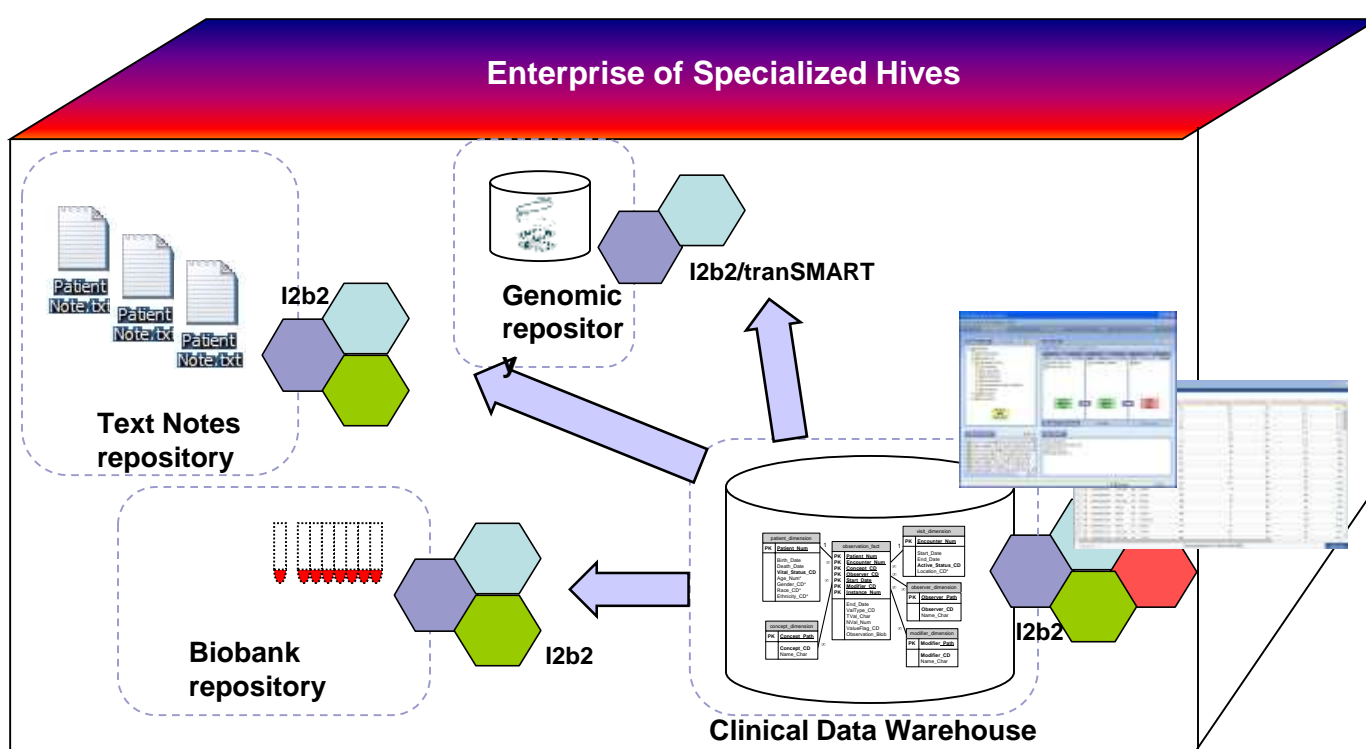


Developing Specialized Hives to Contribute to the i2b2 Ecosystem

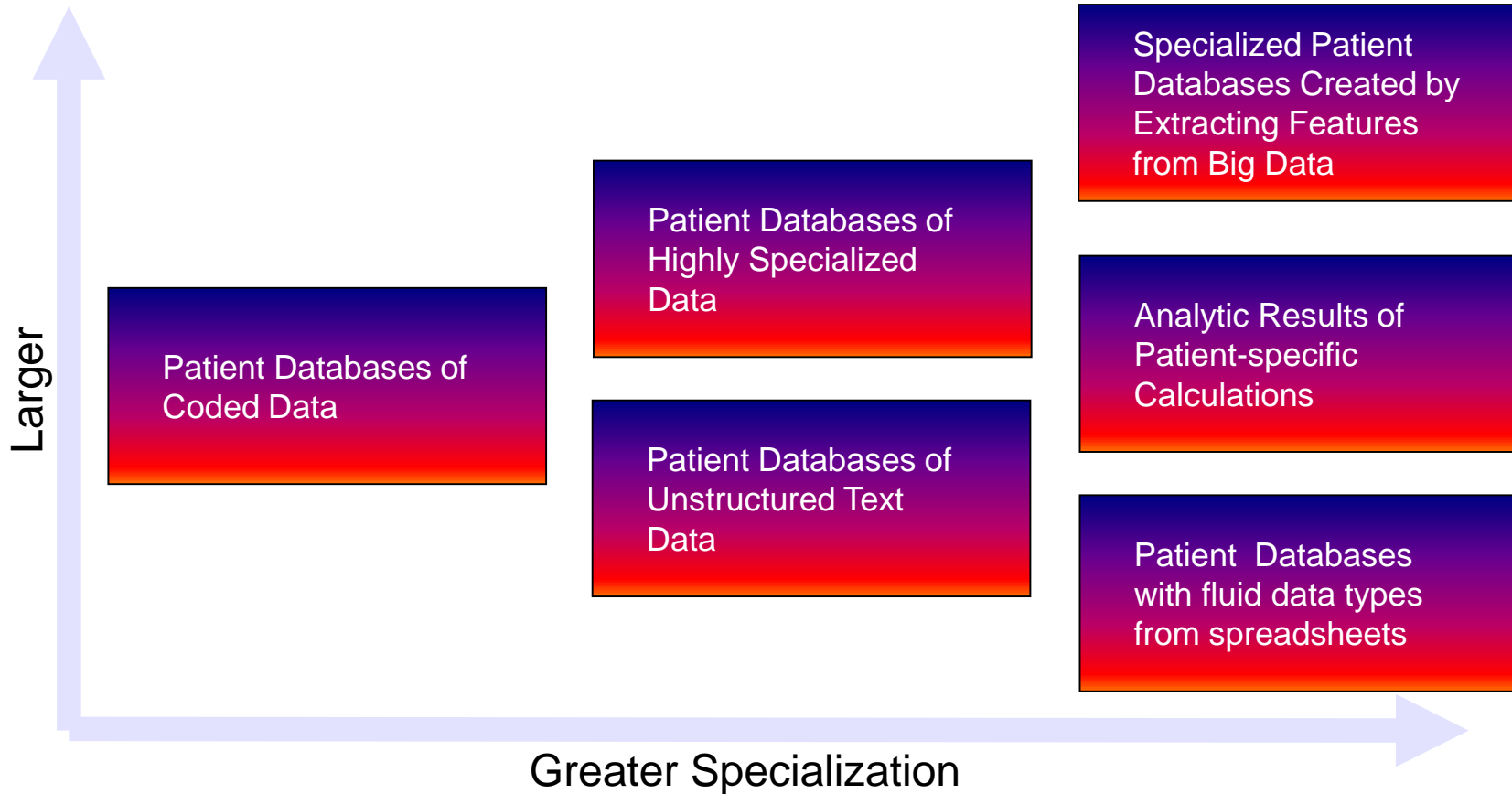
Shawn Murphy

Michael Mendis

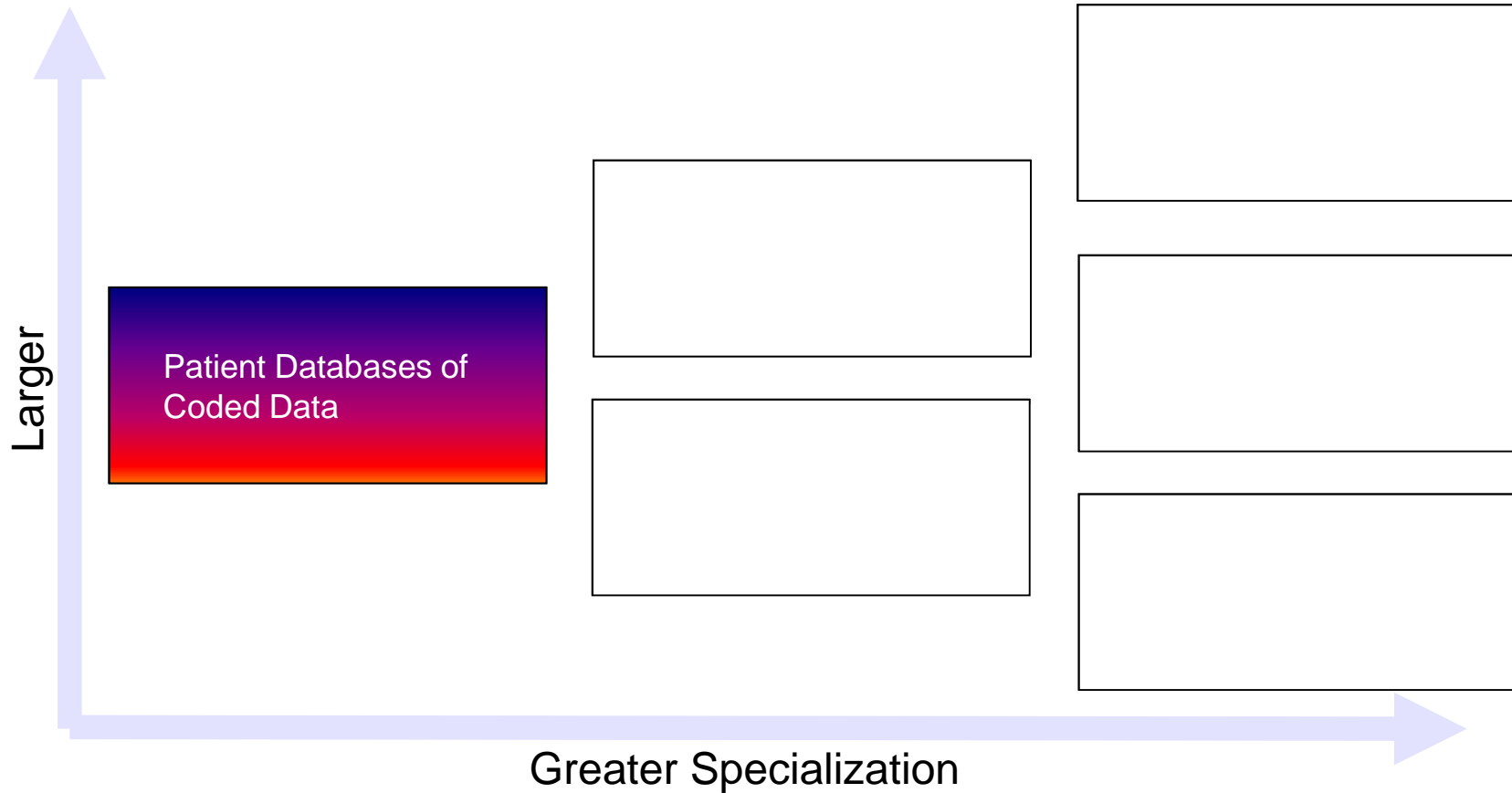
Distributed Query System links together Specialized i2b2 Hives



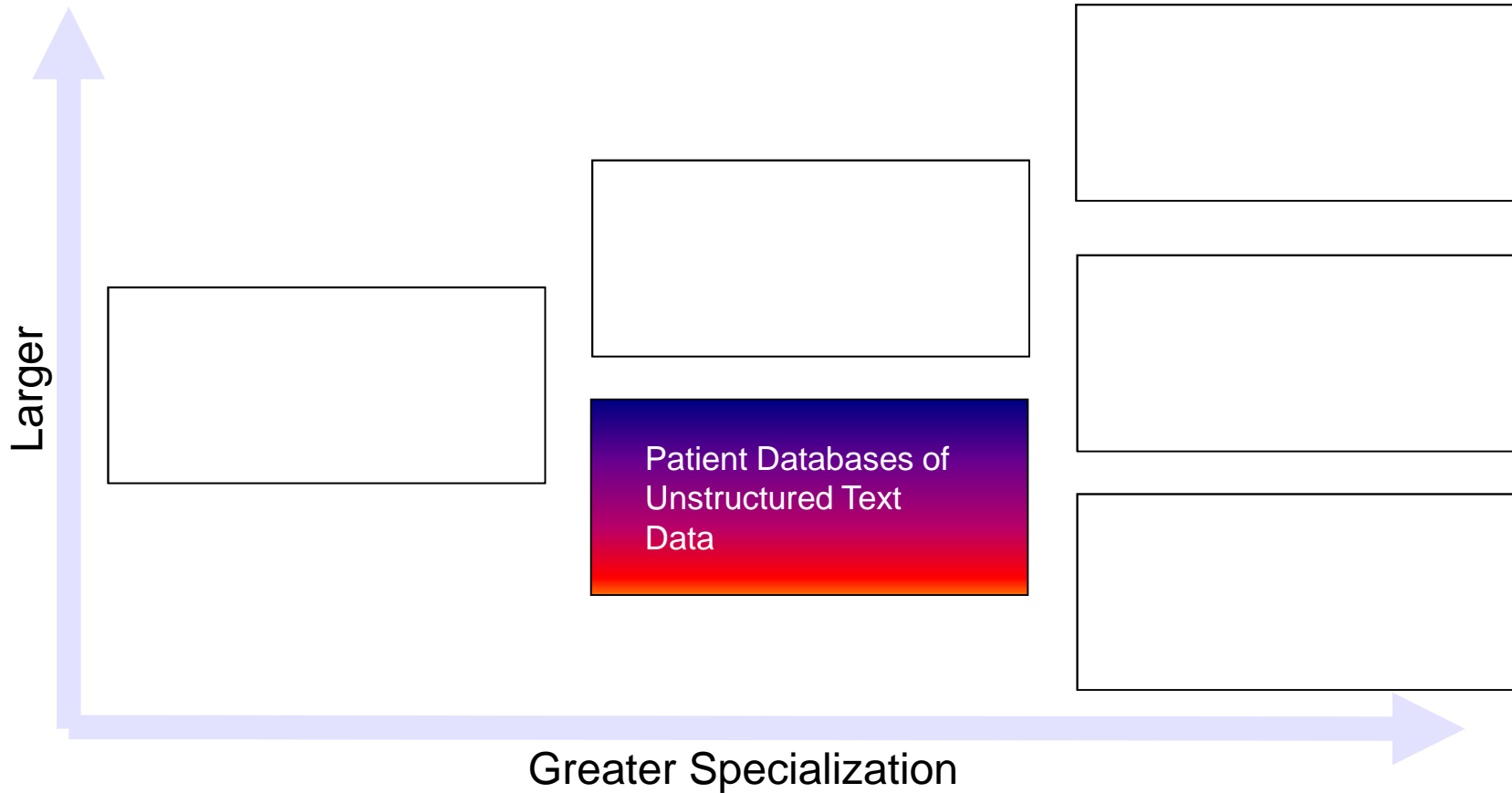
Specialized i2b2 Hives that Exist and are in Development



Specialized i2b2 Hives that Exist and are in Development



Specialized i2b2 Hives that Exist and are in Development



I2b2/RPDR Text-based Queries

- 130 Million Notes Available alongside 2.4 Billion Facts to query in i2b2/RPDR
- Researchers search for phrases of interest in a variety of clinical notes
 - Cardiology, Discharge Summaries, Endoscopy, LMR, Operation, Pathology, Pulmonary, and Radiology
- Transparent to users - add text based searching to query tool
- Notes are updated nightly in special notes repository
- Several steps taken to conceal PHI through query
 - Researchers will not be able to query on numbers of any kind
 - Names are excluded from text index making them “unsearchable”
 - Search terms are also matched against a PHI database
 - Separate database is securely maintained with latest patient PHI, including names, addresses, numbers, nicknames, etc.
 - Search terms are queried in Notes Repository and PHI database
 - Patients that are returned from both databases are excluded from results
 - Final aggregate counts are obfuscated to hide true totals and limit ability to identify one individual
- Policy can be enforced at Query Tool when item is used in query

Search for Diabetes, A1C, and SGLT2 Inhibitors

RPDR Enhanced Query Tool

Logging: Chris Herrick in workgroup of Herrick, Christopher D.

Find Patients | Match Controls | Analyze Patients | Request Detailed Data | Request Specimens | Request Images | Understand Data | Get Help

Query Name: Diabetes m..., Hgb A1C (G..., Invokana (... on 09/28/2015 #3

Independent of Visit | Sensitivity < | Reset all groups to >0 | >Specificity

Group 1 of 3 | Group 2 of 3 | Group 3 of 3

Dates | Recorded > Dx | Exclude

Independent of Visit

The items of this group are joined together, and intersected with other groups.

Diabetes mellitus

Hgb A1C (Group:GHBA1C) > 6.50

Invokana (canagliflozin) - LMR 8989

Reset All | First Group | Shift Left | Shift Right | Last Group | New Group

Run Query | Total count: 414±3 patient(s)

Gender: Female 205±3, Male 210±3

Age: Age 0-9 <3, Age 9-19 <3, Age 20-29 7±3, Age 30-39 15±3, Age 40-49 56±3, Age 50-59 120±3, Age 60-69 147±3, Age 70-79 44±3, Age 80-89 11±3, Age >=90 <3

Race: A. Indian <3, A 8±3, B 31±3, H 10±3, W 332±3, O 14±3, U 17±3

Vital: Alive 413±3, Deceased

Enhance w/ LMR Note Search

The screenshot shows the RPDR Enhanced Query Tool interface. The main window title is "RPDR Enhanced Query Tool" and the URL is "https://rpdrdevsecure.mgh.harvard.edu/rpdrwebclient". The user is logged in as "Chris Herrick in workgroup of Herrick, Christopher D.". The interface includes a navigation bar with buttons for "Find Patients", "Match Controls", "Analyze Patients", "Request Detailed Data", "Request Specimens", "Request Images", "Understand Data", and "Get Help".

The "Query Items" section on the left lists various categories such as "Standard Query Items", "Encounter detail", "Demographics", "Diagnosis", "Diagnosis Related", "Laboratory tests", "Radiology Tests", "Medications", "Microbiology", "Molecular medicine", "Health History", "Providers", "Procedures", "Specimens", "Transfusion Services", "Epic Reason For", "Notes Search", "Cardiology", "Discharge Summary", "Endoscopy R", "LMR Note", "Operation Note", "Pathology", "Pulmonary", "Radiology", "Custom Query Items", "Query Library", and "Patient Sets".

The "Query Name" is "Diabetes m..., Hgb A1C (G..., Invokana (... on 09/28/2015 #3". The "Independent of Visit" dropdown is set to "Independent of Visit". The "Sensitivity" and "Specificity" fields are both set to ">0".

The "Search within LMR Notes" dialog box is open, displaying the following text:

You are allowed to search with a narrative text associated within the selected term/concept.

Enter one or more search phrases (separated by comma) for search.

Each phrase will be searched individually and the results will be unioned together.

Enter Search Text:

Leave search text blank if no search request is requested.

The dialog box has "OK", "Clear", and "Cancel" buttons.

The background interface shows a "Vital" section with a bar chart for "Alive" (413±3) and "Deceased" (0±0). The "Vital" section is checked.

RPDR – Additional Patients Returned

RPDR Enhanced Query Tool

Logging: Chris Herrick in workgroup of Herrick, Christopher D.

Find Patients | Match Controls | Analyze Patients | Request Detailed Data | Request Specimens | Request Images | Understand Data | Get Help

Query Name: Diabetes m..., Hgb A1C (G..., Invokana (... on 09/28/2015 #6

Independent of Visit | Sensitivity < | Reset all groups to >0 | >Specificity

Group 1 of 3

Independent of Visit

The items of this group are joined together, and intersected with other groups.

- Diabetes mellitus

Group 2 of 3

Independent of Visit

The items of this group are joined together, and intersected with other groups.

- Hgb A1C (Group:GHBA1C) > 6.50

Group 3 of 3

Independent of Visit

The items of this group are joined together, and intersected with other groups.

- Invokana (canagliflozin) - LMR 8989
- LMR Note SEARCH FOR Invokana, Farxiga, Jardiance, Canagliflozin, Dapagliflozin, Empagliflozin

Reset All | First Group | Shift Left | Shift Right | Last Group | New Group

Run Query Total count: 727±3 patient(s)

Gender

Female	358±3
Male	371±3

Age

Age 0-9 <3	
Age 9-19 <3	
6±3	22±3
95±3	204±3
239±3	107±3
18±3	3±3

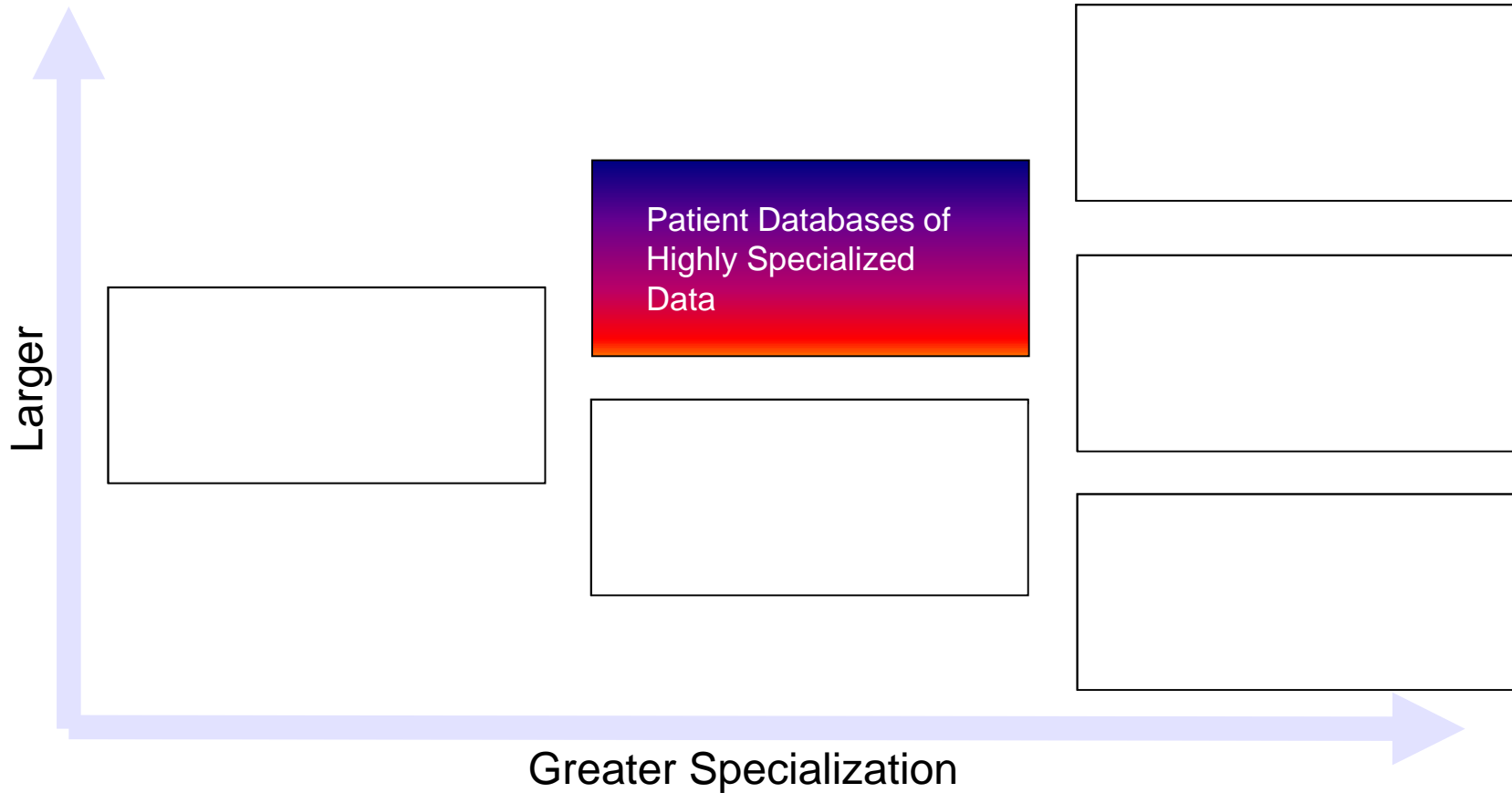
Race

A. Indian <3	
A	23±3
B	66±3
H	20±3
W	569±3
O	27±3
U	22±3

Vital

Alive	719±3
Deceased	6±3

Specialized i2b2 Hives that Exist and are in Development



File Edit View History Bookmarks Tools Help

Partners Biobank Portal

https://biobankportaldev.partners.org/webclient/

Search

PARTNERS HEALTHCARE BIOBANK PORTAL

Find Patients Make Request Help & Support Shawn Murphy, MD

Navigate Terms Find

- Biobank Demographics
- Biobank Genomics Search
 - Illumina MEGA Genotyping (b1_Nov2015) - 4841
 - dbSNP rs Identifier - 4841
 - Gene - 4841
 - indel - 4841
 - SNP - 4841
- Biobank Sample Types

Workplace

- ajp12
- bg84
- lcp5
- nw09
- rm30
- snm0
- vc070
- vsg

Previous Queries

- SNP
- Gene
- Gene
- indel
- indel
- SNP-Asian@08:23:10 [6-19-2016] [snm0]

Query Tool

Query Name: SNP@13:45:14

Temporal Constraint: Treat all groups independently

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
SNP [contains "EGFR AND Homozygous AND (Frameshift OR In_frame OR missense OR nonsense OR start_loss OR stop_loss)"]								

AND

drop a term on here

1 Group

New Group

Results Query Report Download Results

Number of patients

381

For Query "SNP@13:45:14"

Search by Gene - SNP

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

Enter Gene Name: EGFR

Zygoty: Homozygous

Consequence: 6 of 15 selected

- Exon (non-coding)
- Frameshift
- In-frame
- Intron
- Missense

OK Cancel

Navigate Terms Find

- Biobank Demographics
- Biobank Genomics Search
 - Illumina MEGA Genotyping (b1_Nov2015) - 4841
 - dbSNP rs Identifier - 4841
 - Gene - 4841
 - indel - 4841
 - SNP - 4841
- Biobank Sample Types

Workplace

- ajp12
- bg845
- lcp5
- nw096
- rm302
- snm0
- vc070
- vsg

Previous Queries Find

- SNP@13:45:14 [6-21-2016] [snm0]
- Gene@11:25:58 [6-20-2016] [bg845]
- Gene@11:24:43 [6-20-2016] [bg845]
- indel@09:50:44 [6-20-2016] [rm302]
- indel@09:48:56 [6-20-2016] [rm302]
- SNP-Asian@08:23:10 [6-19-2016] [snm0]

Download Data

Specify Data Preview Data File Download Data File History

Download to start processing the entire file.

Process Started: 6-21-2016 @ 13:46:9 | Finished: 6/21/2016 @ 13:47:22

[Proceed to Download](#)

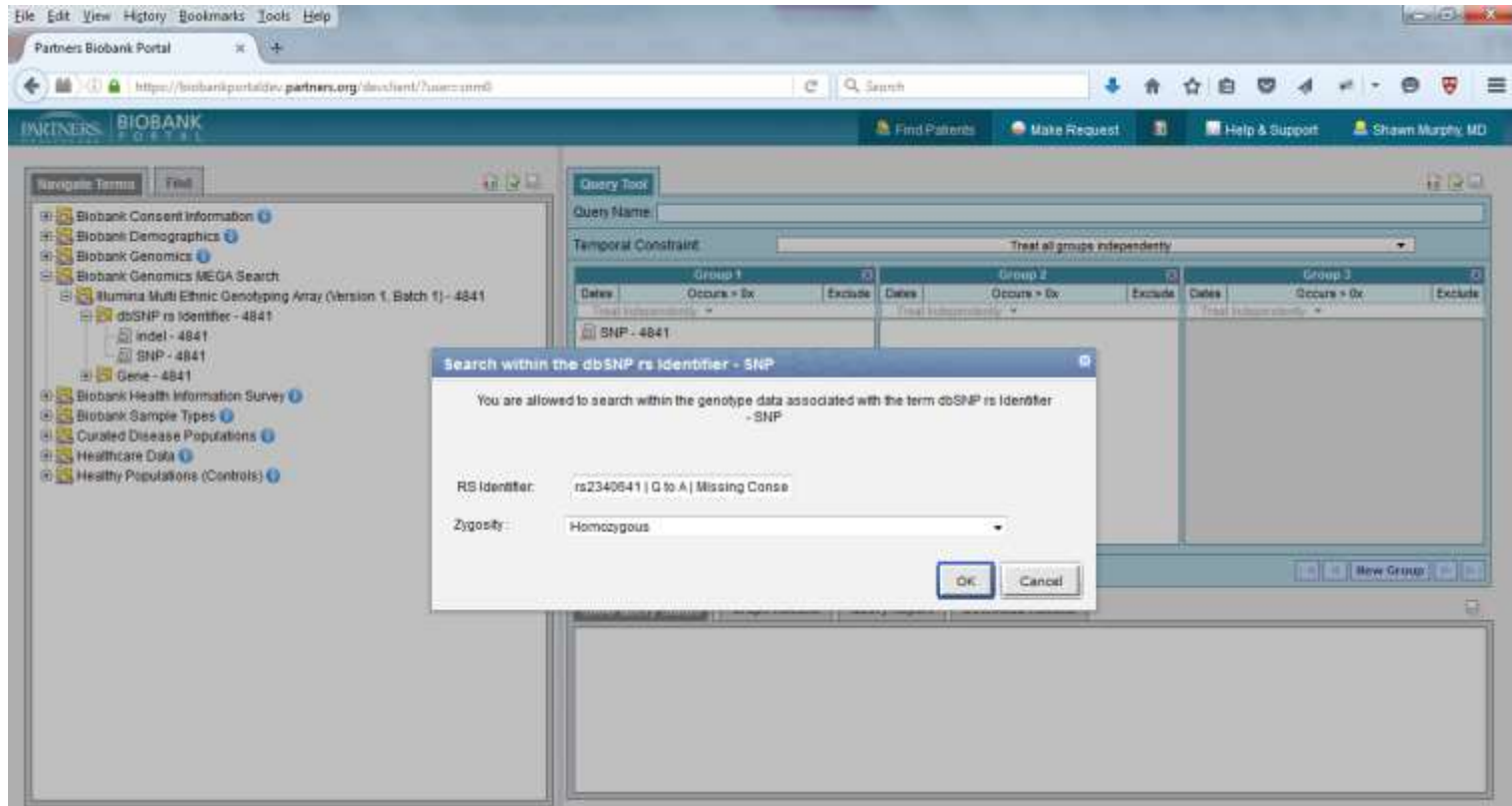
Gender	Age	Race	DNA	Plasma	Serum	SNP [[contains "EGFR AND Homozygous AND (Frameshift OR In_frame OR missense OR nonsense OR start_loss OR stop_loss)"]] [Existence (Yes/No)]
M	67	White	Yes	Yes	Yes	Yes
F	48	Unknown	Yes	No	No	Yes
M	88	White	Yes	No	No	Yes
F	40	White	Yes	No	No	Yes
F	68	White	Yes	No	No	Yes

Plugins

Detailed List View Category: ALL

- Admin Tool: MRN List to Patient Set Creator**
Create a Biobank Portal Patient Set from a list of MRNs
- Biobank Portal Identified Data Request**
Request identified data from the Biobank Portal

Query 1.68 billion rows of Genomic Data for Specific Variants



File Edit View History Bookmarks Tools Help

Partners Biobank Portal

https://biobankportaldev.partners.org/bsvc/ctrl/hsacc:ctrl6

Partners BIOBANK PORTAL

Find Patients Make Request Help & Support Shawn Murphy, MD

Navigate Terms Find

- Biobank Consent Information
- Biobank Demographics
- Biobank Genomics
 - Biobank Genomics MEGA Search
 - Illumina Multi Ethnic Genotyping Array (Version 1, Batch 1) - 4841
 - dbSNP rs Identifier - 4841
 - indel - 4841
 - SNP - 4841
 - Gene - 4841
- Biobank Health Information Survey
- Biobank Sample Types
- Curated Disease Populations
- Healthcare Data
- Healthy Populations (Controls)

Query Tool

Query Name:

Temporal Constraint: Treat all groups independently

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
SNP [contains 'rs2340641 AND Homozygous AND G_A AND missing_consequence']								

one or more of these AND drop a term on here

Run Query Clear 1 Group New Group

Show Query Status Graph Results Query Report Download Results

File Edit View History Bookmarks Tools Help

Partners Biobank Portal

https://biobankportaldev.partners.org/#!/search/3393

Find Patients Make Request Help & Support Shawn Murphy, MD

Partners BIOBANK PORTAL

Query Tool

Query Name: SNP@21:36:43

Temporal Constraint: Treat all groups independently

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
SNP [contains 'rs2340641 AND Homozygous AND G_A AND missing_consequence']								
one or more of these			AND			drop a term on here		

Run Query Clear 1 Group New Group

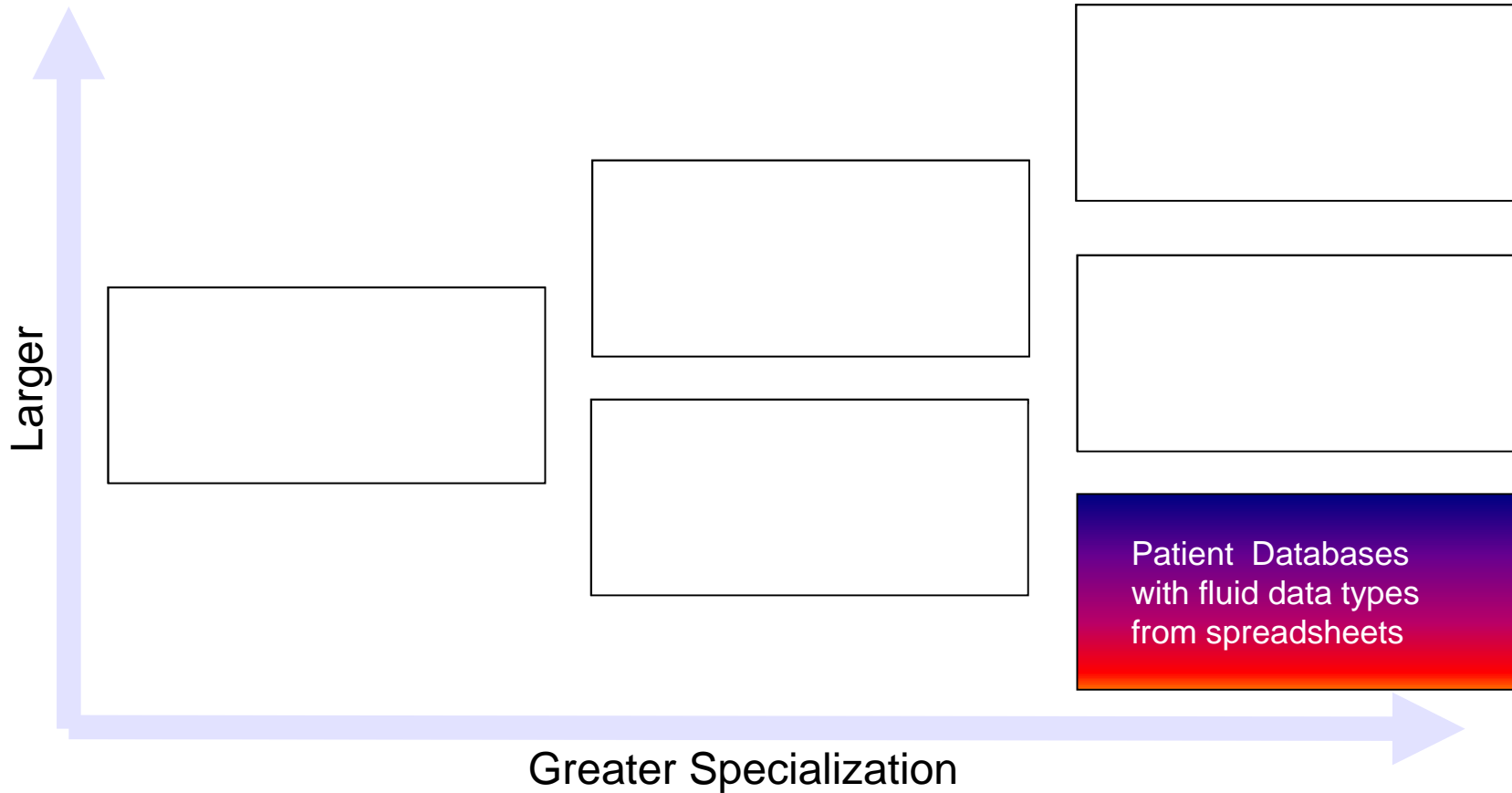
Show Query Status Graph Results Query Report Download Results

Number of patients

3393

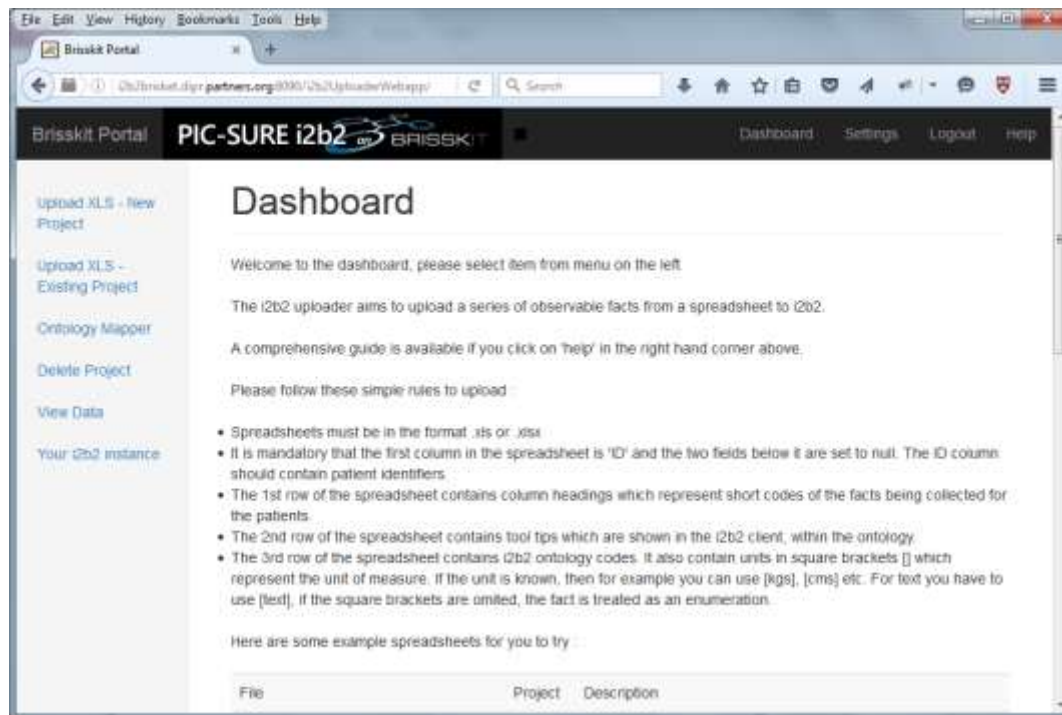
For Query "SNP@21:36:43"

Specialized i2b2 Hives that Exist and are in Development



Put together a new Hive very quickly with BRISSKIT

- BRISSKIT is an open source project that was developed in Leicester, UK
 - Allows Spreadsheet to be uploaded into i2b2 from a web client
 - Creates an instant i2b2 repository of spreadsheet data



File Edit View History Bookmarks Tools Help

Brisskit Portal

i2b2brisket.dipr.partners.org:8090/i2b2UploaderWebapp/h

Brisskit Portal **PIC-SURE i2b2 on BRISSKIT** Dashboard Settings Logout Help

Upload XLS - New Project

Upload XLS - Existing Project

Ontology Mapper

Delete Project

View Data

Your i2b2 instance

File

project1_basic.xls

project2_1_optional

project2_1_optional_fields-1.xls [Read-Only] [Compatibility Mo...]

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW Murphy,...

Clipboard Font Alignment Number Styles

A4 pr-00001

	A	B	C	D	E	F	G	H
1	ID	OBS_START_DATE	AGE	GENDER	RACE	DEATH	MARITAL_STATUS	WORK_STATUS
2	null	null	Age	Sex	Race	Death	Marital Status	Work Status
3	null	null	age	Sex	Race	death	marital status	work status
4	pr-00001	02-06-2014 12:00:00	35	M	Hispanic	1	single	fulltime
5	pr-00002	02-06-2014 12:30:00	46	F	Asian	1	married	not employed
6	pr-00003	03-06-2014 13:23:00	54	M	White	1	widowed	parttime
7	pr-00004	04-06-2014 11:45:00	33	F	White	1	separated	self employed
8	pr-00005	null	65	M	Asian	0	divorced	retired

READY SUM: 84016.8 79%

tab Vital Status is mapped to the column Death in the 'Data' tab. This spreadsheet also contains 5 optional columns which are OBS_START_DATE, AGE, GENDER, RACE and DEATH. OBS_START_DATE is a special column which marks the start

The screenshot shows a web browser window with the following elements:

- Browser Tab:** Brisskit Portal
- Address Bar:** i2b2brissket.dipr.partners.org:8090/i2b2UploaderWebi
- Page Header:** Brisskit Portal | PIC-SURE i2b2 on BRISSKIT | Dashboard | Settings | Logout | Help
- Left Navigation Menu:**
 - Upload XLS - New Project (highlighted)
 - Upload XLS - Existing Project
 - Ontology Mapper
 - Delete Project
 - View Data
 - Your i2b2 instance
- Main Content Area:**

Upload XLS - New Project

Please enter the name of project you wish to create :

alphanumeric characters only (due to database constraints).

File to upload:

project2_1_optional_fields-1.xls

Brisskit, University of Leicester, Jisc accepts no liability for loss or theft of sensitive data submitted through these media and does not recommend the transfer of sensitive data via these methods.

Tick to accept

Press here to upload the file!

File Edit View History Bookmarks Tools Help

Brisskit Portal

i2b2brisket.dipr.partners.org:8090/i2b2UploaderWebapp/i

Search

Dashboard Settings Logout Help

Brisskit Portal **PIC-SURE i2b2 on BRISSKIT**

Upload XLS - New Project

Upload XLS - Existing Project

Ontology Mapper

Delete Project

View Data

Your i2b2 instance

Navigate Terms Find

- bd2kdistributed
 - AGE
 - ALCOHOL_AGE_STARTED
 - ALCOHOL_EVER
 - ALCOHOL_NOW
 - BEER_PER_WEEK
 - CIGARETTES_PER_WEEK
 - CL_STATUS
 - CURRENTLY_SMOKE
 - DEATH
 - DIASTOLIC
 - GENDER
 - HEIGHT
 - MARITAL_STATUS
 - MYINFARC
 - RACE
 - SMOKED_AGE_STARTED
 - SMOKED_EVER
 - SYSTOLIC
 - TEXT_COL
 - WEIGHT
 - WORK_STATUS**
 - fulltime
 - not employed
 - parttime
 - retired
 - self employed

Query Tool

Query Name:

Temporal Constraint: Treat all groups independently

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
fulltime								
parttime								

one or more of these AND drop a term on here

Run Query Clear 1 Group New Group

Show Query Status Graph Results Query Report

File Edit View History Bookmarks Tools Help

Brisskit Portal

Qib2brisket.dgr partners.org:8080/Qib2uploaderWebapp/1

Search

Dashboard Settings Logout Help

PIC-SURE i2b2 on BRISSKIT

Upload XLS - New Project

Upload XLS - Existing Project

Ontology Mapper

Delete Project

View Data

Your i2b2 instance

- ALCOHOL_EVER
- ALCOHOL_NOW
- BEER_PER_WEEK
- CIGARETTES_PER_WEEK
- CL_STATUS
- CURRENTLY_SMOKE
- DEATH
- DIASTOLIC
- GENDER
- HEIGHT
- MARITAL_STATUS
- MYINFARC
- RACE
- SMOKED_AGE_STARTED
- SMOKED_EVER
- SYSTOLIC
- TEXT_COL
- WEIGHT
- WORK_STATUS
 - fulltime
 - not employed
 - parttime
 - retired
 - self employed

Group 1	Group 2	Group 3
Dates Occurs > 0x Exclude	Dates Occurs > 0x Exclude	Dates Occurs > 0x Exclude
Treat independently	Treat independently	Treat independently
fulltime parttime		

one of more of these AND drop a term on here

Run Query Clear 1 Group New Group

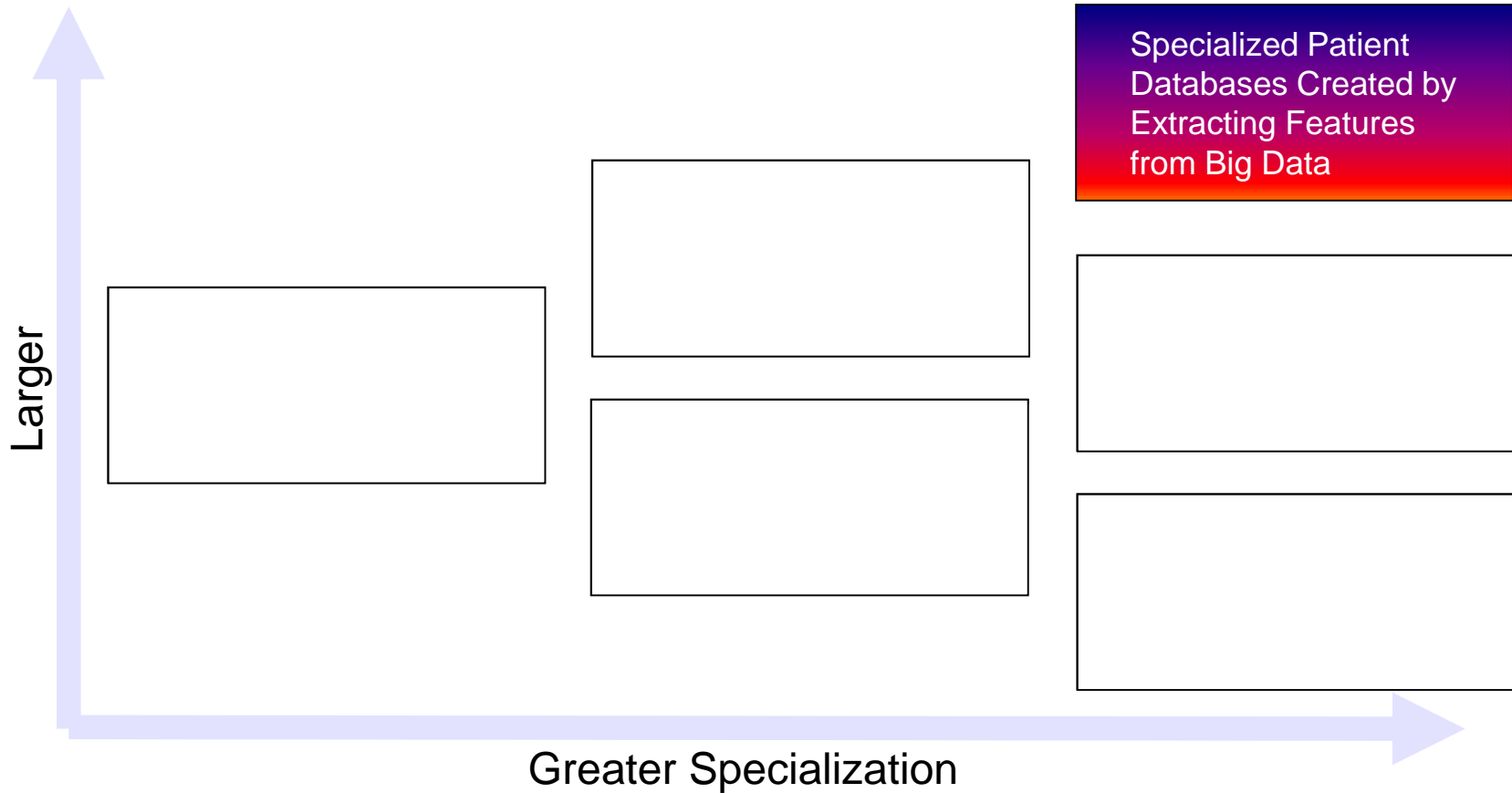
Show Query Status Graph Results Query Report

Number of patients

2

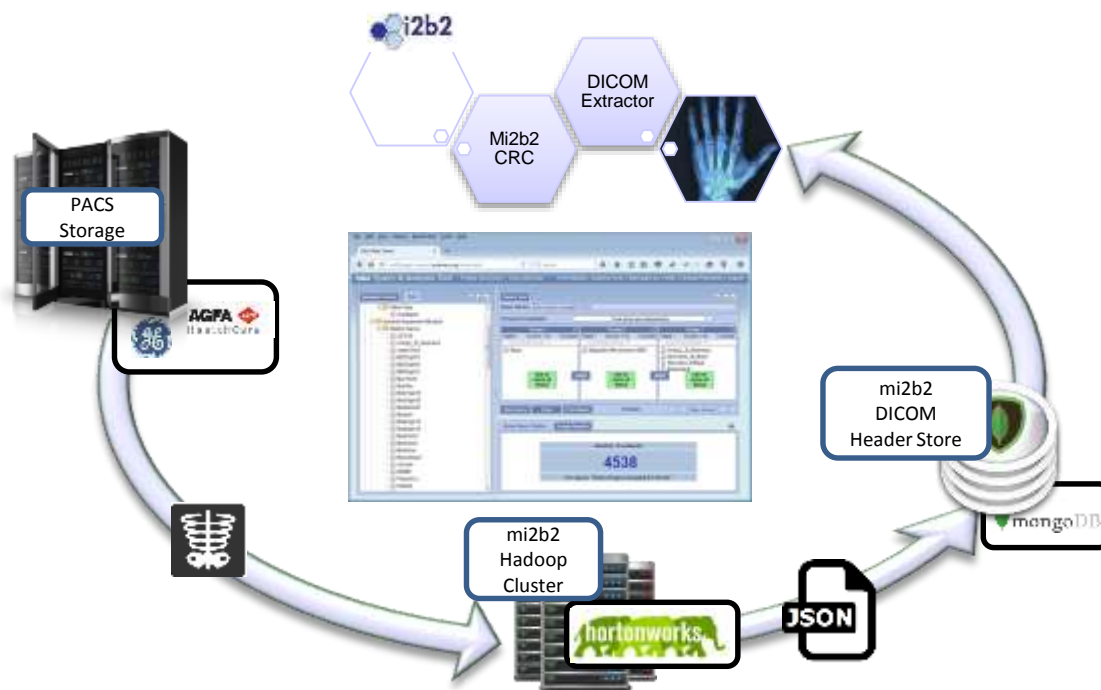
For Query "fulltime@20:17:10"

Specialized i2b2 Hives that Exist and are in Development

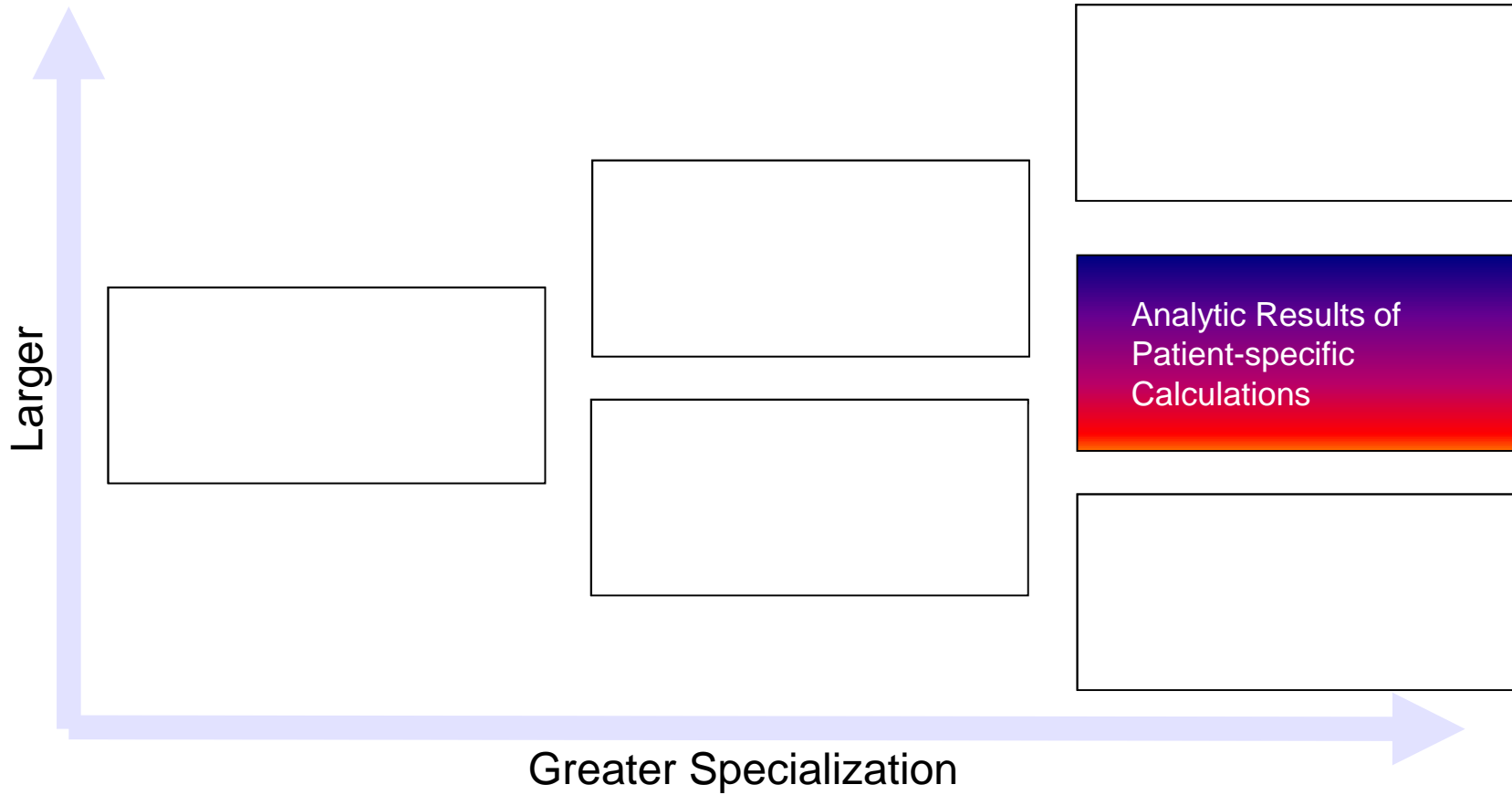


Imaging – *DICOM Index allows images of high quality to be distinguished in a PACS database*

- Investigators will be able to define sets of patients who are relevant to their research by defining the specific type of image required for their analysis (e.g. high resolution).
- Through the Big Data Commons, Investigators will be able to link this patient cohort to other available data (genomic data, biobank samples, other research data, EHR data, etc)



Specialized i2b2 Hives that Exist and are in Development



Unpredictable Quality Using Raw ICD9/10 Codes

Phenotype	Count with ICD-9/ICD-10 Code	Count (90% positive predictive value)	Count with Genotype Data
Asthma	7618	3322	805
Bipolar Disorder	1754	219	84
Breast Cancer	2101	1711	378
Congestive Heart Failure	10160	4597	1859
Coronary Artery Disease	1435	803	236
Crohn's Disease	5177	700	350
Depression	11154	4273	1074
Epilepsy	2351	1211	381
Gout	2464	1828	566
Hypertension	20788	16995	4553
Multiple Sclerosis	602	320	58
Obesity	10245	12179	3191
Rheumatoid Arthritis	3475	878	261
Schizophrenia	509	83	14
Type 1 Diabetes	2196	232	61
Type 2 Diabetes	7123	4385	1268
Ulcerative Colitis	1359	624	157

May 4, 2016, n ~ 40,000

Predictably Quality - Computed Phenotypes

Phenotype	Count with ICD-9/ICD-10 Code	Count (90% positive predictive value)	Count with Genotype Data
Asthma	7618	3322	805
Bipolar Disorder	1754	219	84
Breast Cancer	2101	1711	378
Congestive Heart Failure	10160	4597	1859
Coronary Artery Disease	1435	803	236
Crohn's Disease	5177	700	350
Depression	11154	4273	1074
Epilepsy	2351	1211	381
Gout	2464	1828	566
Hypertension	20788	16995	4553
Multiple Sclerosis	602	320	58
Obesity	10245	12179	3191
Rheumatoid Arthritis	3475	878	261
Schizophrenia	509	83	14
Type 1 Diabetes	2196	232	61
Type 2 Diabetes	7123	4385	1268
Ulcerative Colitis	1359	624	157

May 4, 2016, n ~ 40,000

Definitive Disease States Associated with Genomic Data

Phenotype	Count with ICD-9/ICD-10 Code	Count (90% positive predictive value)	Count with Genotype Data
Asthma	7618	3322	805
Bipolar Disorder	1754	219	84
Breast Cancer	2101	1711	378
Congestive Heart Failure	10160	4597	1859
Coronary Artery Disease	1435	803	236
Crohn's Disease	5177	700	350
Depression	11154	4273	1074
Epilepsy	2351	1211	381
Gout	2464	1828	566
Hypertension	20788	16995	4553
Multiple Sclerosis	602	320	58
Obesity	10245	12179	3191
Rheumatoid Arthritis	3475	878	261
Schizophrenia	509	83	14
Type 1 Diabetes	2196	232	61
Type 2 Diabetes	7123	4385	1268
Ulcerative Colitis	1359	624	157

May 4, 2016, n ~ 40,000

High Quality Data Available for Genomics Queries

The screenshot displays the Partners Biobank Portal interface. The top navigation bar includes 'PARTNERS BIOMEDICAL RESEARCH', 'BIOBANK PORTAL', and buttons for 'Find Patients', 'Make Request', 'Help & Support', and 'Watta'. The left sidebar shows a 'Navigate Terms' tree with categories like 'Biobank Consent Information', 'Biobank Demographics', 'Biobank Genomics', 'Biobank Health Information Survey', 'Biobank Sample Types', 'Curated Disease Populations', 'Healthcare Data', and 'Healthy Populations (Controls)'. The 'Curated Disease Populations' section is expanded to show 'Rheumatoid Arthritis (RA)' with sub-entries for 'RA - current or past history (PPV 0.90) - 717' and 'RA - no history (NPV 0.99) - 24086'. The main 'Query Tool' panel shows a 'Query Name' of 'RA - -Adali-Illum@15:46:12' and a 'Temporal Constraint' of 'Treat all groups independently'. Three groups are defined: Group 1 (RA - current or past history), Group 2 (Adalimumab), and Group 3 (Illumina Multi-Ethnic GWAS/Exome SNP Array). The 'Run Query' button is active, and the 'Graph Results' tab shows a bar chart with the number '70' for the query.

Query Name: RA - -Adali-Illum@15:46:12

Temporal Constraint: Treat all groups independently

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat independently			Treat independently			Treat independently		
RA - current or past history (PPV 0.90) - 717			Adalimumab - 689			Illumina Multi-Ethnic GWAS/Exome SNP Array - 4930		

Run Query **Clear** 3 Groups **New Group**

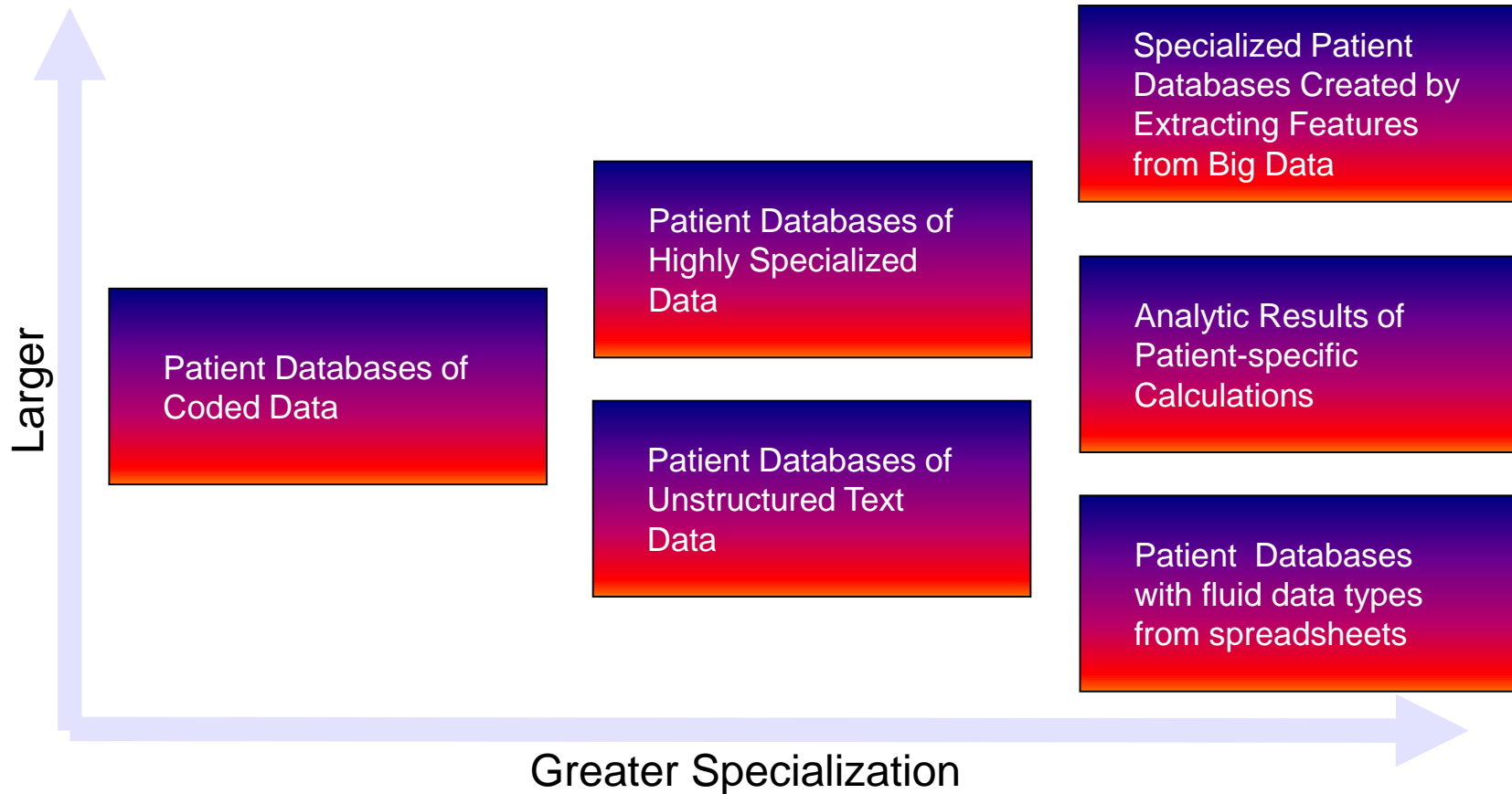
Show Query Status **Graph Results** **Query Report** **Download Data**

Number of patients

70

For Query "RA - -Adali-Illum@15:46:12"

Specialized i2b2 Hives that Exist and are in Development



The Distributed Patient Information Commons is Managed by the Ontologies

The screenshot displays the PIC-SURE i2b2 web client interface. The browser address bar shows the URL `http://i2b2-bioinformatics-partners.org/webclient/`. The page title is "PIC-SURE i2b2". The user is identified as "User: Shawn Murphy". The interface is divided into two main sections:

- Left Panel (Ontology Tree):** A hierarchical tree structure for navigating terms. The root is "Biobank Consent Information", followed by "Biobank Demographics", "Biobank Genomics", and "Biobank Health Information Survey". Under "Biobank Health Information Survey", there is a list of 15 numbered terms, including "1. Language Preference (PC:18508 FC:18509 CC:4)", "2. Age at Survey (PC:18339 FC:18339 CC:1)", "3. Education (PC:18432 FC:18433 CC:10)", "4. Employment (PC:18445 FC:34342 CC:22)", "5. Height, Weight, and BMI (PC:18223 FC:87331 CC:11)", "6. Sun Sensitivity (PC:18415 FC:18417 CC:7)", "7. Smoking (PC:18375 FC:82432 CC:99)", "8. Alcohol (PC:18194 FC:18196 CC:11)", "9. Physical Activity (PC:18189 FC:150884 CC:75)", "10. Sleep Habits (PC:16137 FC:120527 CC:17)", "11. Female Reproductive History (PC:11894 FC:72453 CC:45)", "12. Family History (PC:16026 FC:31772 CC:10)", "13. Family History of Cancer (PC:15825 FC:66408 CC:55)", "14. Family History of Heart Disease (PC:15711 FC:63838 CC:58)", and "15. Family History of Other Medical Problems (PC:15704 FC:123693 CC:52)".
- Right Panel (Query Tool):** A "Query Tool" interface for building queries. It includes a "Query Name:" field, a "Linkage Probability" slider set to 80, and a "% Certainty" field. Below these are three columns for "Group 1", "Group 2", and "Group 3". Each column has a "Dates" field, an "Occurs" field with a dropdown, and an "Exclude" checkbox. A yellow box with the text "drop a term on here" is positioned in the Group 1 area. At the bottom of the query tool are buttons for "Run Query", "Clear", "Print Query", and "New Group".

File Edit View History Bookmarks Tools Help

i2b2 Web Client

bd2k-biobankapprod.partners.org/webclient/ Search

PIC-SURE i2b2 Project: BD2K BioBank Parent User: Shawn Murphy Find Patients | Analysis Tools | Message Log | Help | Change Password | Logout

Navigate Terms Find

- Biobank Demographics
- Biobank Genomics
 - People with genomic data (PC:9358 FC:9358 CC:3)
 - Illumina Multi-ethnic Genotyping Array (MEGA) - b1_Nov2015 (PC:4930 FC:4930 CC:1)
 - Illumina Multi-ethnic Genotyping Array Expanded (MEGA Ex) - b2_May2016 (PC:4428 FC:4428 CC:1)
 - Biobank Genomics MEGA Search
 - Illumina Multi Ethnic Genotyping Array (Version 1, Batch 1) (PC:4841 FC:1698262937 CC:6)
 - dbSNP rs Identifier (PC:4841 FC:1698262937 CC:2)
 - indel (PC:4841 FC:4119726 CC:1)
 - SNP (PC:4841 FC:1694143211 CC:1)
 - Gene (PC:4841 FC:1698262937 CC:2)
 - indel (PC:4841 FC:4119726 CC:1)
 - SNP (PC:4841 FC:1694143211 CC:1)
- Biobank Health Information Survey
- Biobank Sample Types
- Curated Disease Populations
 - Asthma (AST) (PC:37632 FC:37859 CC:3)
 - Bipolar Disorder (BD) (PC:38184 FC:38183 CC:4)
 - Breast Cancer (BRCA) (PC:39573 FC:39812 CC:3)
 - Chronic Obstructive Pulmonary Disease (COPD) (PC:38709 FC:38708 CC:3)
 - Congestive Heart Failure (CHF) (PC:36725 FC:36724 CC:4)
 - Coronary Artery Disease (CAD) (PC:46428 FC:46427 CC:6)
 - Crohn's Disease (CD) (PC:39573 FC:39812 CC:3)
 - Depression (DEP) (PC:39351 FC:39587 CC:3)
 - Epilepsy (EPIL) (PC:39573 FC:39812 CC:3)
 - Epilepsy - current or past history (PPV 0.90) (PC:1214 FC:1223 CC:1)
 - Epilepsy - no history (NPV 0.99) (PC:38359 FC:38589 CC:1)

Query Tool

Query Name:

Linkage Probability > % Certainty

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
Epilepsy - current or past history (PPV 0.90) (PC:1214 FC:1223 CC:1)			Illumina Multi-ethnic Genotyping Array (MEGA) - b1_Nov2015 (PC:4930 FC:4930 CC:1) Illumina Multi-ethnic Genotyping Array Expanded (MEGA Ex) - b2_May2016 (PC:4428 FC:4428 CC:1)			NOT SNP [contains "rs6054 AND Homozygous AND C_to_T AND missense"] (PC:4841 FC:1694143211 CC:1)		
one or more of these			AND			one or more of these		
			AND			none of these		

Run Query Clear Print Query 3 Groups New Group

Show Query Status Graph Results

Number of patients

143

For Query "Epileps-Illumin@23:06:21"

Getting Data from the Patient Information Commons

Using spreadsheets to extract matrixes of data

Using data transfer from one node to another to run the analysis on one node

Mega-analysis by Julia with direct access to data

Using SMART to get insights into clinical care

Partners Biobank Portal – Specify Data for Download

Patient Set Viewer View Result Table

Setting up Parameters > [View Result Table](#)

1. Drop a *Patient Set* or a *Previous Query* into the Patient Set box. The ontology terms (*Concepts*) used in the *Previous Query* will automatically populate the Concepts box.
2. Additional *Concepts* can be dropped into the same Concepts box by dragging them from the Navigate Terms panel or the Workplace panel to the left.
3. Change the default options by clicking on the Options box to suit your display needs.
4. Click on Create Table to start getting the Patients and the Concepts you entered.

Patient Set:

Diabetes mellit@12:09:29 [10-5-2015] [demo]	Size = 22093
---	--------------

Concept(s):

Diabetes mellitus	Count <input type="text" value=""/>
Family history [Contains: diabetes]	Count <input type="text" value=""/>
HGB A1C (LOINC:4548-4)	Count <input type="text" value=""/>
HGB A1C (LOINC:4548-4)	Min <input type="text" value=""/>
HGB A1C (LOINC:4548-4)	Max <input type="text" value=""/>
HGB A1C (LOINC:4548-4)	Average <input type="text" value=""/>

Options:

Starting with the 1st Patient, retrieve 100 Patients

Table Size: Max = 200, Min = 1

Starting with the st Patient

Create Table ▶

Download table and select patients

0 / 100 Selected

ID	Gender	Age	Race	Diabetes mellitus	Family history (Contains Diabetes)	HGB A1C (LOINC:4548-4)	HGB A1C (LOINC:4546-4)	HGB A1C (LOINC:4546-4)
1000000003	Male	41	asian	12	54	43	4.1	6.9
1000000008	Male	23	hispanic	3	12	54	4.2	6.7
1000000011	Female	60	white	21	21	43	4.6	7.1
1000000013	Female	83	black	34	9	65	4.7	5.9
1000000015	Male	32	hispanic	1	10	34	4.4	6.3
1000000029	Male	25	hispanic	3	25	25	5.1	6.8
1000000043	Male	61	hispanic	23	14	21	4.9	6.5
1000000045	Female	89	asian	37	15	54	4.9	6.6
1000000054	Female	66	asian	43	6	45	4.8	6.6
1000000066	Female	22	black	32	4	33	5.2	7.0
1000000070	Male	38	black	43	11	38	5.5	6.5
1000000077	Male	13	black	12	8	42	4.4	6.7
1000000080	Male	33	indian	32	13	48	4.2	6.5
1000000083	Female	65	black	33	4	39	4.7	6.6
1000000087	Male	55	asian	36	6	42	4.8	6.9
1000000096	Female	76	black	54	17	43	4.9	6.5
1000000099	Male	36	black	32	0	39	5.2	6.6
1000000108	Male	66	asian	43	1	25	5.3	6.7
1000000109	Male	63	hispanic	31	4	29	4.9	5.9
1000000112	Male	17	black	17	0	14	4.7	6.8
1000000119	Female	72	white	43	32	45	4.4	5.9
1000000133	Female	16	asian	26	3	75	4.9	6.6

Showing Patients 1 to 100 of total 22093

Tribute to...

- RPDR/I2b2 Core Team

- Christopher Herrick
- Michael Mendis
- Lori Phillips
- Janice Donahoe
- Nich Wattanasin
- Wayne Chan
- Vivian Gainer
- Alyssa Goodson
- Mariah Mitchell
- Martin Rees
- Charles Wang
- Laurie Bogosian
- Stacey Duey
- Andrew Cagan
- David Wang

- Biobank Team

- Natalie Boutin
- Victor Castro
- Scott Weiss
- Beth Karlson

- SMART Team

- Ken Mandl
- Josh Mandel
- Kavi Wagholikar

- Genomics Innovation Team

- Sandy Aronson
- Heidi Rehm
- Calum MacRea

... and many more.



Centre for Health Technologies
University of Pavia

4th i2b2 EU AUG Meeting

Sept 15th 14:30 - Sept 16th 12:30

Pavia – ITALY

Also featuring

Ulli Prokosch
Ulrich Sax
Christel Daniel, ...

further info <http://cht.unipv.it>

Invited Talks by
Paul Avillach
Shawn Murphy

Sponsored by:





FOUNDED BY BRIGHAM AND WOMEN'S HOSPITAL
AND MASSACHUSETTS GENERAL HOSPITAL



I2b2 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>)