

# Employing SNOMED CT and LOINC to make EHR data sensible and interoperable for clinical research

James R. Campbell MD

W. Scott Campbell PhD

Hubert Hickman MS

James McClay MD

Implementation Showcase

October 31, 2014



# Disclosure

We receive no financial support from any interests vending electronic health records or vocabulary services



# Agenda

- Interoperability for research and public health; framework of ONC
- Clinical research data networks: PCORI
- Tooling and communication protocols for big data research; i2b2
- Greater Plains Cooperative and deployment of data standards
- Integrating SNOMED CT and LOINC observables
- Expectations of LOINC-SNOMED CT harmonization



# Meaningful Use Vocabulary Requirements

- Demographics (80% stage 2): HL7/OMB code set
- Smoking status: SNOMED CT
- Problem list/encounter diagnoses: SNOMED CT, ICD\*
- Structured lab results (55% stage 2): Lab LOINC
- Physical findings: LOINC, SNOMED CT observables
- E-prescribing (50% formulary check stage 2): RxNORM
- Medications: RxNORM
- Orders: LOINC, RxNORM
- Immunizations (Immunization registries): CVX, MVX
- Procedures: CPT, HCPCS
- Documents: LOINC





# ONC HIT: Objectives of Interoperability

- Facilitate transitions of care
- Promote free flow of results data
- Engage patients and families in managing their EHR
- Support Public Health
- Contribute to research within Learning Healthcare Environment



# Interoperable Learning Health Ecosystem



**How to merge data from the EHR into integrated, interoperable research datasets?**



# PCORnet: the National Patient-Centered Clinical Research Network



The goal of PCORI's National Patient-Centered Clinical Research Network Program is to improve the nation's capacity to conduct CER efficiently, by creating a large, highly representative, national patient-centered clinical research network for conducting clinical outcomes research.

The vision is to support a learning US healthcare system, which will allow for large-scale research to be conducted with enhanced accuracy and efficiency.

PCORnet will function through a single, interoperable data research network, the PCORnet DRN.



# What is Patient Centered Outcome Research Institute(PCORI)?

- Funded by trust established Affordable Care act (2010); \$650M per year through 2015
- Supports both Clinical Data Networks (CDRN) and Patient Research networks (PPRN)
- Recent funding cycle concluded 2013 with 11 CDRNs and 18 PPRNS

## PCORI Funding Awards

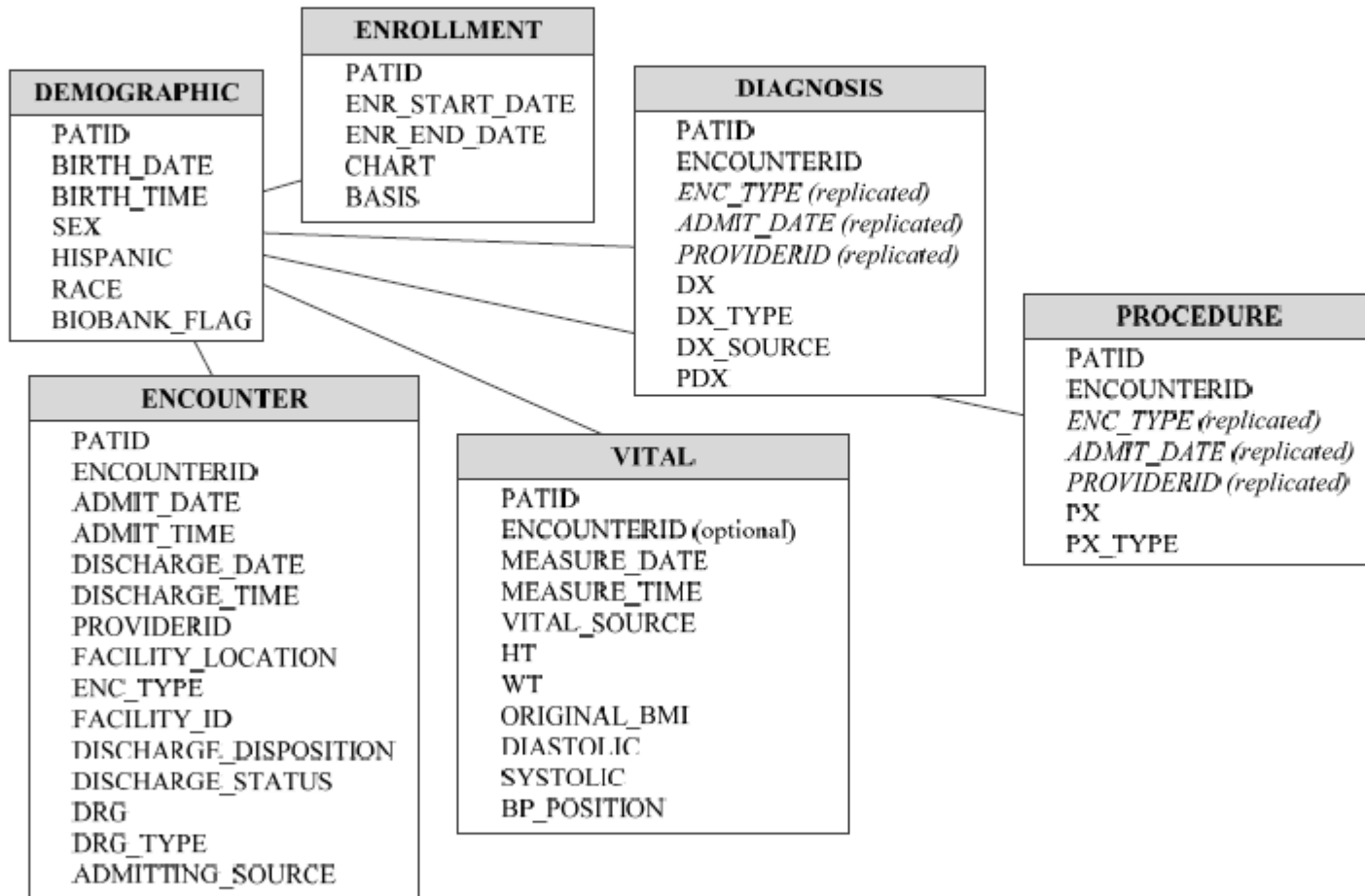


PCORI has approved 279 awards totaling more than \$464.4 million to fund patient-centered comparative clinical effectiveness research projects to date.

# Clinical Data Research Network

- Collaboration of (mostly) academic medical centers supported by network funding from PCORI
- Expected to standardize research data access across sites, establish query communication function and share data access across the CDRN and PCORNET
- Research interoperability = Reference information model + (Content standards) + Popmednet network query tools

# PCORnet Common Data Model v1.0





# Tooling for 'Big Data'

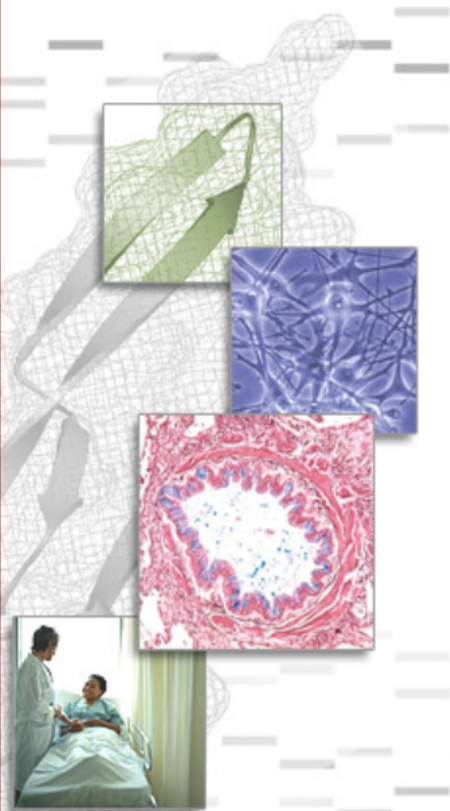
## i2b2

Informatics for Integrating Biology & the Bedside

A National Center for Biomedical Computing

i2b2: Informatics for Integrating Biology & the Bedside - A National Center for Biomedical Computing

[About Us](#) | [Driving Biology Projects](#) | [Software](#) | [Resources](#) | [Events](#) | [Training](#) | [News](#) | [Collaborations](#) | [Publications](#)



### MISSION

i2b2 (Informatics for Integrating Biology and the Bedside) is an NIH-funded National Center for Biomedical Computing based at Partners HealthCare System. The i2b2 Center is developing a scalable informatics framework that will enable clinical researchers to use existing clinical data for discovery research and, when combined with IRB-approved genomic data, facilitate the design of targeted therapies for individual patients with diseases having genetic origins. This platform currently enjoys wide international adoption by the CTSA network, academic health centers, and industry. i2b2 is funded as a cooperative agreement with the National Institutes of Health.

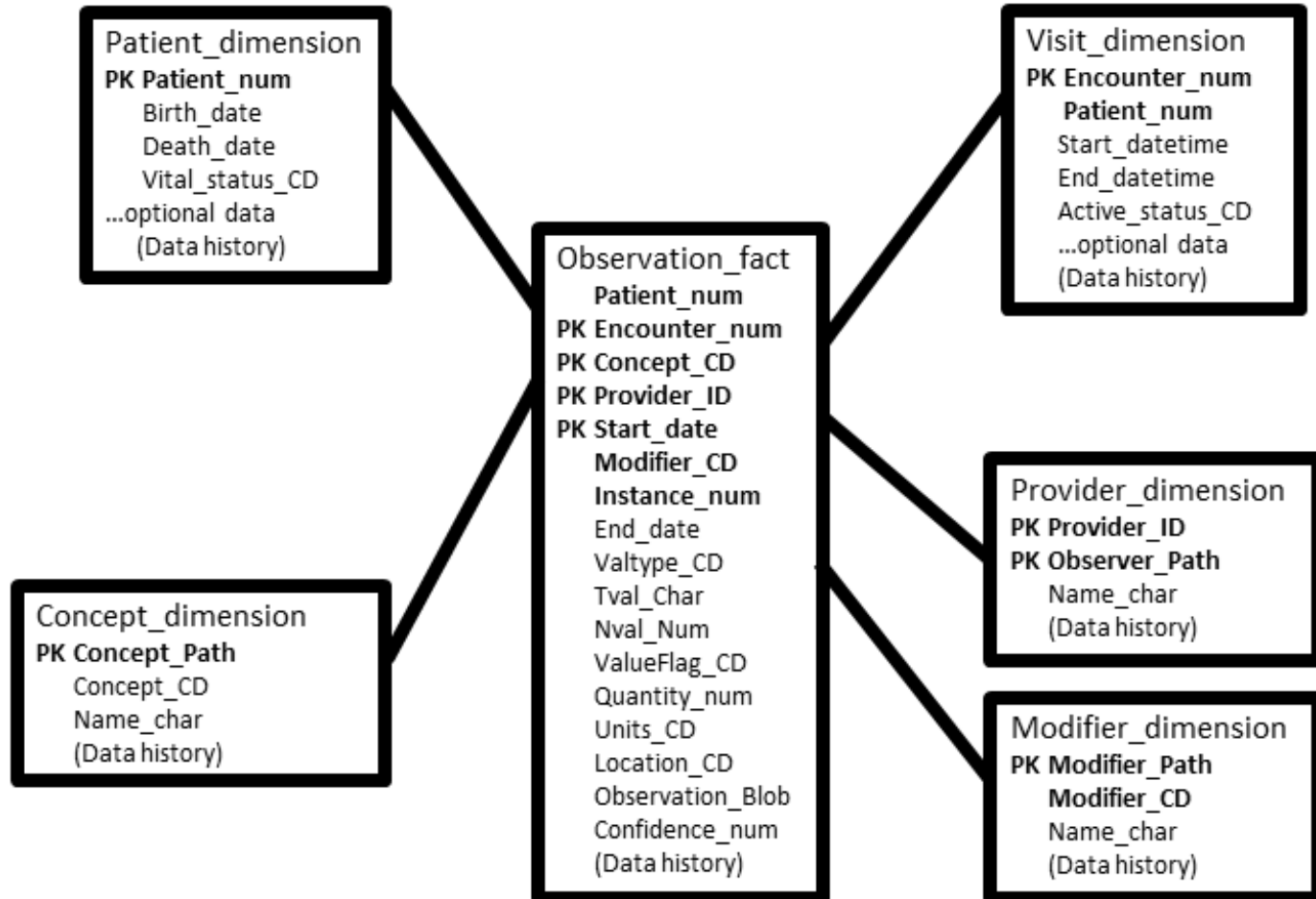
### DRIVING BIOLOGY PROJECTS

- ⌘ Overview
- ⌘ Current DBPs
  - ⌘ Autoimmune/CV Diseases
  - ⌘ Diabetes/CV Diseases
- ⌘ Past DBPs
  - ⌘ Airways Diseases
  - ⌘ Hypertension
  - ⌘ Type 2 Diabetes Mellitus
  - ⌘ Huntington's Disease
  - ⌘ Major Depressive Disorder
  - ⌘ Rheumatoid Arthritis
  - ⌘ Obesity

### RESOURCES

- ⌘ Overview
- ⌘ Computational Tools
- ⌘ De-Identification Demo
- ⌘ Software
- ⌘ NLP Research Data Sets
- ⌘ NLP Shared Tasks
- ⌘ Academic Users' Group
- ⌘ Publication Data

# i2b2 'Out of the box'





# I2b2 workbench

**i2b2 Query & Analysis Tool** Project: I2b2 Demo User: James R. Campbell Find Patients | Analysis Tools | Message Log | Help | Logout

**Navigate Terms** Find Terms

- Clinical finding (finding)
  - Administrative statuses (finding)
  - Adverse incident outcome categories (finding)
  - Bleeding (finding)
  - Calculus finding (finding)
  - Clinical history and observation findings (finding)
  - Clinical stage finding (finding)
  - Cyanosis (finding)

**Workplace**

- SHARED
- campbell

**Previous Queries**

- Age Range [1,73@08:09:25 [9-21-2014] [campbell]
- Age Range [1,73@08:08:46 [9-21-2014] [campbell]
- Encounter Type @15:38:51 [9-14-2014] [campbell]
- Admission sourc@15:37:18 [9-14-2014] [campbell]
- Admission sourc@15:36:39 [9-14-2014] [campbell]
- Admission sourc@15:35:54 [9-14-2014] [campbell]

**Query Tool**

Query Name: \_\_\_\_\_

Temporal Constraint: \_\_\_\_\_

Groups independently

Exclude	Dates	Occurs > 0x	Exclude
	Treat Independently		

one or more of these AND drop a term on here

Run Query Clear Print Query 1 Group New Group

**Query Status**

150%

# i2b2 workbench

**Query  
formulation  
tool**

The screenshot shows the i2b2 Query & Analysis Tool interface. At the top, the browser window displays the URL and various tabs. The tool's header includes the title "i2b2 Query & Analysis Tool", the project name "Project: i2b2 Demo", the user "User: James R. Campbell", and navigation links for "Find Patients", "Analysis Tools", "Message Log", "Help", and "Logout".

The main interface is divided into several sections:

- Navigate Terms:** A sidebar on the left with a tree view of clinical terms, including "Clinical finding (finding)", "Administrative statuses (finding)", "Adverse incident outcome categories (finding)", "Bleeding (finding)", "Calculus finding (finding)", and "Clinical history and observation findings (finding)".
- Query Tool:** The central area for defining queries. It features a "Query Name" field, a "Temporal Constraint" dropdown set to "Treat all groups independently", and three query groups (Group 1, Group 2, and Group 3). Each group has a "Dates" field, an "Occurs > 0x" field, and an "Exclude" checkbox. Group 1 contains the term "Body mass index [1,662,134 facts; 252,442]". Below the groups are buttons for "one or more of these", "AND", and "drop a term on here". At the bottom of the query tool are "Run Query", "Clear", and "Print Query" buttons, along with a "New Group" button.
- Previous Queries:** A list at the bottom left showing a history of queries, such as "Age Range [1,73@08:09:25 [9-21-2014] [campbell]", "Age Range [1,73@08:08:46 [9-21-2014] [campbell]", "Encounter Type @15:38:51 [9-14-2014] [campbell]", "Admission sourc@15:37:18 [9-14-2014] [campbell]", "Admission sourc@15:36:39 [9-14-2014] [campbell]", and "Admission sourc@15:35:54 [9-14-2014] [campbell]".
- Query Status:** A section at the bottom right for monitoring query execution.



# i2b2 workbench

**Query history and datasets**

# CDRN 4: Greater Plains Collaborative

- University of Kansas
- U Wisconsin Madison
- Medical College Wisconsin
- Childrens Mercy (KC)
- Marshfield Clinic
- University of Iowa
- University of Minnesota
- University of Nebraska
- U Texas Southwestern
- U Texas San Antonio



# GPC Architecture for Research Interoperation

- Research data integration and query management tooling: i2b2
- Information model: Star schema
- Domain ontology/code sets:
  - Demographics, Clinical findings/biometrics, Lab findings, Radiology findings, Diagnoses, Allergies, Procedures, Orders/procedure/medications, Medications/pharmaceuticals administered, Registry data
- Value sets for coded data



# (Some) Research Data Warehouse Requirements

- Organization, presentation and query data access understandable to the clinical researcher
- Data aggregation queries easy to do; relevant to research questions; standards driven but not requiring knowledge of the standard
- Query of quantitative results and conversion of data units must be supported for heterogeneous data sources
- Interoperability standards inherent to design but not obtrusive or confusing



# Research (QA) Data Warehouse: Query use cases

- How many diabetics have had a glycosylated hemoglobin test?
- How many hypertensive patients have a recent blood pressure below 140/90?
- Which patients with ALS have a drooling score  $>5$  ?
- What fraction of children who had BMI %ile  $>95$  at any time are obese at age 30?





# GPC: i2b2 Reference Ontologies

Top level ontology			Concept dimension code sets
<b>Reference standard ontologies</b>			
i2b2\Demographics	LOINC(clinical)		Age 30525-0.....
	SNOMED CT (observables)		Religion SNOMED CT:160538000.....
i2b2\Diagnoses	ICD9CM		Billing; encounter diagnoses
	SNOMEDCT (clinical findings, events, situations)		Problem list; PMH
	ICD10CM		Billing; encounter diagnoses
\Lab measurements	LOINC (Lab)		All labs
\Clinical Measurements	LOINC (Clinical)		Vital signs, physical findings
	SNOMED CT (Observables)		Physical findings and symptoms
\Procedures	HCPCS level 1: CPT4		Professional services
	HCPCS level 2: ICD9CM		Hospital procedures and services
	HCPCS level 3		Administrative procedures
	SNOMEDCT (Procedures)		Surgical history
\Medications	NDF-RT; RxNORM		Medication orders, administered
	NDC		Medications dispensed
<b>UNMC Enterprise data</b>			
\UNMC Demographics			LOINC(Clinical); SNOMEDCT
\UNMC Common labs\	\Hematology \Chemistry \Microbiology ...		LOINC(Laboratory)
UNMC Diagnoses\Problem list			SNOMED CT(Clinical findings, events,



# UNMC: i2b2 Clinical Ontologies

<b>UNMC Enterprise data</b>					
\UNMC Demographics					LOINC(Clinical); SNOMEDCT
\UNMC Common labs\	\Hematology \Chemistry \Microbiology ...				LOINC(Laboratory)
UNMC Diagnoses\Problem list					SNOMED CT(Clinical findings, events,
UNMC Diagnoses\Encounter diagnoses					ICD-9-CM; ICD-10-CM
\UNMC Radiology					Clinical LOINC; SNOMED CT; CPT
\UNMC Procedures					CPT; HCPCS; SNOMED CT
\Medications administered					NDF-RT; RxNORM; NDC
\Medications dispensed					(NDF-RT+ RxNORM); NDC
\Immunizations administered					CVX; MVX
UNMC Orders\Procedure orders					CPT; HCPCS; SNOMED CT; LOINC
UNMC Orders\Medication Orders					NDF-RT; RxNORM; NDC
UNMC Orders\Prescriptions					NDF-RT; RxNORM
UNMC\Allergies&Sensitivities\Reaction type					Allergens(SNOMED CT); Reaction(SNOMED CT); Severity(SNOMED CT)
<b>PCORI metadata</b>					
\i2b2\PCORI\					
\i2b2\PCORI\Demographics\	Demographics				
\i2b2\PCORI\Demographics\Patient identifier\	Patient identifier	LOINC:56794-1			
\i2b2\PCORI\Demographics\Biobank flag\	Biobank flag	LOINC:75520-7			
\i2b2\PCORI\Demographics\Gender\	Gender	LOINC:72143-1			
\i2b2\PCORI\Demographics\Birth date\	Birth date	LOINC:21112-8			



# GPC:

# PCORI CDM Mapped Ontology

PCORI metadata		
\i2b2\PCORI\Demographics\Patient identifier\	Patient identifier	LOINC:56794-1
\i2b2\PCORI\Demographics\Biobank flag\	Biobank flag	LOINC:75520-7
\i2b2\PCORI\Demographics\Gender\	Gender	LOINC:72143-1
\i2b2\PCORI\Demographics\Birth date\	Birth date	LOINC:21112-8
\i2b2\PCORI\Demographics\Birth time\	Birth time	LOINC:57715-5
\i2b2\PCORI\Demographics\Race\	Race	LOINC:72826-1
\i2b2\PCORI\Demographics\Hispanic\	Hispanic	LOINC:69490-1
\i2b2\PCORI\Vitals\Vitals source\	Source of vitals data	LOINC:75521-5
\i2b2\PCORI\Vitals\Blood pressure\	Blood pressure	
\i2b2\PCORI\Vitals\Blood pressure\Raw data\		
\i2b2\PCORI\Vitals\Blood pressure\Systolic pressure\	Systolic blood pressure	LOINC:8480-6
\i2b2\PCORI\Vitals\Blood pressure\Diastolic pressure\	Diastolic blood pressure	LOINC:8462-4
\i2b2\PCORI\Vitals\Blood pressure\BP_position\	BP Position	LOINC: 8361-8
\i2b2\PCORI\Vitals\BMI\	BMI	LOINC:39156-5
\i2b2\PCORI\Vitals\BMI percentile\	BMI percentile	LOINC:59574-4
\i2b2\PCORI\Vitals\Height\		LOINC:8302-2
\i2b2\PCORI\Vitals\Weight\		LOINC:3141-9
\i2b2\PCORI\Diagnosis\ICD9CM/Encounter)		ICD-9-CM/Encounter or Billing diagnoses
\i2b2\PCORI\Procedure\CPT4; ICD9CM)		HCPCS(CPT4, ICD-9-CM)
\i2b2\PCORI\Encounters\Encounter identifier\	Admission date	LOINC:75519-9
\i2b2\PCORI\Encounters\Admit date\	Admission date	LOINC:52455-3
\i2b2\PCORI\Encounters\Admit time\	Admission time	LOINC:75522-3

# Agenda

- Interoperability for research and public health; framework of ONC
- Clinical research data networks: PCORI
- Tooling and communication protocols for big data research; i2b2
- Greater Plains Cooperative and deployment of data standards
- Integrating SNOMED CT and LOINC observables
- Expectations of LOINC-SNOMED CT harmonization





# Metataxml required for units conversion and numeric queries

i2b2 Query & Analysis Tool      Project: i2b2 Demo      User: James R. Campbell      Find Patients | Analysis Tools | Message Log | Help | Logout

**Navigate Terms**    **Find Terms**

- National emergency medical services information system (nemsis) [2,209,100 facts; 2,209,100 patients] - 2209100
- Neonatal apgar status
- Obstetrical studies and measures
- Ophthalmology/optometry variables
- Organ transplant
- Patient health record panels [27,733,793 facts; 455,134 patients] - 455134
- Patient safety
- Radiology studies
- Respiratory measures and ventilator management [4,683,593 facts; 1,111,111 patients] - 1111111
- Survey instruments [53,674,616 facts; 508,345 patients] - 508345
- Tumor registry (naaccr)
- Urology studies and measures
- Vaccination records
- Vital signs [36,947,763 facts; 270,970 patients] - 270970
  - Blood pressure measurements [16,453,392 facts; 249,111 patients] - 249111
    - Blood pressure panels [10,968,928 facts; 249,111 patients] - 249111
      - 34553-8: Orthostatic blood pressure panel [10,968,928 facts; 249,111 patients] - 249111
      - 35094-2: Blood pressure panel [10,968,928 facts; 249,111 patients] - 249111
        - 8357-6: Blood pressure method
        - 8358-4: Blood pressure device cuff size
        - 8359-2: Peripheral artery measurement
        - 8462-4: Diastolic blood pressure [5,484,464 facts; 5,484,464 patients] - 5484464
        - 8478-0: Mean blood pressure
        - 8480-6: Systolic blood pressure [5,484,464 facts; 5,484,464 patients] - 5484464
        - 9856-6: Blood pressure device institution
        - 9857-4: Blood pressure device vendor name
        - 9858-2: Blood pressure device vendor name
  - Central
  - General blood pressure [16,453,392 facts; 249,111 patients] - 249111
  - Position
  - Specific blood pressure
  - Timed specific blood pressure

**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 ▾		
			8480-6: Systolic blood pressure [5,484,464 facts; 5,484,464 patients] - 5484464					

**Choose value of Systolic blood pressure (Test:LOINC:8480-6)**

Searches by Lab values can be constrained by the high/low flag set by the performing laboratory, or by the values themselves.

No Value      Please select operator: **GREATER THAN (>)**

By Value      Please enter a value: **140**

Click on a bar segment to help specify a value or range:  
Range in mm Hg

Units = **mm Hg**

**OK**    **Cancel**

# Local taxonomies for research user: labor intensive for LOINC

File Edit View Favorites Tools Help

i2b2 Query & Analysis Tool Project: i2b2 Demo User: James R. Campbell Find Patients | Analysis Tools | Message Log | Help | Logout

Navigate Terms Find Terms

- Clinical Measurements [67,475,602 facts; 518,830 patients]
- DRG [34,356 facts; 23,199 patients]
- Demographics [50,929,758 facts; 1,737,382 patients]
- Diagnoses (ICD9CM) [6,887,084 facts; 498,807 patients]
- Encounters [34,576,070 facts; 954,980 patients]
- Laboratory Measurements [146,001,219 facts; 544,898 patients]
- PCORI [99,487,439 facts; 1,736,608 patients]
- Procedures [5,458,439 facts; 233,641 patients]
- SNOMEDCT
- UNMC One Chart [31,915,785 facts; 415,134 patients]
  - Laboratory tests [31,915,785 facts; 415,134 patients] - 415134
    - Hematology [26,109,507 facts; 390,463 patients] - 390463
      - Complete blood count without differential [26,109,507 facts; 390,463 patients] - 390463
        - Hematocrit [3,606,065 facts; 379,643 patients] - 379643
        - Hemoglobin [3,644,385 facts; 388,871 patients] - 388871
        - MCH
        - MCHC [3,130,745 facts; 369,826 patients] - 369826
        - MCV [3,135,238 facts; 369,904 patients] - 369904
        - Platelet count [3,167,093 facts; 364,792 patients] - 364792
        - RDW [3,135,209 facts; 369,902 patients] - 369902
        - Red blood cell count [3,141,903 facts; 370,004 patients] - 370004
        - White blood cell count [3,148,869 facts; 370,003 patients] - 370003
    - Urinalysis [5,806,278 facts; 229,608 patients] - 229608

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		

drop a term on here

Run Query Clear Print Query 0 Groups New Group

Query Status



# Task: Map Research Data Model to Reference Ontologies

i2b2 Query & Analysis Tool      Project: i2b2 Demo      User: James R. Campbell      Find Patients | Analysis Tools | Message Log | Help | Logout

**Navigate Terms**      Find Terms

- Laboratory Measurements [146,001,219 facts; 544,898 patients]
- PCORI [99,487,439 facts; 1,736,608 patients]
  - Demographics [12,155,584 facts; 1,736,512 patients] - 1736512
    - Age [1,736,512 facts; 1,736,512 patients] - 1736512
    - Biobank flag [1,736,512 facts; 1,736,512 patients] - 1736512
    - Birth date [1,736,512 facts; 1,736,512 patients] - 1736512
    - Birth time [1,736,512 facts; 1,736,512 patients] - 1736512
    - Gender [1,736,512 facts; 1,736,512 patients] - 1736512
    - Hispanic [1,736,512 facts; 1,736,512 patients] - 1736512
    - Race [1,736,512 facts; 1,736,512 patients] - 1736512
  - Diagnosis [7,085,512 facts; 245,317 patients] - 245317
  - Encounter [64,386,044 facts; 920,460 patients] - 920460
    - Admission datetime [31,760,976 facts; 920,460 patients] - 920460
    - Admission source [208,854 facts; 95,045 patients] - 95045
    - Diagnosis related group [34,483 facts; 23,262 patients] - 23262
    - Discharge datetime [205,831 facts; 94,438 patients] - 94438
    - Discharge disposition [209,305 facts; 95,313 patients] - 95313
    - Discharge status [205,619 facts; 94,370 patients] - 94370
    - Encounter Type [31,760,976 facts; 920,460 patients] - 920460
    - Encounter provider
    - Facility identifier
    - Facility location
  - Procedure
  - Vitals [15,860,299 facts; 292,865 patients] - 292865
    - Blood pressure [10,968,928 facts; 249,139 patients] - 249139
    - Body mass index [1,662,134 facts; 252,442 patients] - 252442
    - Height [1,193,630 facts; 228,253 patients] - 228253
    - Weight [2,035,607 facts; 265,189 patients] - 265189
  - Procedures [5,458,439 facts; 233,641 patients]
  - SNOMEDCT
  - UNMC One Chart [31,915,785 facts; 415,134 patients]
  - Laboratory tests [31,915,785 facts; 415,134 patients] - 415134

**Query Tool**

Query Name: \_\_\_\_\_

Temporal Constraint:

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
drop a term on here								

Run Query      Clear      Print Query      0 Groups      New Group

**Query Status**

**LOINC:72826-1|Race|Type|Pt|^Patient|OMB.1997|**



# Problems with deploying Observables

- Release data sets from IHTSDO and Regenstrief clumsy to deploy and not harmonized; user requirements of aggregation for i2b2 better served with transitive closure tables
- Organizing infrastructure with NRC for maintenance release cycles TBD
- Metadatasets required to manage datatype, units conversion and normals range must be developed



# Research Data Requirements of LOINC – SNOMED CT Harmonization

- Make sure the world hears about this!!! This is a clandestine op as far as the informatics community is concerned!
- Publish as a fully classified expression dataset that will support local ontology deployment
- Agreement between principles and publication of uniform identifier for ALL expression elements
- Grouper concepts to support aggregation by lab department, sub-department and panel code; clinical observables organized by findings class
- Coordination of release datasets with NRCs to assure useful datasets (transitive closure tables) and software for deployment

Questions?  
Comments?



# Observation fact: Race of patient

