-35.1
SQ-PTT	APTT	33.6			sec	22.1-35.1
SQ-XPTT	Superstat APTT	34.3			sec	22.1-35.1
SQ-PTT	APTT	37.9		H	sec	22.1-35.1
SQ-PTT	APTT	22.6			sec	22.1-34.1
SQ-PTT	APTT	37.4		H	sec	22.1-34.1
SQ-PTT	APTT	37.2	SLT HEMOLYS	H	sec	22.1-34.1
SQ-PTT	APTT	35.1		H	sec	22.1-34.1
SQ-PTT	APTT	36.4	MODERATE HE	H	sec	22.1-34.1

Record: 1 of 108164

Datasheet View

Files include Personal Database

# The Partners Biobank



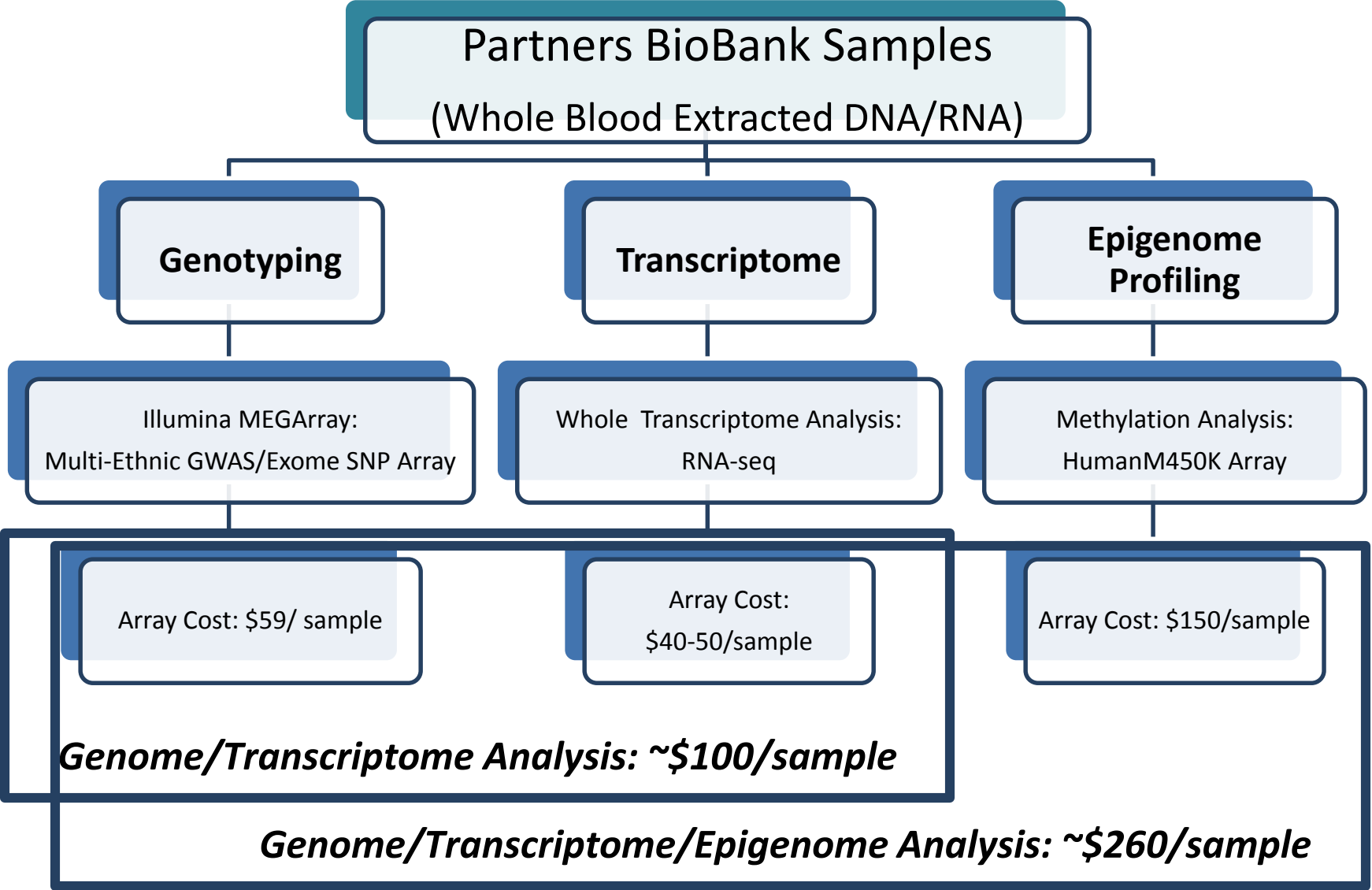
Research Discoveries

**Improved Clinical Care for All Patients**

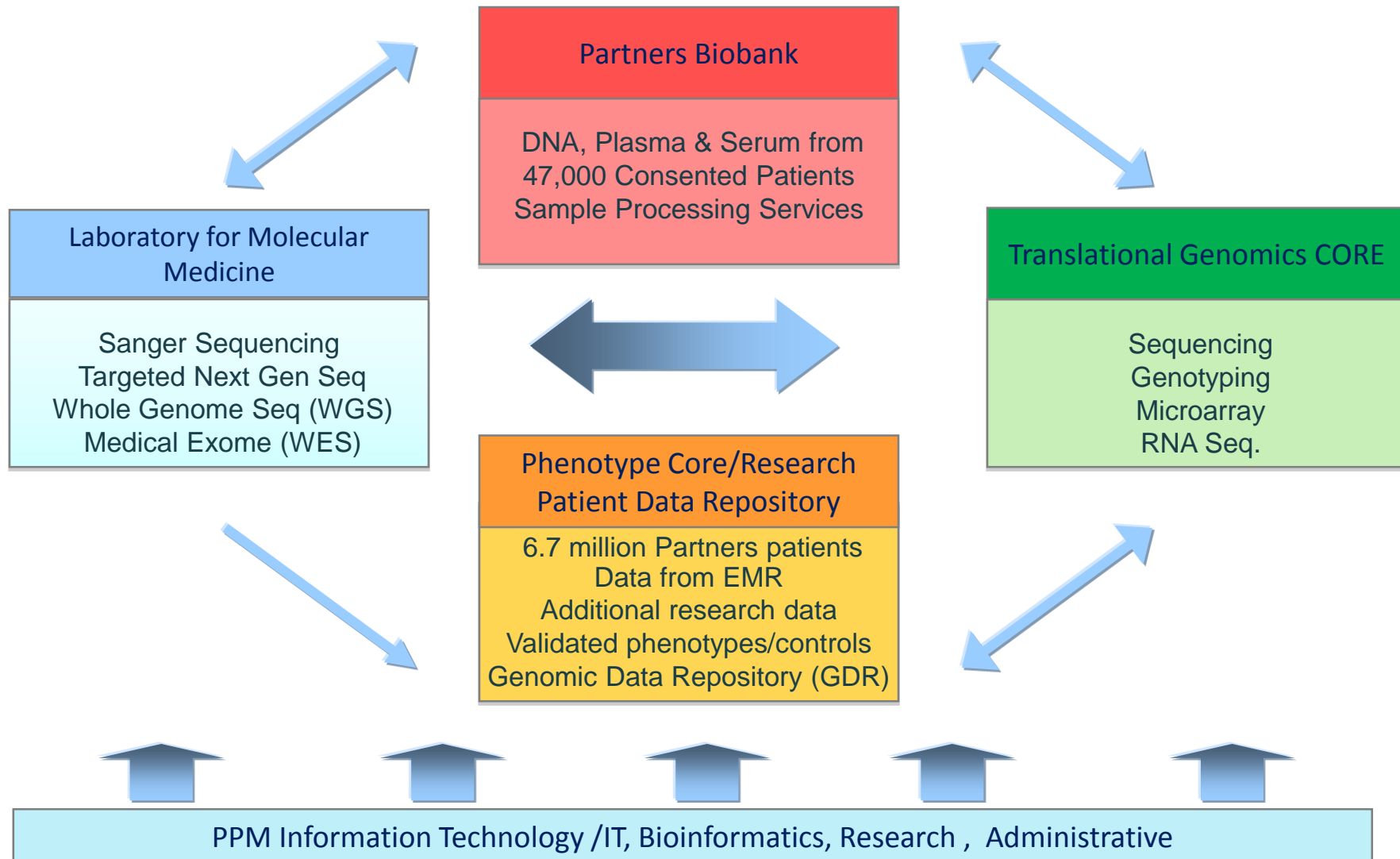
- The Partners Biobank provides samples (plasma, serum, and DNA) collected from consented patients.
- 64,000 patients have consented to date
- Samples are available for distribution to Partners investigators\* to help identify novel Personalized Medicine opportunities that reduce cost and provide better care

*\*with required approval from the Partners Institutional Review Board (IRB).*

# Biobank Integrative Genomics Strategy



# Partners Personalized Medicine Components



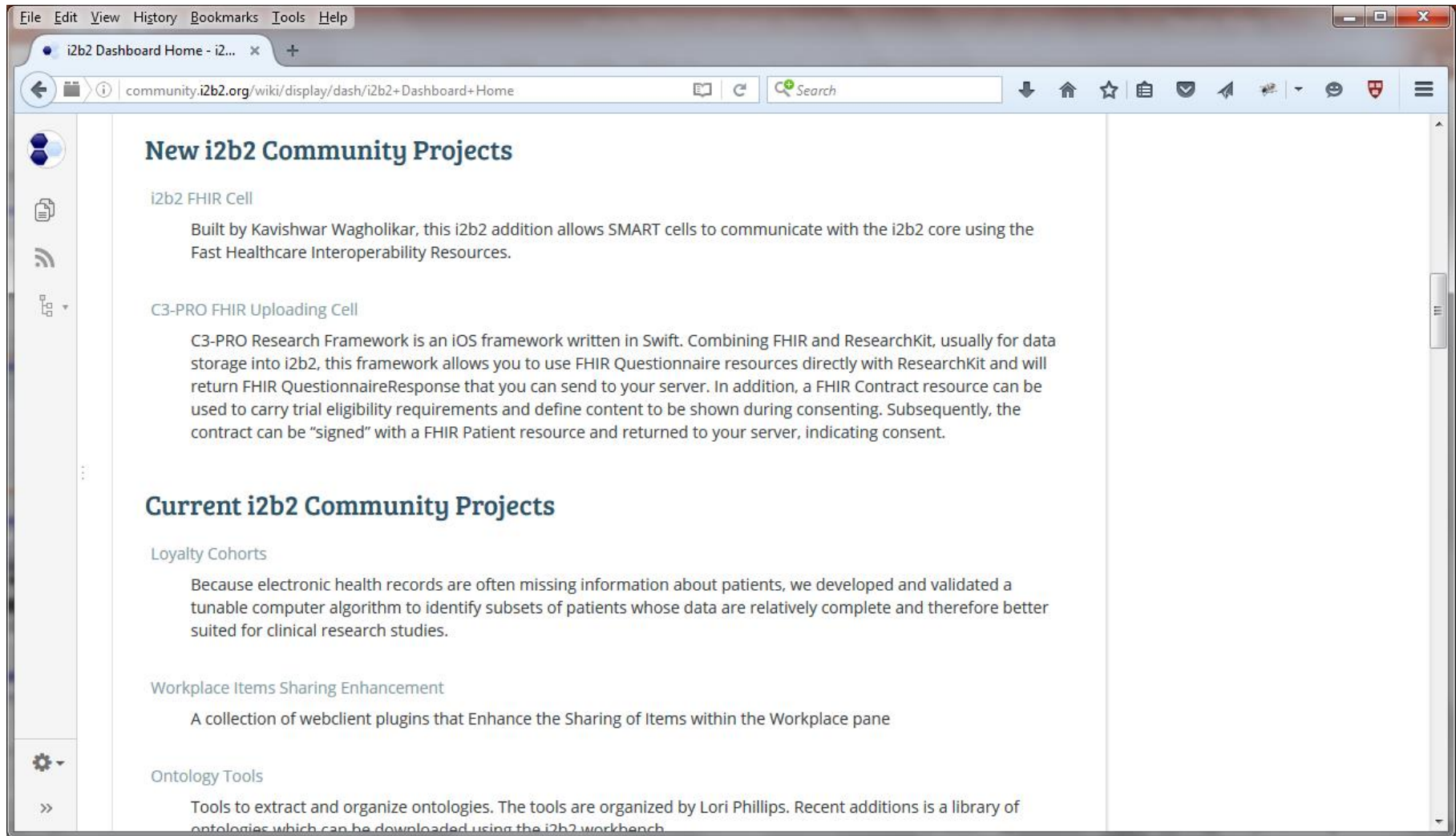
# High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion
- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data
- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing
- Personalized medicine is delivered into clinical care

## **RPDR Evolved into international “Informatics for Integrating Biology and the Bedside (i2b2)” sponsored by the National Institutes of Health, what is it?**

- Software for explicitly organizing and transforming person-oriented clinical data to a way that is optimized for clinical genomics research
  - Allows integration of clinical data, trials data, and genotypic data
- A portable and extensible application framework
  - Software is built in a modular pattern that allows additions without disturbing core parts
  - Available as open source at <https://www.i2b2.org>

# I2b2 Community Software Modules contributed as “Cells”



The screenshot shows a web browser window with the address bar displaying `community.i2b2.org/wiki/display/dash/i2b2+Dashboard+Home`. The page content is organized into two main sections:

## New i2b2 Community Projects

- i2b2 FHIR Cell**

Built by Kavishwar Wagholikar, this i2b2 addition allows SMART cells to communicate with the i2b2 core using the Fast Healthcare Interoperability Resources.
- C3-PRO FHIR Uploading Cell**

C3-PRO Research Framework is an iOS framework written in Swift. Combining FHIR and ResearchKit, usually for data storage into i2b2, this framework allows you to use FHIR Questionnaire resources directly with ResearchKit and will return FHIR QuestionnaireResponse that you can send to your server. In addition, a FHIR Contract resource can be used to carry trial eligibility requirements and define content to be shown during consenting. Subsequently, the contract can be “signed” with a FHIR Patient resource and returned to your server, indicating consent.

## Current i2b2 Community Projects

- Loyalty Cohorts**

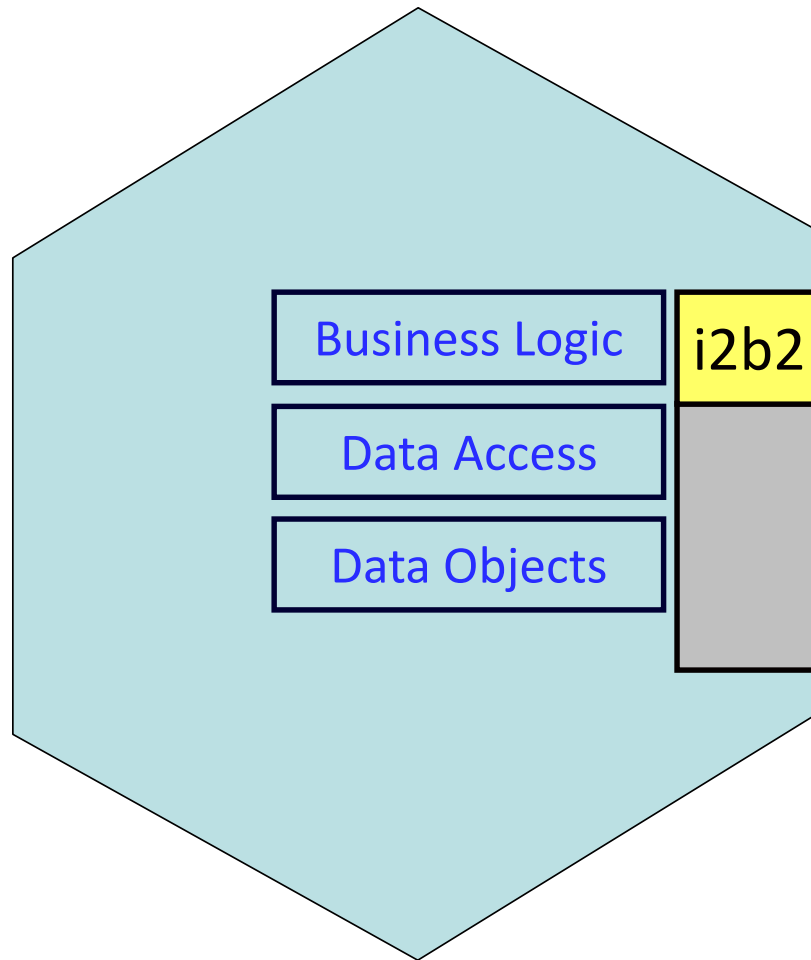
Because electronic health records are often missing information about patients, we developed and validated a tunable computer algorithm to identify subsets of patients whose data are relatively complete and therefore better suited for clinical research studies.
- Workplace Items Sharing Enhancement**

A collection of webclient plugins that Enhance the Sharing of Items within the Workplace pane
- Ontology Tools**

Tools to extract and organize ontologies. The tools are organized by Lori Phillips. Recent additions is a library of ontologies which can be downloaded using the i2b2 workbench



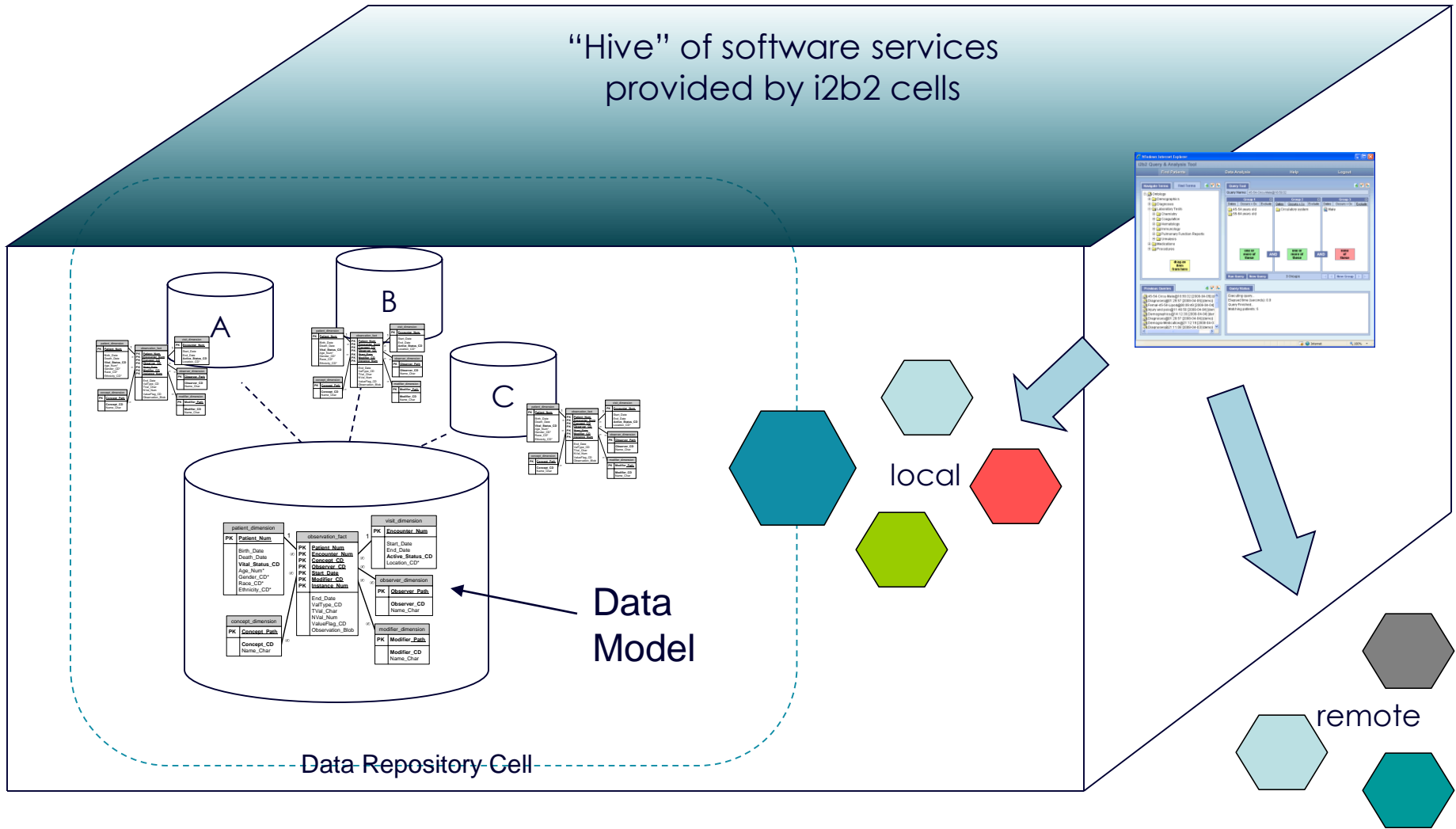
# i2b2 Cell: The Canonical Software Module



**HTTP XML**  
(minimum: RESTful)

# An i2b2 Environment (the Hive) is built from i2b2 Cells

“Hive” of software services provided by i2b2 cells



# I2b2 Software components are distributed as open source

The screenshot shows a web browser window titled "i2b2: Informatics for Integrating Biology & the Bedside - Windows Internet Explorer". The address bar shows the URL "https://www.i2b2.org/software/index.html". The page content includes a navigation menu on the left with links for "Archived Source Code", "Contributed", "Tutorial", "Guestbook \*", and "Statistics \*". The main content area features a "Documentation" section with the text "Hover over the modules below for the latest documentation:". Below this is a key for the components: "i2b2 Core Cell" (dark blue), "i2b2 Optional Cell" (light blue), "Workben" (green), "Web Client" (grey), and "CRC Plug" (cyan). The central diagram consists of a cluster of hexagonal components: "Project Management", "File Repository", "Natural Language Processing", "PFT Processing", "Ontology Management", "Identity Management", "Data Repository (CRC)", "Workflow Framework", "High Performance Computing Plug-in", "Correlation Analysis Plug-in", "Export Data Plug-in", "Table View Plug-in", "Text Analyzer Plug-in", "Import Data Plug-in", and "Annotator Plug-in". A tooltip for "i2b2 Workbench" is displayed over the "Workflow Framework" component, containing the text: "The i2b2 Workbench is a collection of client-side components designed as Eclipse-based java plug-ins that communicate with i2b2 Cells and allow the investigator to query, analyze, and display the data of the hive, generally in greater depth than the web client." The tooltip also lists links for "Installation Guide", "Tutorial Document", "Developer's Guide", "Go to Download Client", and "Go to Download Source".

# Implementations

## CTSA's

- Boston University
- Case Western Reserve University (*including Cleveland Clinic*)
- Children's National Medical Center (GWU), Washington D.C.
- Duke University
- Emory University (*including Morehouse School of Medicine and Georgia Tech*)
- Harvard University (*including Beth Israel Deaconess Medical Center, Brigham and Women's Hospital, Children's Hospital Boston, Dana Farber Cancer Center, Joslin Diabetes Center, Massachusetts General Hospital*)
- Medical University of South Carolina
- Medical College of Wisconsin
- Oregon Health & Science University
- Penn State Milton S. Hershey Medical Center
- Tufts University
- University of Alabama at Birmingham
- University of Arkansas for Medical Sciences
- University of California Davis
- University of California, Irvine
- University of California, Los Angeles\*
- University of California, San Diego\*
- University of California San Francisco
- University of Chicago
- University of Cincinnati (*including Cincinnati Children's Hospital Medical Center*)
- University of Colorado Denver (*including Children's Hospital Colorado*)
- University of Florida
- University of Kansas Medical Center
- University of Kentucky Research Foundation
- University of Massachusetts Medical School, Worcester
- University of Michigan
- University of Pennsylvania (*including Children's Hospital of Philadelphia*)
- University of Pittsburgh (*including their Cancer Institute*)
- University of Rochester School of Medicine and Dentistry
- University of Texas Health Sciences Center at Houston
- University of Texas Health Sciences Center at San Antonio
- University of Texas Medical Branch (Galveston)
- University of Texas Southwestern Medical Center at Dallas
- University of Utah
- University of Washington
- University of Wisconsin - Madison (*including Marshfield Clinic*)
- Virginia Commonwealth University
- Weill Cornell Medical College

## Academic Health Centers (does not include AHCs that are part of a CTSA):

- Arizona State University
- City of Hope, Los Angeles
- Georgia Health Sciences University, Augusta
- Hartford Hospital, CN
- HealthShare Montana
- Massachusetts Veterans Epidemiology Research and Information Center (MAVERICK), Boston
- Nemours
- Phoenix Children's Hospital
- Regenstrief Institute
- Thomas Jefferson University
- University of Connecticut Health Center
- University of Missouri School of Medicine
- University of Tennessee Health Sciences Center
- Wake Forest University Baptist Medical Center

## HMOs:

- Group Health Cooperative
- Kaiser Permanente

## International:

- Georges Pompidou Hospital, Paris, France
- Hospital of the Free University of Brussels, Belgium
- Inserm U936, Rennes, France
- Institute for Data Technology and Informatics (IDI), NTNU, Norway
- Institute for Molecular Medicine Finland (FIMM)
- Karolinska Institute, Sweden
- Landspítali University Hospital, Reykjavik, Iceland
- Tokyo Medical and Dental University, Japan
- University of Bordeaux Segalen, France
- University of Erlangen-Nuremberg, Germany
- University of Goettingen, Goettingen, Germany
- University of Leicester and Hospitals, England (Biomed. Res. Informatics Ctr. for Clin. Sci)
- University of Pavia, Pavia, Italy
- University of Seoul, Seoul, Korea

## Companies:

- Johnson and Johnson (TransMART)
- GE Healthcare Clinical Data Services

# Interrogation can occur through i2b2 web client

The screenshot displays the i2b2 Query & Analysis Tool web client interface within a Windows Internet Explorer browser window. The browser's address bar shows the URL: `http://phsi2b2appdev.mgh.harvard.edu/webclient/#`. The browser's menu bar includes File, Edit, View, Favorites, Tools, and Help. The browser's toolbar contains various icons for navigation and utility. The i2b2 interface itself has a header with the title "i2b2 Query & Analysis Tool" and navigation links for "Project: RA Mart Test", "User: Iori", "Find Patients", "Analysis Tools", "Message Log", and "Help | Logout".

The main interface is divided into several panels:

- Navigate Terms:** A tree view showing a hierarchy of terms: Visit, Clinic, Hospital, Brigham and Womens, Faulkner, Massachusetts General, and Newton Welleslev.
- Workplace:** A panel showing a file structure with folders "Icp5" and "SHARED".
- Previous Queries:** A list of recent queries, all starting with "Arterial vascul@" followed by a timestamp and "[Icp5]".
- Query Tool:** The central panel for building queries. It includes a "Query Name" field, a "Temporal Constraint" dropdown set to "Selected groups occur in the same financial encounter", and three query groups:
  - Group 1:** Contains the term "Brigham and Womens".
  - Group 2:** Contains the term "Malignant melanoma of sk".
  - Group 3:** Contains the term "Female".Each group has a dropdown menu for "Occurs in Same Encounter" and a "Dates" field. The groups are connected by "AND" operators. Below the groups, there are buttons for "Run Query", "Clear", "Print Query", and "New Group".
- Query Status:** A panel at the bottom right for displaying the results of the query.

The browser's status bar at the bottom shows a warning icon and the text "Done, but with errors on page." The system tray includes the Internet Explorer icon and a zoom level of 100%.

# I2b2 Workbench provides a detailed patient view for Investigator

The screenshot displays the 'Patient Recruitment Workbench' interface. At the top, the user is identified as 'Shawn Murphy' with a 'Status' indicator and 'Log out' and 'Help' buttons. The interface is divided into several panels:

- Concept trees:** A hierarchical list of medical concepts on the left, including Demographics, Diagnoses, and Respiratory system. The 'Respiratory system' is expanded to show 'Acute respiratory infections', 'Chronic obstructive diseases', and 'Asthma'.
- Group Selection:** Three panels labeled 'Group 1', 'Group 2', and 'Group 3'. Group 1 contains 'Bronchodilators', Group 2 contains 'Prednisolone' and 'Prednisone', and Group 3 is empty. 'and' connectors are placed between the groups. Buttons for 'Reset' and 'Run Query' are at the bottom.
- Timeline:** A central panel showing a timeline for five patients. The patients listed are:
  - Person #82032\_Female\_68yroid\_White
  - Person #82098\_Female\_37yroid\_White
  - Person #82202\_Male\_Deaf\_Black
  - Person #82229\_Female\_63yroid\_WhiteEach patient entry shows a vertical bar representing a time period, with colored segments indicating events for 'Bronchodilators' (blue) and 'Prednisolone-Pr...' (green).

Below the timeline, there are controls for 'Patient Set: Patient Set - 3894 Patients', 'start: 0', and 'Increment: 10'.

On the right side, there are several informational panels:

- Help:** Contains 'Related Topics' and 'Dynamic Help' sections.
- Patient Schedule:** Displays details for 'John Smith', including dates and times for visits (e.g., 10/15/2006 11:00 am - 12:00 am) and search results for 'Dematology'.
- Report Diagnosis Detail Image:** Shows two medical images: a top-down view of a brain scan with orange highlights, and a sagittal view of a brain scan.

At the bottom, a workflow diagram titled 'Drop Patient Object' shows a sequence of steps: 'Drop Patient Object' -> 'ObjectToRecord' -> 'RecordDisassembler' -> 'Synchronizer' -> 'Display' and 'Display2'. The author is listed as 'Tobin T. Fricke'.

# Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model

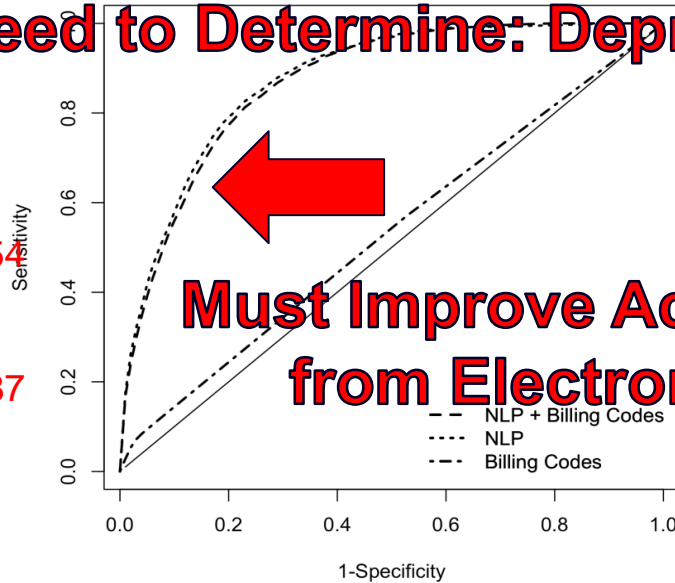
R. H. Perlis<sup>1,2\*</sup>, D. V. Iosifescu<sup>1,3</sup>, V. M. Castro<sup>4</sup>, S. N. Murphy<sup>5</sup>, V. S. Gainer<sup>4</sup>, J. Minnier<sup>6</sup>, T. Cai<sup>6</sup>,  
S. Goryachev<sup>4</sup>, Q. Zeng<sup>7</sup>, P. J. Gallagher<sup>2</sup>, M. Fava<sup>1</sup>, J. B. Weillburg<sup>1</sup>, S. E. Churchill<sup>8</sup>,  
I. S. Kohane<sup>9</sup> and J. W. Smoller<sup>2</sup>

## Use Phenotyping Algorithms to define cohorts of treatment-resistant and treatment-responsive depression

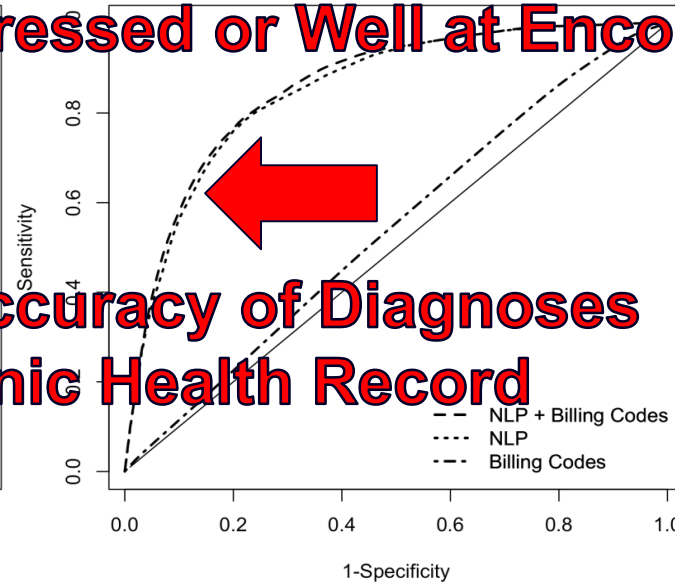
**Need to Determine: Depressed or Well at Encounter**

Initially:  
AUC = 0.54

Finally:  
AUC = 0.87



**Must Improve Accuracy of Diagnoses from Electronic Health Record**

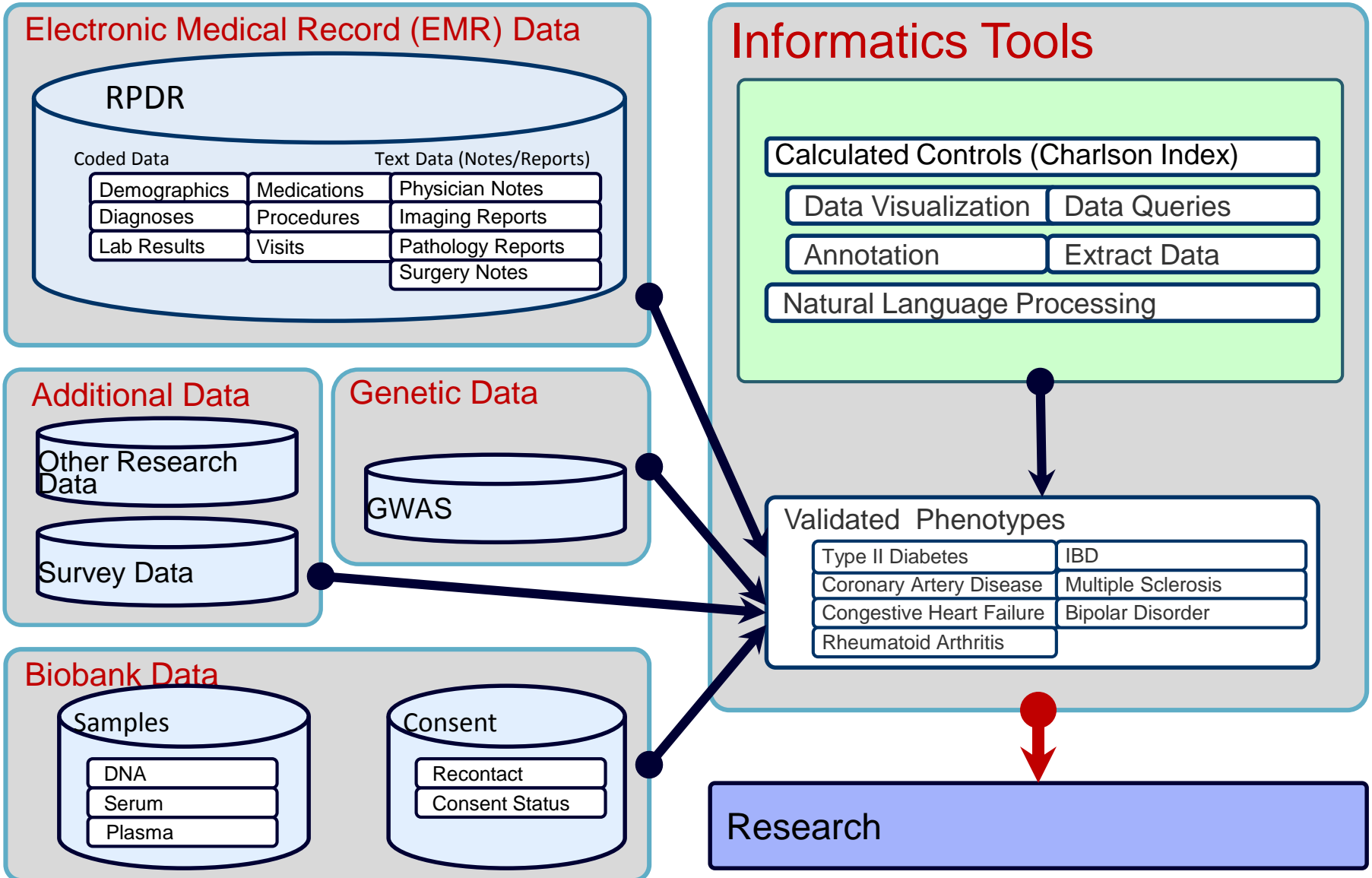


Initially:  
AUC = 0.55

Finally:  
AUC = 0.86

Clinical Status	Model	Specificity	Sensitivity	Precision	AUC
Depressed	Billing Codes	0.95	0.09 (0.03)	0.57 (0.14)	0.54 (0.02)
Depressed	NLP	0.95	0.42 (0.05)	0.78 (0.02)	0.88 (0.02)
Depressed	NLP + Billing Codes	0.95	0.39 (0.06)	0.78 (0.02)	0.87 (0.02)
Well	Billing Codes	0.95	0.06 (0.02)	0.26 (0.27)	0.55 (0.03)
Well	NLP	0.95	0.37 (0.06)	0.86 (0.02)	0.85 (0.02)
Well	NLP + Billing Codes	0.95	0.39 (0.07)	0.85 (0.02)	0.86 (0.02)

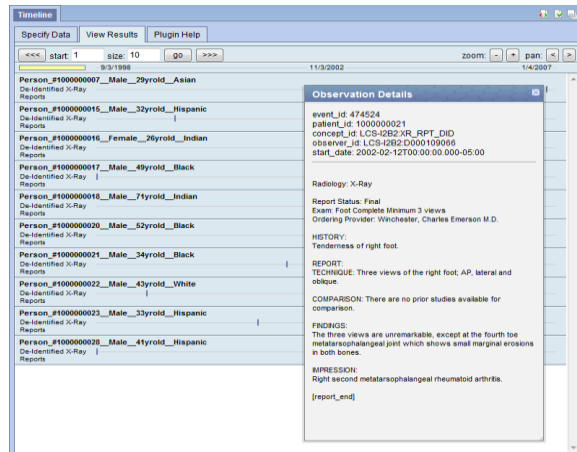
# Data Integration | Phenotype Discovery Center



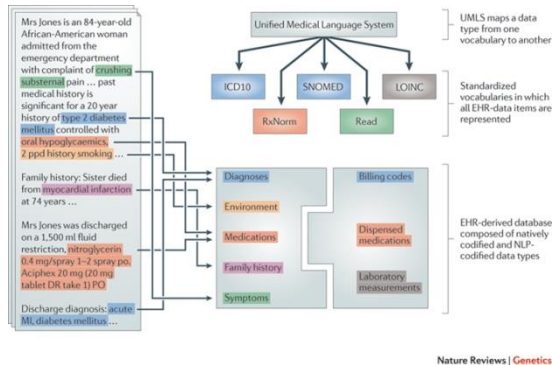


# Curating a Disease Algorithm

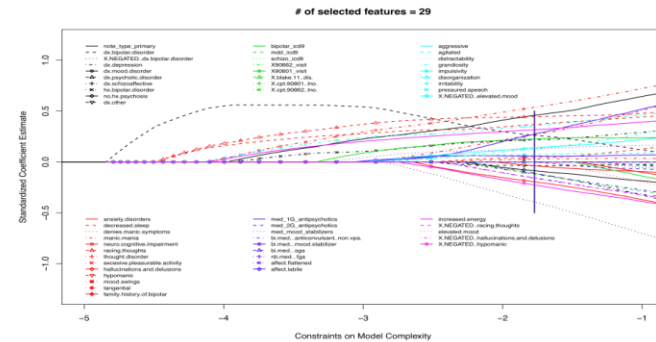
## 1. Create a gold standard training set.



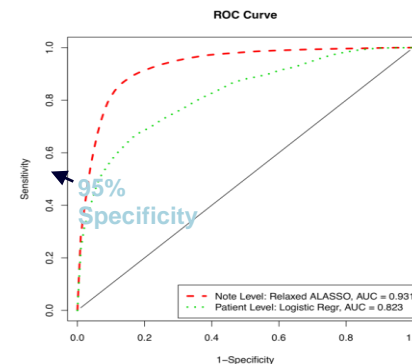
## 2. Create a comprehensive list of features from patient's electronic data that describe the disease of interest



## 3. Develop the classification algorithm. Using the data analysis file and the training set from step 1, assess the frequency of each variable. Remove variables with low prevalence. Apply adaptive LASSO penalized logistic regression to identify highly predictive variables for the algorithm

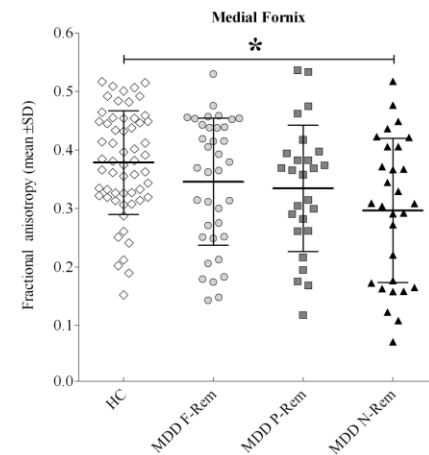
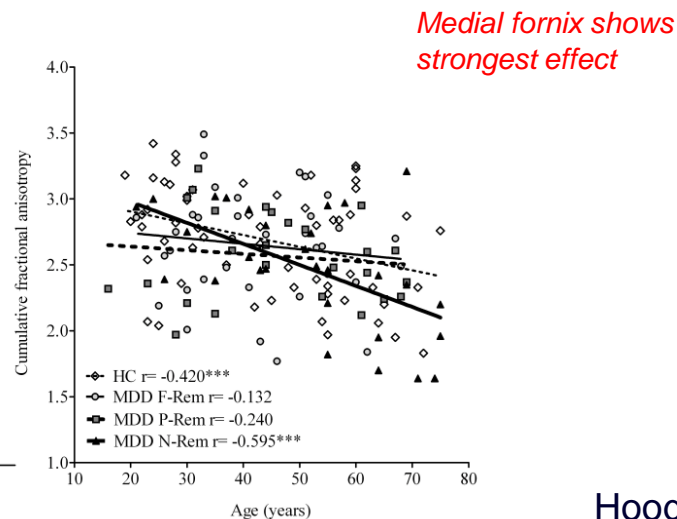
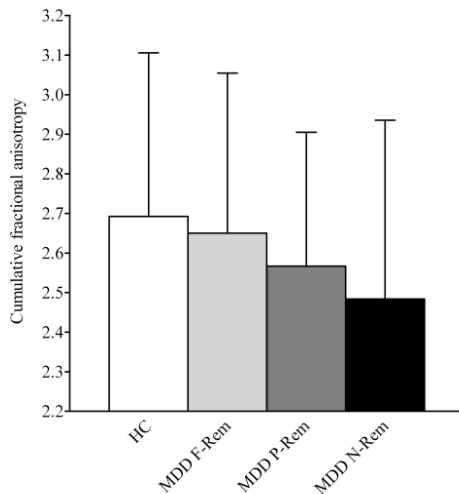
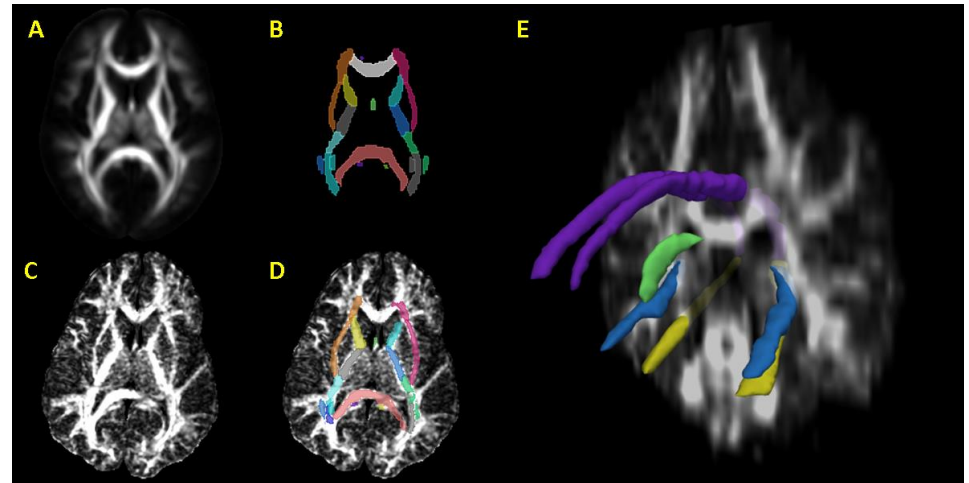


## 4. Apply the algorithm to all subjects in the superset and assign each subject a probability of having the phenotype



# White matter abnormalities associated with treatment-resistant depression

- Scans collected as part of routine clinical care
- Diffusion tensor imaging in 150 pts
- Age-related decline in white matter integrity increases with treatment resistant depression



# Biobank Portal | Curated Diseases

Validated Phenotype	Count*	Predictive Positive Value
Bipolar Disease	71	89%
Congestive Heart Failure	387	90%
Coronary Artery Disease	2,420	97%
Crohn's Disease	453	90%
Multiple Sclerosis	94	90%
Rheumatoid Arthritis	550	90%
Type 2 Diabetes Mellitus	1,887	97%
Ulcerative Colitis	330	90%

Healthy Controls based on Charlson Index	Count**
0 – 10-year survival probability is >98.3%	2,206
1 – 10-year survival probability is >95.87%	4,343
2 – 10-year survival probability is >90.15%	6,545

\* Based on 15,880 patients

\*\* Based on 21,300 patients

# High Quality Phenotypes available for Genetic Studies

The screenshot displays the Partners Biobank Portal interface. On the left is a navigation tree with categories like Biobank Consent Information, Demographics, Genomics, Health Information Survey, Sample Types, and Curated Disease Populations. The main area features a 'Query Tool' with a query name 'Prima-CHF --Gene@14:22:40' and a temporal constraint of 'Treat all groups independently'. The query is structured into three groups, each with a table of criteria:

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
			Primary dilated cardiomyopathy - 4002	CHF - current or past history (PPV 0.90) - 700		Gene [contains "TTN AND Homozygous AND (Frameshift OR missense OR nonsense OR start_loss OR stop_loss)"]		

Each group is linked by 'AND' operators, and each group's criteria are linked by 'one or more of these' operators. Below the query tool, there are buttons for 'Run Query', 'Clear', and 'New Group'. The results section shows '3 Groups' and includes buttons for 'Show Query Status', 'Graph Results', 'Query Report', and 'Download Results'. A large blue box displays the result: 'Number of patients: 70 For Query "Prima-CHF --Gene@14:22:40"'. The user's name 'Shawn Murphy, MD' is visible in the top right corner.

# Genotype Data

Search by Gene

Use the gene name box to specify the variant for which to search. When you begin typing in the search box below, a selection list will appear after you type the first characters.

Gene Name\*: APOJ

Please note the zygosity to query for patients without a particular variant (frameshift, missense, nonsense, start loss, stop loss)

Zygosity\*: APOA1, APOA1BP, APOA2, APOA4, APOA5

Consequence: APOB, APOBEC1, APOBEC2, APOBEC3A, APOBEC3A\_B, APOBEC3B

drop a term on here

Search by dbSNP rs Identifier

Use the rs identifier box to specify the variant for which to search. When you begin typing in the search box below, a selection list will appear after you type the first three numbers.

rs identifier\*: rs1234

Please note the zygosity to query for patients without a particular variant (frameshift, missense, nonsense, start loss, stop loss) after you start typing (nucleotide on the right) to query for patients without a particular variant (frameshift, missense, nonsense, start loss, stop loss) identifier that appears in the box on the left) to the alternate

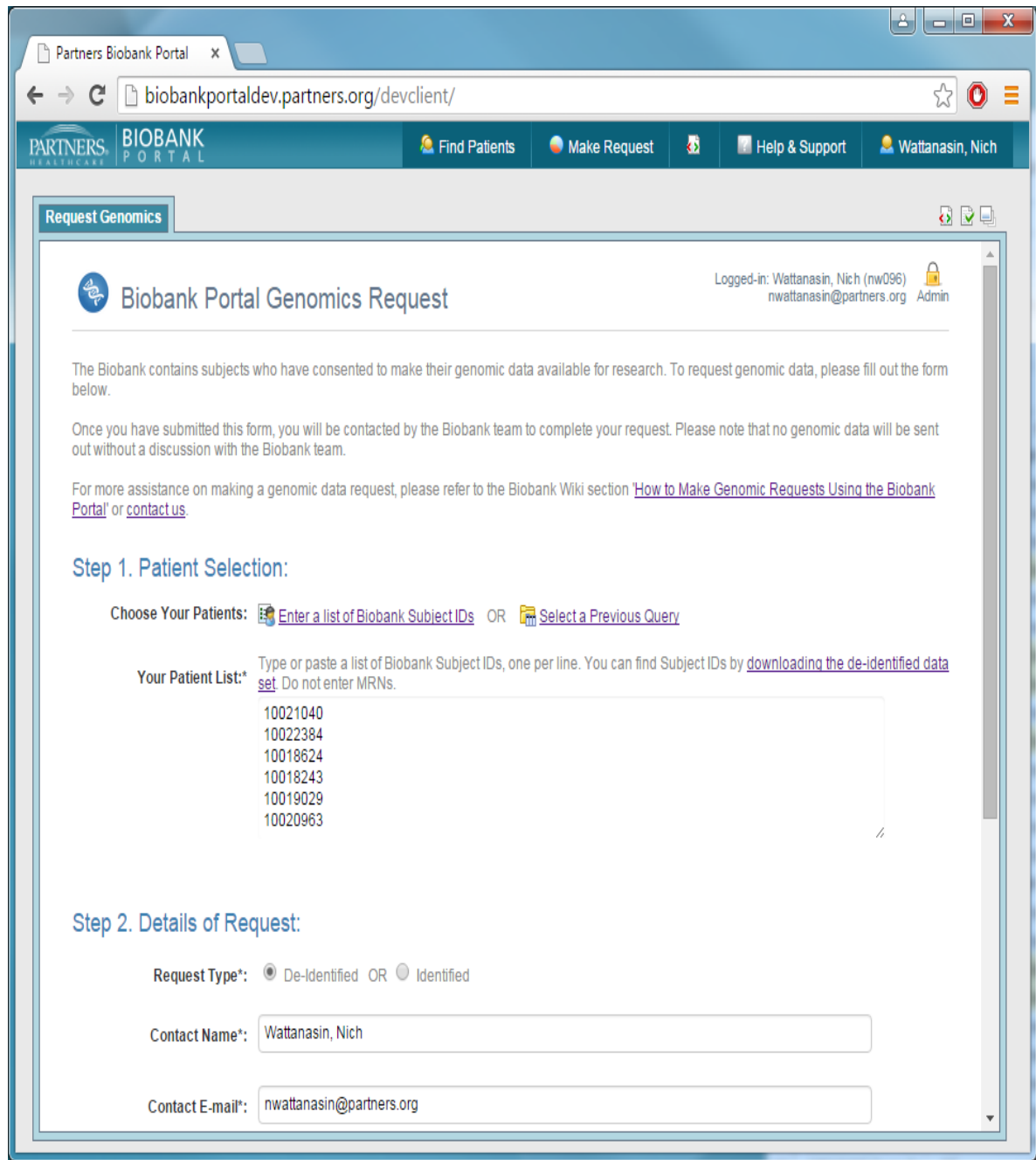
Zygosity\*: rs12340033 | C to G, rs12340061 | G to A, rs12340067 | C to T, rs12340088 | T to G, rs12340105 | A to C, rs12340107 | G to T, rs12340117 | G to A, rs12340120 | G to A, rs12340129 | A to G, rs12340149 | G to A, rs12340158 | T to A

drop a term on here

TTN

- 3349 SNP or indels
- 1680 Homozygous
- 1336 subjects with protein altering (frameshift, missense, nonsense, start loss, stop loss) variant

# Partners Biobank Portal – Request Genetic Data



Partners Biobank Portal

biobankportaldev.partners.org/devclient/

Find Patients Make Request Help & Support Wattanasin, Nich

Request Genomics

Logged-in: Wattanasin, Nich (nw096)  
nwattanasin@partners.org Admin

## Biobank Portal Genomics Request

The Biobank contains subjects who have consented to make their genomic data available for research. To request genomic data, please fill out the form below.

Once you have submitted this form, you will be contacted by the Biobank team to complete your request. Please note that no genomic data will be sent out without a discussion with the Biobank team.

For more assistance on making a genomic data request, please refer to the Biobank Wiki section ['How to Make Genomic Requests Using the Biobank Portal'](#) or [contact us](#).

### Step 1. Patient Selection:

Choose Your Patients:  OR

Your Patient List\* Type or paste a list of Biobank Subject IDs, one per line. You can find Subject IDs by [downloading the de-identified data set](#). Do not enter MRNs.

10021040  
10022384  
10018624  
10018243  
10019029  
10020963

### Step 2. Details of Request:

Request Type\*:  De-Identified OR  Identified

Contact Name\*:

Contact E-mail\*:

# High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion
- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data
- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing
- Personalized medicine is delivered into clinical care

# emerge network

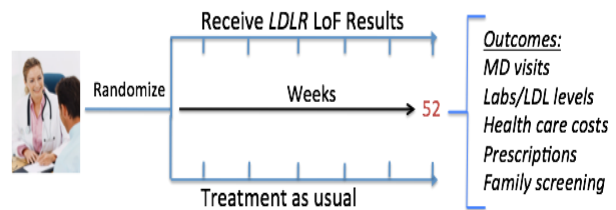
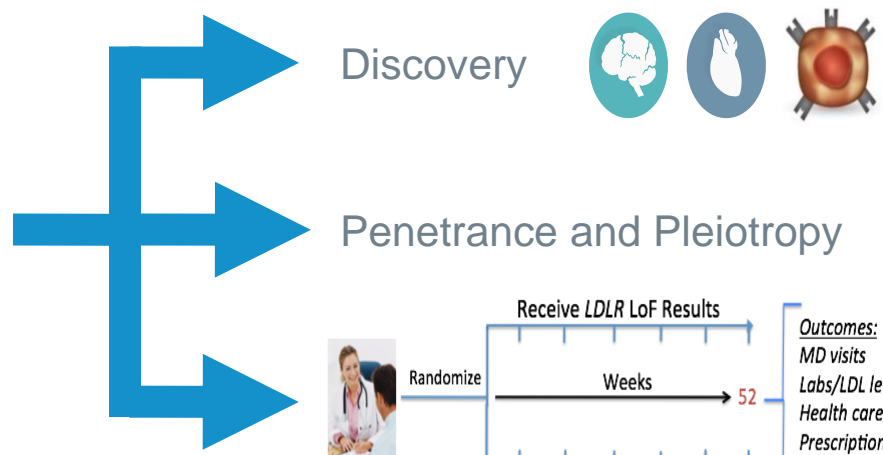
ELECTRONIC MEDICAL RECORDS AND GENOMICS



Sequencing  
100 high-  
priority genes



25,000 Network-wide



- MD visits
- Labs/LDL levels
- Health care costs
- Prescriptions
- Family screening



# Federated Queries in PCORNet

**SHRINE** Project: SHRINE User: Shawn Murphy Find Patients | Message Log | Help | Logout

**Query Tool**

Query Name: Acute hepatitis@12:45:50

Search by Names Search by Codes

Containing hepatitis c

Find Any Category

- Hepatitis b core antibody measurement
- Hepatitis b core igm antibody test
- Hepatitis b surface antigen (hbsag) carrier
- Hepatitis b vaccine injection administered or previous
- Hepatitis b virus (hbv) status assessed and results i
- Hepatitis c antibody
- Hepatitis c antibody
- Hepatitis c antibody confirmatory test
- Hepatitis c antibody confirmatory test (eg, immunobi
- Hepatitis c antibody measurement
- Hepatitis c antibody test

Previous Queries

- Arthropathies a@17:23:51 [2-11-2015] [smurphy]
- Nephritis, neph@15:56:29 [2-11-2015] [smurphy]
- Major-Femal-18-34@17:36:09 [2-4-2015] [smurphy]
- Major d-Female@17:34:30 [2-4-2015] [smurphy]
- Major depressio@17:32:43 [2-4-2015] [smurphy]
- Appendi-Black o@15:57:24 [1-28-2015] [smurphy]
- Appendicitis@15:55:07 [1-28-2015] [smurphy]

Query Status

Finished Query: "Acute hepatitis@12:45:50" [63.5 secs]

Temple - 10 patients or fewer	FINISHED [2.0 secs]
Wake - 4274 ±3 patients	FINISHED [3.0 secs]
BCH - 501 ±3 patients	FINISHED [4.1 secs]
BMC - 8922 ±3 patients	FINISHED [4.6 secs]
CC-HMC - 212 ±3 patients	FINISHED [8.0 secs]
Partners HealthCare - 18750 ±3 patients	FINISHED [6.1 secs]
UT - 3516 ±3 patients	FINISHED [26.8 secs]

- Partners HealthCare System
- Boston Children's Hospital
- BIDMC
- Boston Health Net (BMC and Community Health Centers)

University of California, Davis

Columbia U. Medical Center and New York Presbyterian Hospital

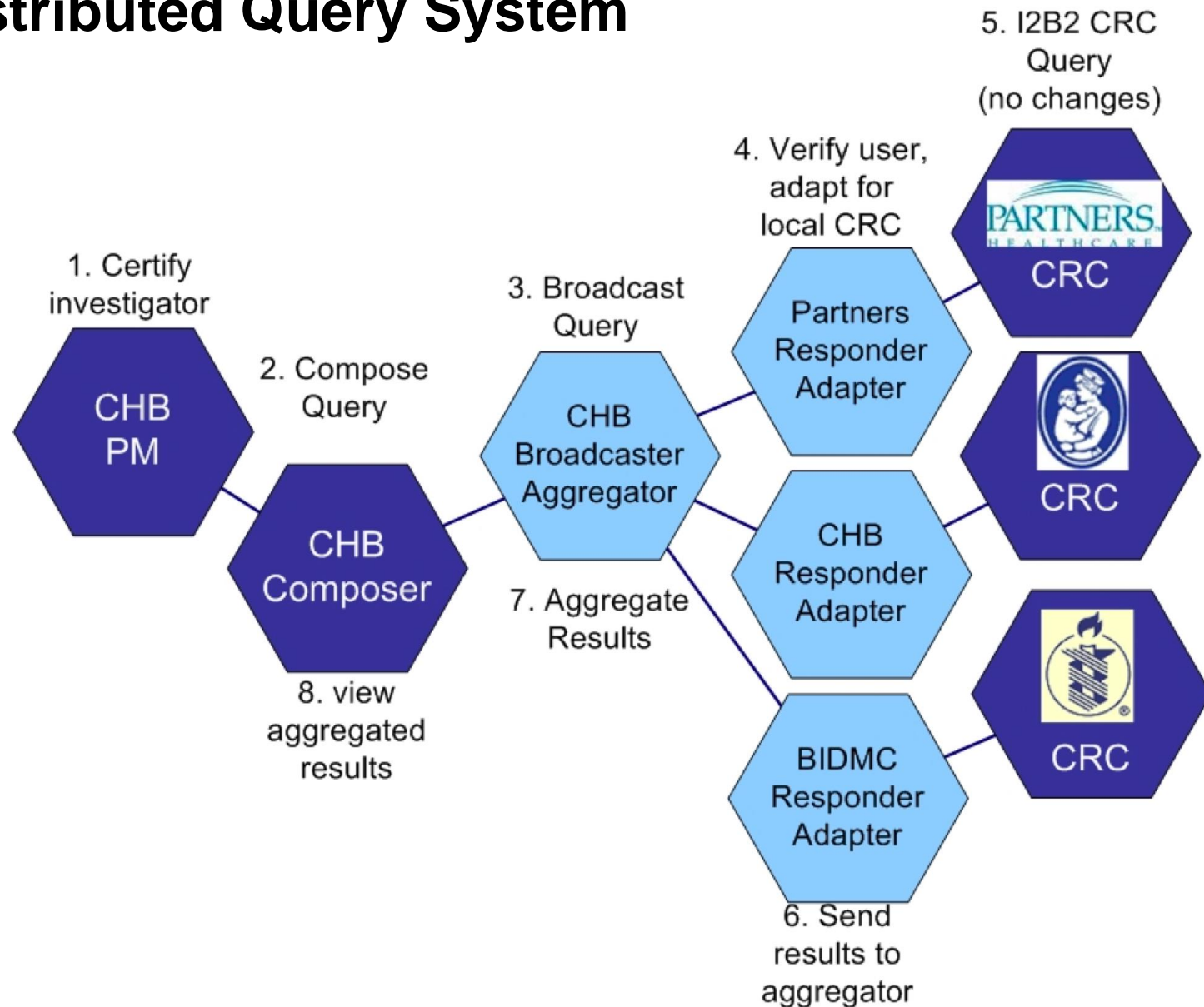
Washington University in St. Louis

Wake Forest Baptist Medical Center

Morehouse/Grady/RCMI

U Texas Health Science Center/Houston

# Distributed Query System



# Run Query Using SCILHS-SHRINE

**SHRINE**
Project: SHRINE
User: Shawn Murphy
Find Patients | Message Log | Help | Logout

Navigate Terms
Find Terms

Search by Names

Search by Codes

Containing

Find Any Category

- Hepatitis b core antibody measurement
- Hepatitis b core igm antibody test
- Hepatitis b surface antigen [hbsag] carrier
- Hepatitis b vaccine injection administered or previous
- Hepatitis b virus (hbv) status assessed and results i
- Hepatitis c antibody**
- Hepatitis c antibody
- Hepatitis c antibody
- Hepatitis c antibody confirmatory test
- Hepatitis c antibody; confirmatory test (eg, immunobl
- Hepatitis c antibody measurement
- Hepatitis c antibody test

Query Tool

Query Name:

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
<input type="checkbox"/> Acute hepatitis c with hepatic co <input type="checkbox"/> Acute hepatitis c without mentio <input type="checkbox"/> Unspecified viral hepatitis c <input type="checkbox"/> Unspecified viral hepatitis <input type="checkbox"/> Unspecified viral hepatitis			<b>AND</b>			<div style="background-color: yellow; padding: 2px; display: inline-block;">drop a term on here</div>		
<input type="checkbox"/> Unspec								

one or more of these

AND

Run Query
Clear
Print Query

1 Group
New Group

Previous Queries

- Arthropathies a@17:23:51 [2-11-2015] [smurphy]
- Nephritis, neph@15:56:29 [2-11-2015] [smurphy]
- Major-Femal-18-34@17:36:09 [2-4-2015] [smurphy]
- Major d-Female@17:34:30 [2-4-2015] [smurphy]
- Major depressiv@17:32:43 [2-4-2015] [smurphy]
- Appendi-Black o@15:57:24 [1-28-2015] [smurphy]
- Appendicitis@15:55:07 [1-28-2015] [smurphy]

Query Status

**Finished Query: "Acute hepatitis@12:45:50"** [63.5 secs]

Temple - 10 patients or fewer	FINISHED [2.0 secs]
Wake - 4274 ±3 patients	FINISHED [3.0 secs]
BCH - 501 ±3 patients	FINISHED [4.1 secs]
BMC - 8922 ±3 patients	FINISHED [4.6 secs]
CCHMC - 212 ±3 patients	FINISHED [8.0 secs]
Partners HealthCare - 18750 ±3 patients	FINISHED [6.1 secs]
UT - 3516 ±3 patients	FINISHED [26.8 secs]

# Workflow at the sites to find patients for a clinical trial:

- After a query is run across the “SHRINE” network, the query is automatically saved at every site
- The query saved at each site is transformed into a patient set
- The patient set is studied and sorted for the specific patients eligible for the Clinical Trial

# Review Patients at Sites

i2b2 Workbench

File Window Help

i2b2 Workbench for i2b2 Demo (Oracle) i2b2 User Status: ● i2b2

Navigate Term Find Terms Timeline View Analysis View Patient Mapping View CT Viewer Workplace

Select Patients Select Concepts Render Tables

Search by Names Search by Codes

Containing reductase inhibitors

Find Any Category

HMG-CoA reductase inhibitors

- Atorvastatin
- Cerivastatin
- Fluvastatin
- Lovastatin
- Lovastatin-niacin
- Pravastatin
- Rosuvastatin
- Simvastatin

Previous Query Patient Sets

Search By Name

Containing

Find Any Category

Patient Set for "Asthma-Albuter@01:06:16"

- 1000000001 [24 y/o F black]
- 1000000002 [44 y/o F white]
- 1000000003 [41 y/o M asian]
- 1000000004 [34 y/o M black]
- 1000000005 [37 y/o F hispanic]
- 1000000010 [33 y/o F hispanic]

Be... 02-14-2014 0:45:29

Set #	Patient Set Name
1	Patient Set for "Asthma-Albuter@01:06:16"

SMART	Patient ID	PSet #	Patient Name	Gender	Race	Date of Birth	Age	Obesity	Acute Myocardial..	HMG-CoA reduc
✖	1000000001	1+1	xxxxx, xxxxx	F	black	1985-11-17T...	21	✓	✓	✓
✖	1000000002	1+1	xxxxx, xxxxx	F	white	1966-08-29T...	40	✗	✓	✗
✖	1000000003	1	xxxxx, xxxxx	M	asian	1969-03-07T...	38	✓	✗	✗
✖	1000000004	1	xxxxx, xxxxx	M	black	1976-08-13T...	30	✓	✓	✓
✖	1000000005	1	xxxxx, xxxxx	F	hispanic	1973-06-25T...	33	✓	✓	✓
✖	1000000006	1	xxxxx, xxxxx	F	black	1981-08-05T...	25	✓	✗	✓
✖	1000000007	1	xxxxx, xxxxx	M	asian	1981-07-10T...	25	✓	✓	✗
✖	1000000008	1	xxxxx, xxxxx	M	hispanic	1986-12-16T...	20	✓	✓	✓
✖	1000000009	1	xxxxx, xxxxx	F	hispanic	1989-01-11T...	18	✗	✓	✓
✖	1000000010	1	xxxxx, xxxxx	F	hispanic	1976-11-23T...	30	✗	✓	✗
✖	1000000011	1	xxxxx, xxxxx	F	white	2049-10-19T...	57	✓	✗	✓
✖	1000000012	1	xxxxx, xxxxx	F	black	1991-03-29T...	16	✓	✗	✓
✖	1000000013	1	xxxxx, xxxxx	F	black	2027-02-27T...	80	✗	✗	✓
✖	1000000014	1	xxxxx, xxxxx	M	white	1956-05-22T...	50	✗	✗	✓
✖	1000000015	1	xxxxx, xxxxx	M	hispanic	1978-09-11T...	28	✓	✗	✓
✖	1000000016	1	xxxxx, xxxxx	F	indian	1983-11-19T...	23	✓	✓	✓

SHARED

- Clinical Trial 'Asthma'
  - Maybe
  - NO
  - YES
- demo

# High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion
- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data
- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing
- **Personalized medicine is delivered into clinical care**

# Personalized Medicine for the Everyday Clinician - Finding similar patients

- Finding similar patients help us understand what is disease and what is normal, to distinguish between several disease states, help predict successful therapies, and to help determine next steps in potentially very expensive diagnostic pathways
- This is an opportunity for combining the EHR, Big Data Queries, and SMART Apps
- Apply the approaches we have used to conduct scientific research to Provider and Patient engaging visualizations

# Designing the App Store for Health



# SMART

## Bloodwork Cardiology Result

### Patient info

NAME: John Doe

GENDER: M AGE: 49 DOB: 01/10/1961

ORDERED BY: Dr. Francis Pulaski

Bellevue Medical Centre  
famar.d@bmcmed.edu  
(602) 555-54321 x1523

COLLECTED: 11/02/2010, 10:40 a.m.

RECEIVED: 11/02/2010, 1:03 p.m.

### 1 About this test

This report evaluates your potential risk of heart disease, heart attack, and stroke.

### 2 Your results

#### CRP level test



#### Total cholesterol level



#### LDL "bad" cholesterol



#### HDL "good" cholesterol



### 3 Your risk

You show an elevated risk of cardiovascular disease

If you're a smoker with normal blood pressure (130 mm/Hg) but family history of heart attack before age 60 (one or both parents) your risk over 10 years is:

**15%**

Your risk would be lowered to

12% if your blood pressure were 120mm/Hg

10% if you quit smoking

6% if you reduced cholesterol to 160mg/DL

Use your CRP results and cholesterol level to calculate your 10 risk of a cardiovascular event at ReynoldsRisk.com

Factor in or other s that can sterid

Consider retesting in 1 to 2 weeks to exclude a temporary spike in blood levels

information@beautiful.net



## The NEW ENGLAND JOURNAL of MEDICINE

### No Small Change for the Health Information Economy

Kenneth D. Mandl, M.D., M.P.H., and Isaac S. Kohane, M.D., Ph.D.

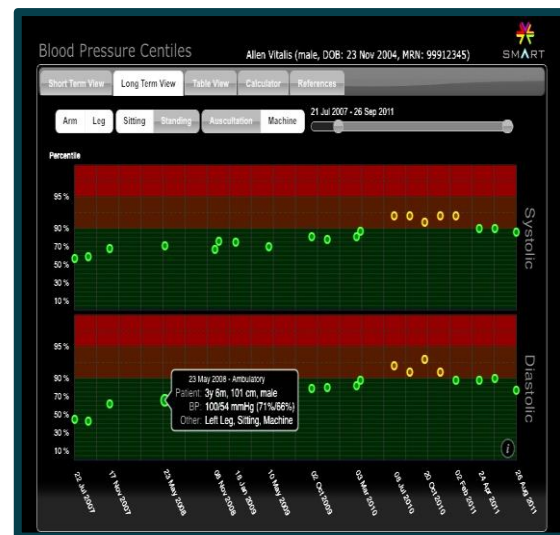
The economic stimulus package signed by President Barack Obama on February 17 included a \$19 billion investment in health information technology. How can we best take advantage of this unprecedented opportunity to computerize health care and stimulate the health information economy while also stimulating the U.S. economy? A health care system adapting to the effects of an aging population, growing expenditures, and a diminishing primary care workforce needs the support

of a flexible information infrastructure that facilitates innovation in wellness, health care, and public health.

Flexibility is critical, since the system will have to function under new policies and in the service of new health care delivery mechanisms, and it will need to incorporate emerging information technologies on an ongoing basis. As we seek to design a system that will constantly evolve and encourage innovation, we can glean lessons from large-scale information-

technology successes in other fields. An essential first lesson is that ideally, system components should be not only interoperable but also substitutable.

The Apple iPhone, for example, uses a software platform with a published interface that allows software developers outside Apple to create applications; there are now nearly 10,000 applications that consumers can download and use with the common phone interface. The platform separates the system from the functional-





# State-of-the-Art ???

<input checked="" type="checkbox"/> ALIGN HERE		SEND TO:					
PATIENT NAME DOE, JOE			PATIENT ID NO. NOT GIVEN		DATE COLLECTED	TIME	
ACCESSION NO. 36904447	BIRTH DATE 55	GENDER MALE	SAMPLE ID NO. NOT GIVEN	OTHER ID NO.	RECEIVED 06/11/2010	09:41	
REMARKS SAMPLE REPORT, NO SAMPLE SENT				REFERRING PHYSICIAN	REPORTED 06/11/2010	10:00	
					STATUS FINAL		
TEST	RESULT (* = OUT OF RANGE)		UNITS	REFERENCE RANGE			
Cardio CRP			0.4 mg/L				
For Ages > 17 Years:							
CCRP mg/L	Risk According to AHA/CDC Guidelines						
<1.0	Lower Relative Cardiovascular Risk.						
1.0-3.0	Average Relative Cardiovascular Risk						
3.1-10.0	Higher Relative Cardiovascular Risk. Consider retesting in 1 to 2 weeks to exclude a benign transient elevation in the baseline CRP value secondary to infection or inflammation.						
>10.0	Persistent elevations upon retesting, may be associated with infection and inflammation.						

An  
**Inspired Design**  
 from Dave McCandless  
 (cc license)

## Bloodwork Cardiology Result



ORDERED BY: Dr. Francis Pulaski  
 Bellevue Medical Centre  
 lamar.d@bactamed.edu  
 (603) 555-54321 x1523  
 COLLECTED: 11/02/2010, 10:40 a.m.  
 RECEIVED: 11/02/2010, 1:03 p.m.

### Patient info

NAME: **John Doe**  
 GENDER: **M** AGE: **49** DOB: **01/10/1961**

### 1 About this test

This report evaluates your potential risk of heart disease, heart attack, and stroke.

### 2 Your results

#### CRP level test



#### Total cholesterol level



#### LDL "bad" cholesterol



#### HDL "good" cholesterol



### 3 Your risk You show an elevated risk of cardiovascular disease

If you're a smoker with normal blood pressure, (130 mm/Hg) but family history of heart attack before age 60 (one or both parents) your risk over 10 years is:

**15%**

Your risk would be lowered to

**12%** if your blood pressure were 120mm/Hg  
**10%** if you quit smoking  
**6%** if you reduced cholesterol to 160mg/DL

Use your CRP results and cholesterol level to calculate your 10 risk of a cardiovascular event at [ReynoldsRisk.org](http://ReynoldsRisk.org)

### 4 What now?



**Diet & exercise-** can improve your cholesterol levels



**Quitting smoking-** can decrease your heart disease risk by 50% or more



**Ask your doctor** about statins or other medications that can lower cholesterol



**Consider retesting** in 1 to 2 weeks to exclude a temporary spike in blood levels

# Out of the Box - SMART Apps can link Big Data to the EMR

- Substitutable Medical Application and Reusable Technology – Started with grant from the Office of the National Coordinator
- Paradigm is similar to Mobile Apps with a proposed standard interface using FHIR (Fast Healthcare Interoperable Resource)

**Blood Pressure Centiles SMART Application**

Short Term View | Long Term View | Table View | Calculator | References

Inpatient | Ambulatory | Arm | Leg | Sitting | Standing | Auscultation | Machine | 18 May 2007 - 23 Jan 2012

Percentile

90 %  
80 %  
70 %  
60 %

Systolic  
Diastolic

**Medication**

www.meducation.com

0 1 N W Q C 5 D

**Celebrex Capsule 200 mg**

**How to take medicine**  
Take the medicine by mouth once a day.  
Take one (1) pill each time.

Breakfast | Lunch | Dinner | Bedtime

1

**Instructions**  
Swallow the medicine without crushing or chewing it.  
Take the medicine with 250 mL (1 cup) of water.  
Keep the medicine at room temperature. Avoid heat and direct light.  
This medicine may cause you to become more sensitive to the sun. Use sunscreen or wear protective clothing when you are exposed to the sun.  
If you are using this medicine regularly, it is important that you keep taking each dose of this medicine on time even if you are feeling well.

Your Logo Here  
Your Pharmacy Here  
1 Main Street  
Anytown, USA 12345  
555-555-5555

**Prescriptions**

Name	Start Date	End Date	Instructions
Carvedilol 25 MG Oral Tablet	Wed Jul 16 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 daily
Celecoxib 200 MG Oral Capsule (Celebra)	Fri Aug 22 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 daily
Cyclobenzaprine Hydrochloride 30 MG Oral Tablet	Wed Jul 30 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 Three times a day pm
Digoxin 0.25 MG Oral Tablet	Wed Jul 30 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 daily
Digoxin 0.25 MG Oral Tablet (Lansin)	Wed Jul 30 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 daily
Furosemide 20 MG Oral Tablet	Thu Feb 07 2009 00:00:00 GMT-0800 (Pacific Standard Time)		1 daily
Ramipril 5 MG Oral Capsule	Thu Aug 08 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 daily
Warfarin 5 MG Oral Tablet (Coumadin)	Wed Jul 30 2008 00:00:00 GMT-0700 (Pacific Daylight Time)		1 daily

**Your Results**

CRP level test

Use risk: High | Average: 0.5 | 2.1 | High risk of cardiovascular disease

Total cholesterol level

Desirable: 190 | 141 | Borderline High: 160-190 | High: 200+

LDL "bad" cholesterol

Desirable: 100 | 107 | Borderline High: 130-160 | Very High: 160+

HDL "good" cholesterol

Desirable: 40 | 34 | Normal: 40-129 | High: 130+

**Your risk** You show an elevated risk of cardiovascular disease

If you're a smoker with normal blood pressure (120/80 mmHg) but family history of heart disease before the age of 60 (one or both parents) your risk over 10 years to see your test results to calculate your risk of a cardiovascular event at [MyRiskTool.org](#)

**13%** Your risk would be lowered to: 8% if your blood pressure were 120mm/Hg or 9% if you don't smoke and all levels were optimal

**What now?**

- Stay smoke free: One of the best ways to improve your heart disease risk
- Ask your doctor about statins or other medications that can lower cholesterol
- Consider retesting: In 1 or 2 weeks to include a temporary spike in blood levels

Original Design: David McCandless & Stephen Trevisan for Wired Magazine // [informationisbeautiful.net](#)  
Development and maintenance of improved algorithms for the assessment of global cardiovascular risk in women: The Reynolds Risk Score. *Baker et al. JAMA 2007;297:812-819*  
C-reactive protein and parental history improve global cardiovascular risk prediction: The Reynolds Risk Score for Men. *Baker et al. Circulation 2008;118:2201-2211*

# 1 SMART App in 3 SMART Systems



Indiviohealth™ for John Smith

Settings Logout

Mary John D. Smith Joshua Lewis

### Bloodwork Cardiology Result

NOTE: These results are valid for non-diabetic only

**1 About this test**  
This report evaluates your potential risk of heart disease, heart attack, and stroke.

**2 Your Results**

CRP level test: 2.4

Total cholesterol level: 141

LDL "bad" cholesterol: 107

CHP - HMS - © 2011

SMART

< Alex Lewis > Challenge Judge - Logout

### Bloodwork Cardiology Result

NOTE: These results are valid for non-diabetic only

**1 About this test**  
This report evaluates your potential risk of heart disease, heart attack, and stroke.

**2 Your Results**

CRP level test: 2.4

Total cholesterol level: 141

LDL "bad" cholesterol: 107

HDL "good" cholesterol: 34

**3 Your risk** You show an elevated risk of cardiovascular disease

If you're a smoker with normal blood pressure, (130 mm/Hg) but family history of heart attack before the age of 60 (one or both parents) your risk over 10 years is: **12%**

Your risk would be lowered to: **12%** if your blood pressure were 120mm/Hg  
**2%** if you didn't smoke and all levels were optimal  
**12%** if you quit smoking

Use your test results to calculate your risk of a cardiovascular event at ReynoldsRisk.org

**4 What now?**

Manage Apps  
CHP - HMS - © 2011

SMART Query & Analysis Tool

Step 1: Drag up/down to include Step 2: Choose a SMART App below

Step 3: Drag up/down to include Step 4: Choose a SMART App below

### Bloodwork Cardiology Result

NOTE: These results are valid for non-diabetic only

**1 About this test**  
This report evaluates your potential risk of heart disease, heart attack, and stroke.

**2 Your Results**

CRP level test: 2.4

Total cholesterol level: 141

LDL "bad" cholesterol: 107

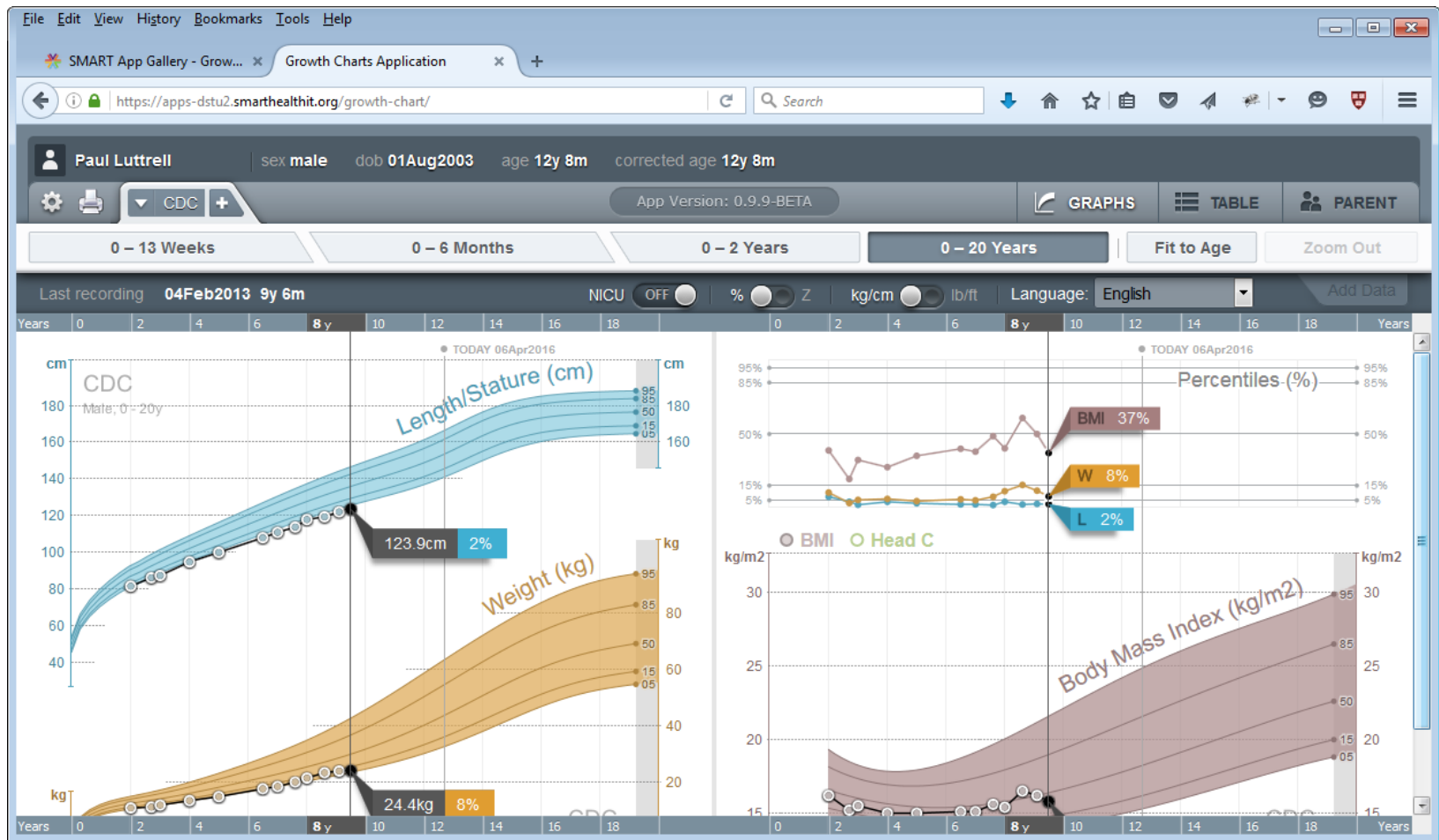
CHP - HMS - © 2011

# What Big Data can do for the Everyday Clinician

## Finding Similar Patients

- Looking at similar patients can help predict:
  - Future outcomes and responses to therapy
  - Course of disease
  - Penetrance of genetic variants
  - Likelihood that a diagnostic pathway might be fruitful
- Big Data Commons is an opportunity for combining data from the Electronic Health Record, Specialized Health Databases, Analytics from Big Data Queries, and presentation in SMART Apps
- Presentation of results can be greatly enhanced with engaging visualizations for the provider making difficult, complex decisions

# <https://gallery.smarthealthit.org/boston-childrens-hospital/growth-chart>



File Edit View History Bookmarks Tools Help

SMART App Gallery - Grow... x Growth Charts Application x +

https://apps-dstu2.smarthealthit.org/growth-chart/ Search

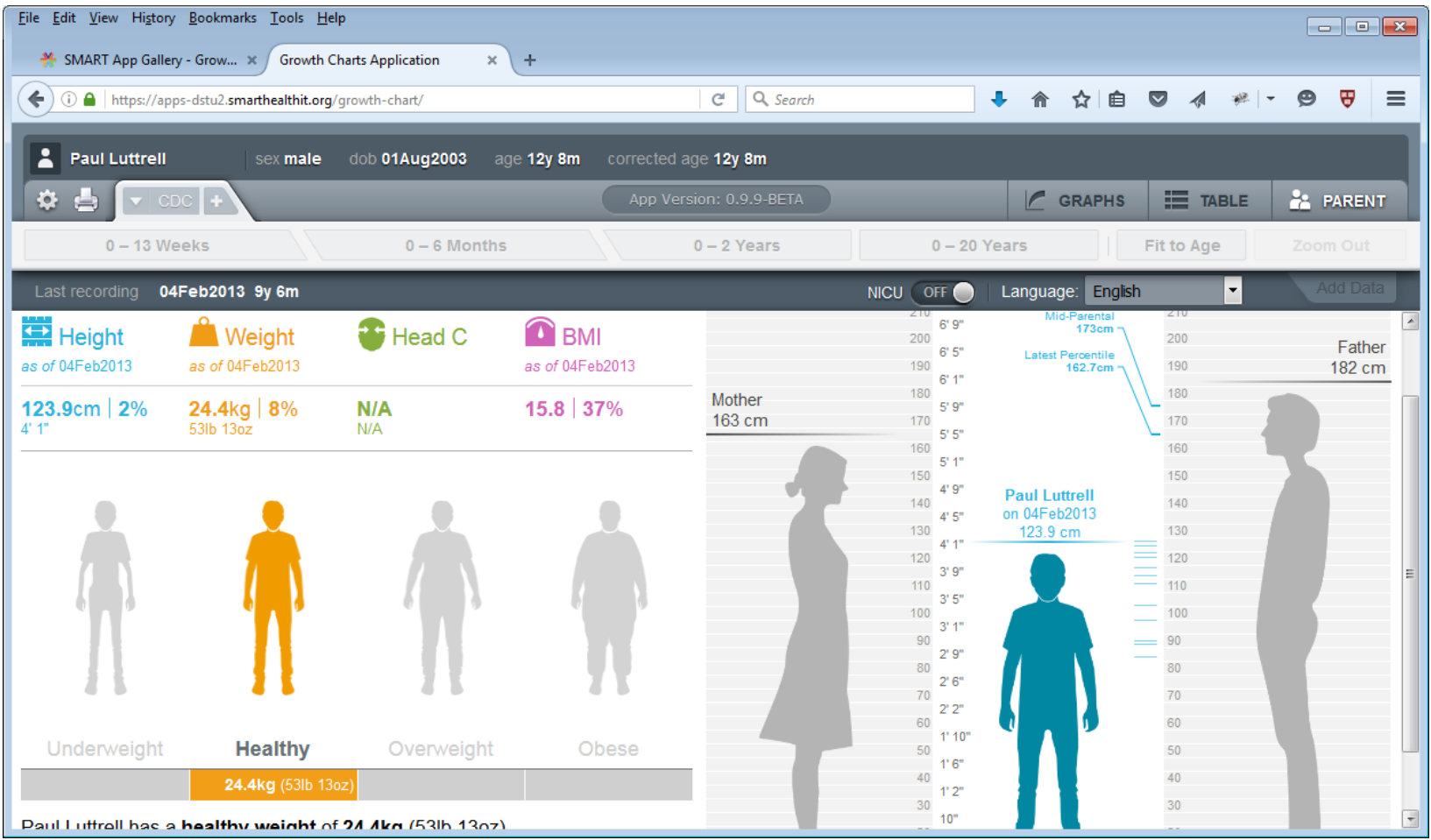
Paul Luttrell sex male dob 01Aug2003 age 12y 8m corrected age 12y 8m

CDC + App Version: 0.9.9-BETA GRAPHS TABLE PARENT

0 - 13 Weeks 0 - 6 Months 0 - 2 Years 0 - 20 Years Fit to Age Zoom Out

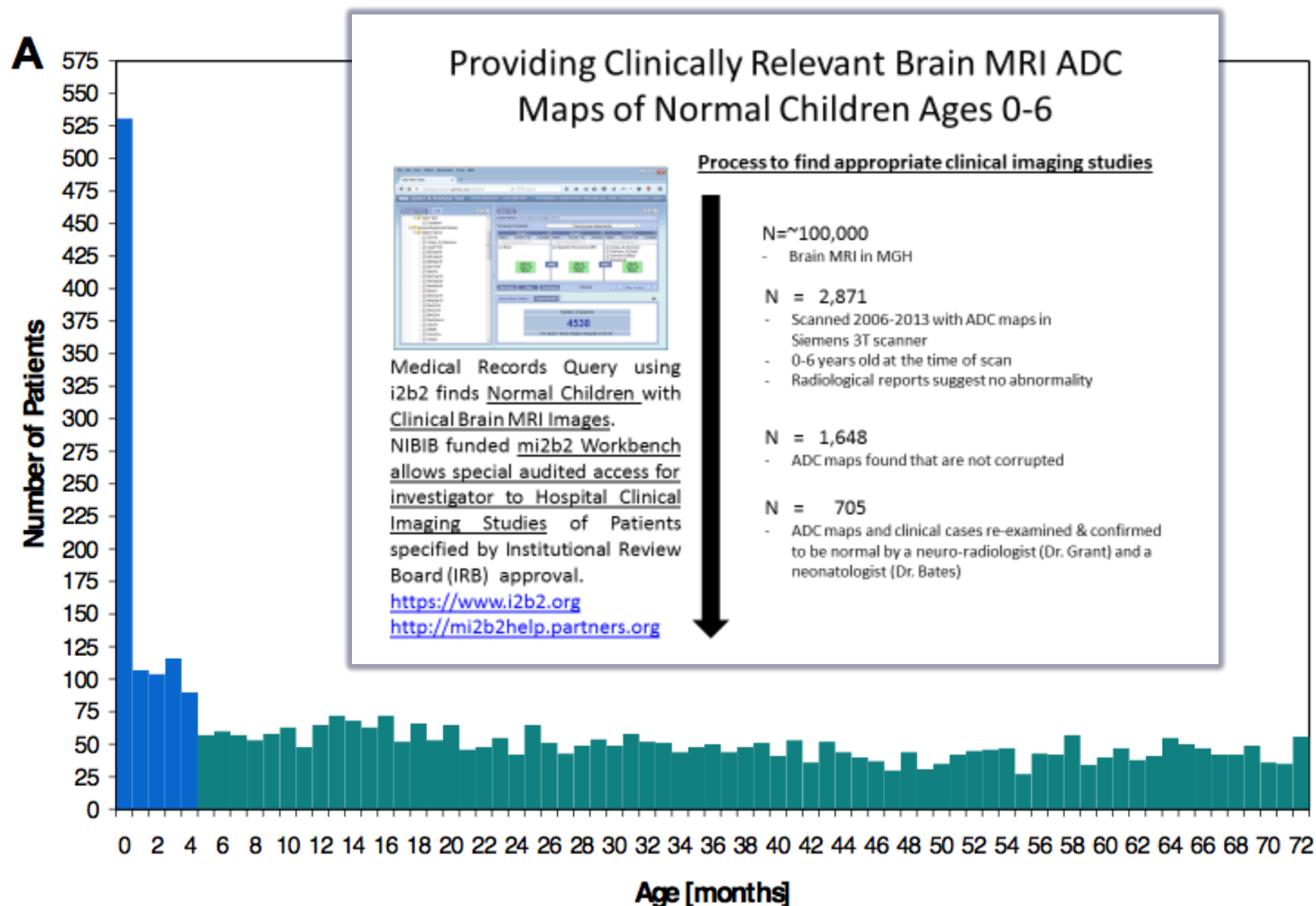
Last recording 04Feb2013 9y 6m NICU OFF kg/cm lb/ft Language: English Add Data

Entry Date	08Aug2006	08Aug2007	07Aug2008	05Feb2010	06Aug2010	14Mar2011	06Aug2011	13Mar2012	12Sep2012	04Feb2013	
Age	3y 6d	4y 6d	5y 5d	6y 6m	7y 5d	7y 7m	8y 5d	8y 7m	9y 1m	9y 6m	
Annotation	See all										
<b>Length</b>	<b>cm</b>	<b>87.6</b>	<b>95</b>	<b>100.2</b>	<b>108.2</b>	<b>111.1</b>	<b>114</b>	<b>118.1</b>	<b>119.6</b>	<b>122.3</b>	<b>123.9cm</b>
Percentile	%	2	4	3	2	2	2	4	2	3	2
Z Score	Z	-2.1	-1.8	-1.9	-2	-2	-2.1	-1.7	-2	-2	-2
Velocity	cm/yr	5.6	5.3	5.3	5.2	5.1	5.2	3.9	4.8	4	To here
<b>Weight</b>	<b>kg</b>	<b>12.1</b>	<b>13.7</b>	<b>15.1</b>	<b>17.9</b>	<b>18.7</b>	<b>20.3</b>	<b>21.7</b>	<b>23.6</b>	<b>24.2</b>	<b>24.4kg</b>
Percentile	%	5	6	4	6	5	7	11	15	12	8
Z Score	Z	-1.6	-1.6	-1.7	-1.6	-1.6	-1.4	-1.2	-1	-1.2	-1.4
Velocity	kg/yr	1.9	1.9	2.1	2.2	2.3	2.2	1.8	0.9	0.5	To here
<b>Head C</b>	<b>cm</b>	—	—	—	—	—	—	—	—	—	—
Percentile	%	—	—	—	—	—	—	—	—	—	—
Z Score	Z	—	—	—	—	—	—	—	—	—	—
Velocity	cm/yr	—	—	—	—	—	—	—	—	—	To here
<b>BMI</b>	<b>kg/m2</b>	<b>15.5</b>	<b>15</b>	<b>15</b>	<b>15.1</b>	<b>15.1</b>	<b>15.6</b>	<b>15.4</b>	<b>16.5</b>	<b>16.2</b>	<b>15.9</b>



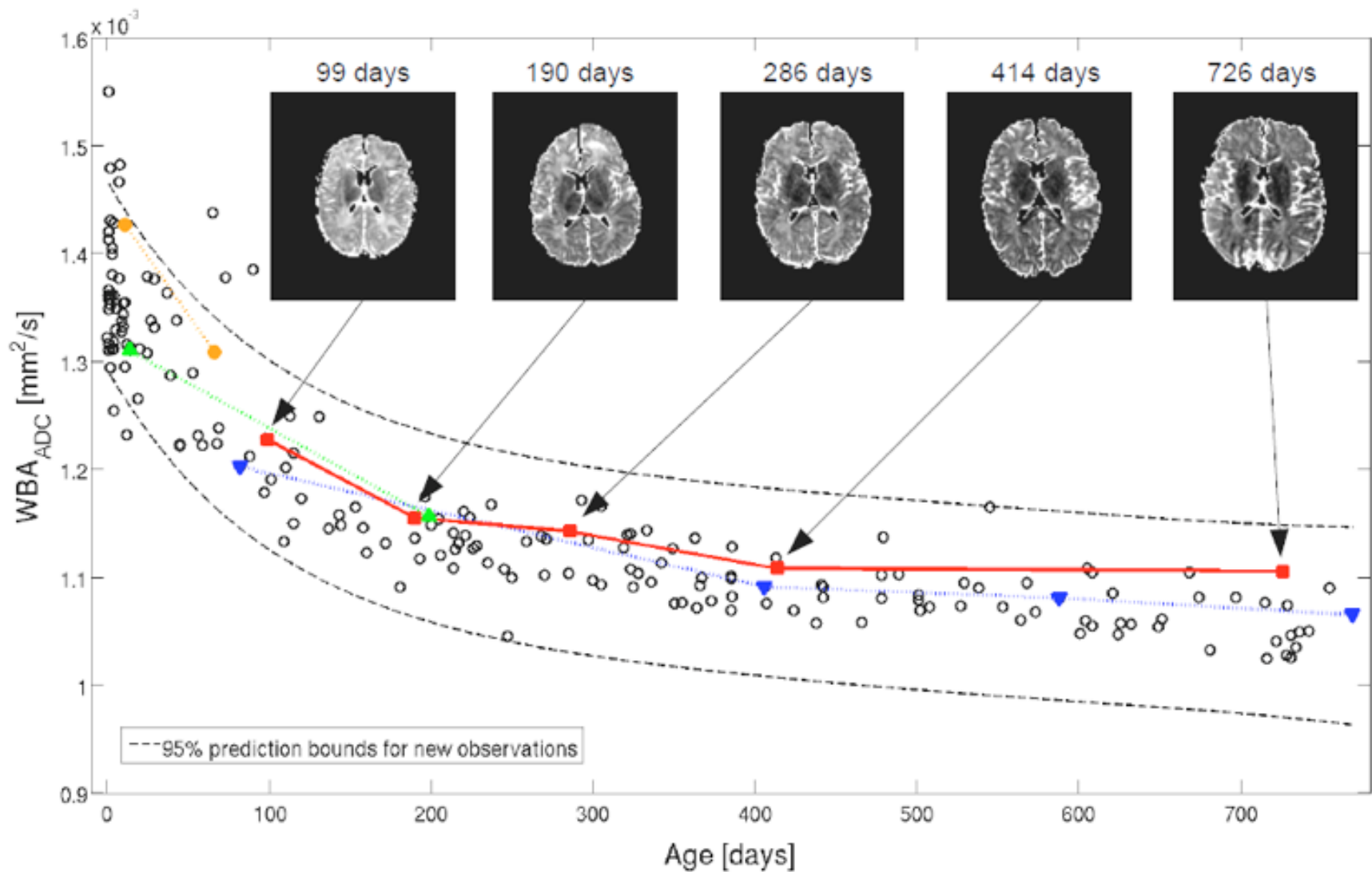


# Find Normal MRI's at All Ages 0-6 y/o



Number of patients who had a brain MRI scan at a particular age in months from 0 to 6 years (A) and in weeks from 0 to 4 months (B)

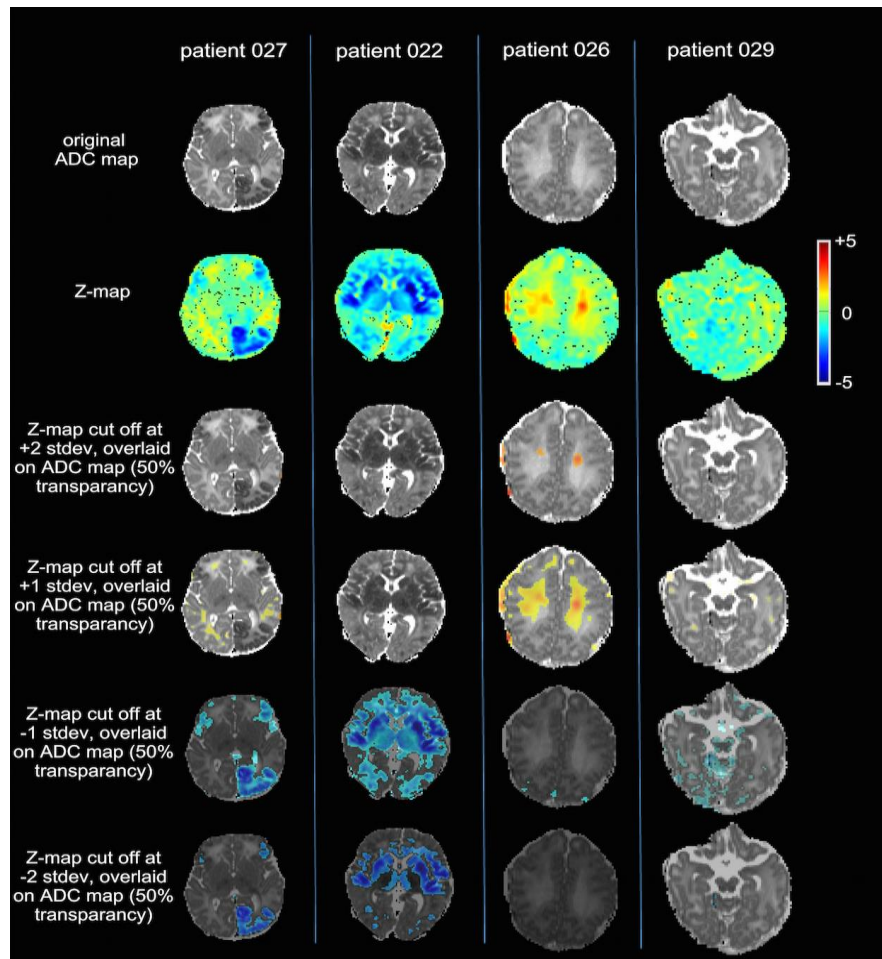
# Determining a Normal Child's MRI



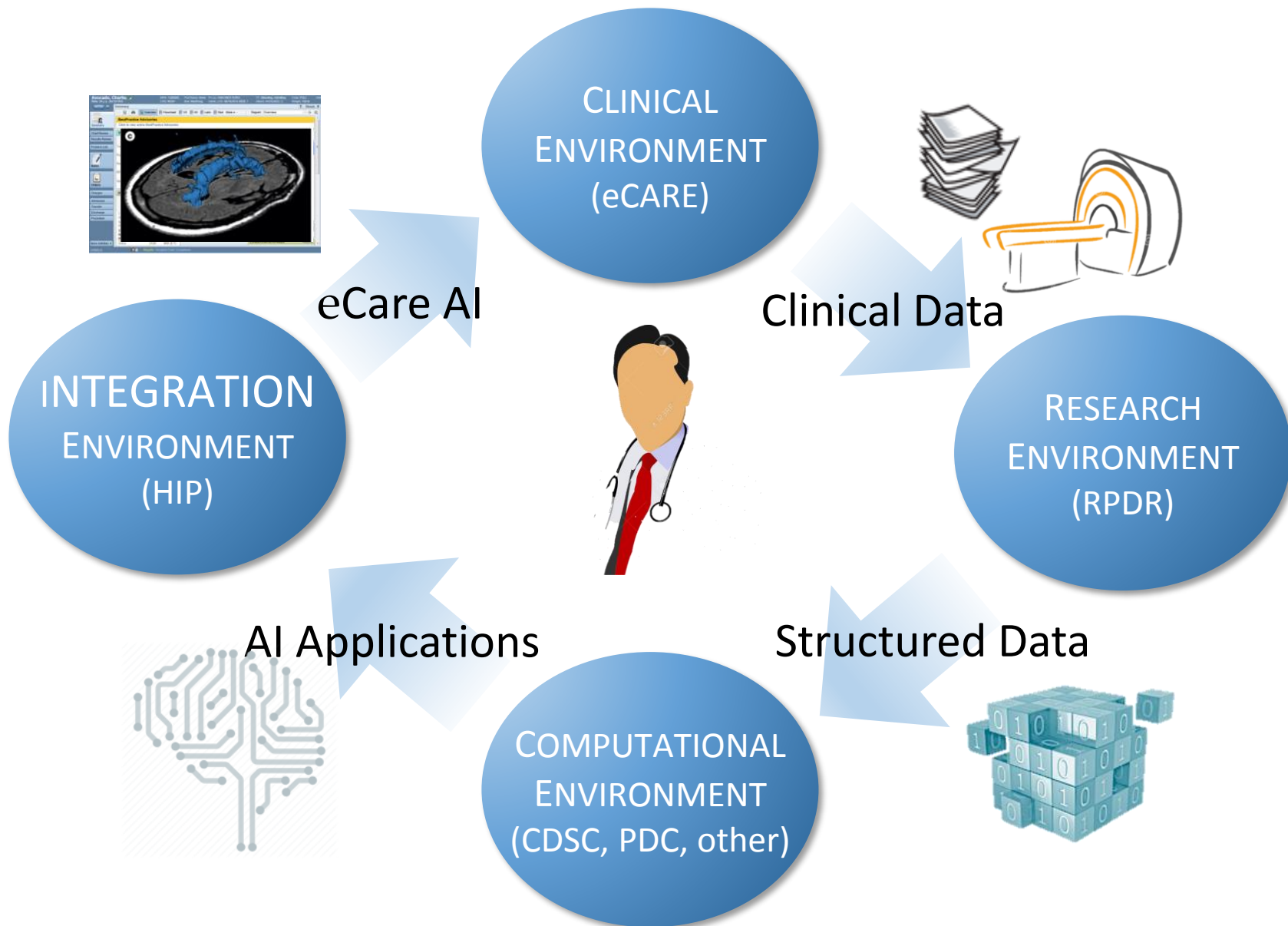
# Atlases provide a visual guide for Radiology Decision Support, such as determining Perinatal Hypoxic Ischemic Encephalopathy

ADC map from 4 infants:  
Each statistically compared to age matched atlas yields visual guide to pathology

**Quantitative analysis tools + large data sets = Great insights for practicing doctors**



# PARTNERS CLINICAL DEVELOPMENT ENVIRONMENT



# Tribute to...

## ■ I2b2 Core Team

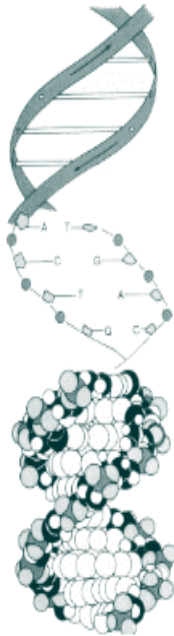
- Isaac Kohane
- Susanne Churchill
- Michael Mendis
- Christopher Herrick
- Griffin Weber
- Paul Avillach
- Lori Phillips
- Janice Donahoe
- Nich Wattanasin
- David Wang
- Vivian Gainer
- Victor Castro
- Andrew Cagan
- Wayne Chan

## ■ SMART Apps Team

- Calum MacRae
- Sandy Aronson
- Mike Oats
- Layne Ainsworth
- Kenneth Mandl
- Joshua Mandel

## ■ Radiology (mi2b2) Team

- Randy Gollub
- P Ellen Grant
- Kathy Andriole
- Kallirroi Retzepi
- Rudolph Pienaar
- Lilla Zollei
- Yangming Ou



# **I2b2, SHRINE, and SMART Information and Software on the Web**

***i2b2 Homepage (<https://www.i2b2.org>)***

***i2b2 Software (<https://www.i2b2.org/software>)***

***i2b2 Community Site (<https://community.i2b2.org>)***

***SMART Platforms Homepage (<http://smarthealthit.org>)***

Partners Healthcare, NIH/NCBC/BD2K; /NIMH; /NCATS; /NIBIB; /NHGRI

NIH R01 EB014947

NIH U54 LM008748

NIH U01 HG008685

PCORI 282364.5077585.0007

NIH U54 HG007963

NIH R01 AT006364

NIH R01 AT005280

NIH P01 AT006663

THANK YOU