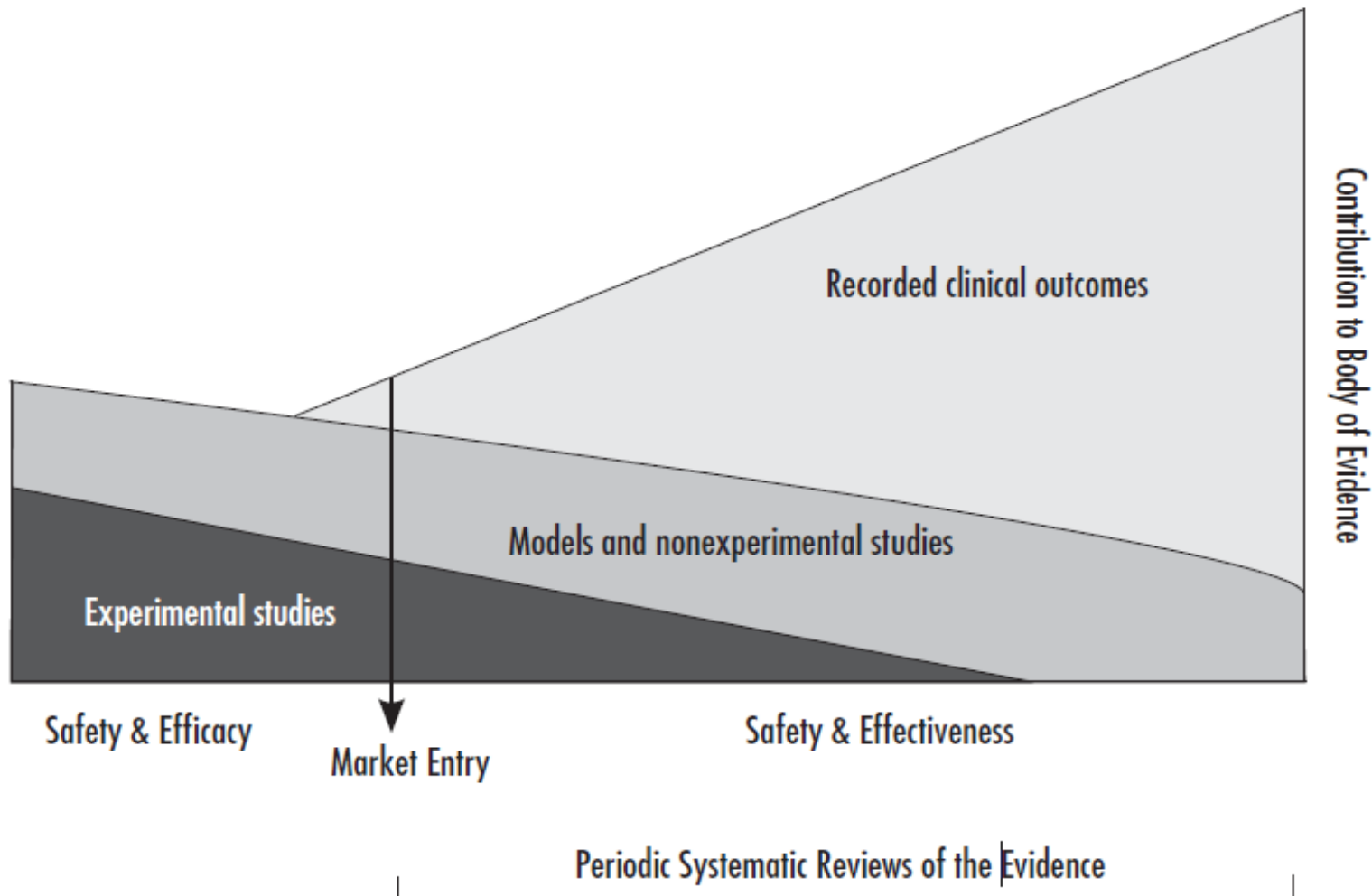


Data Integration in the SCILHS PCORNet Clinical Data Research Network

Shawn Murphy, Ken Mandl, Sebastian Schneeweiss
Principle Investigators

Learning Health System: Evidence Generated from a Well-Instrumented Health System



Patient Centered Outcomes Research Institute

- ❑ An **independent, non-profit health research organization** authorized by the Patient Protection and Affordable Care Act of 2010.
- ❑ PCORI funds **patient-centered research** to assist patients, caregivers, and other stakeholders in making informed health decisions.

Mission

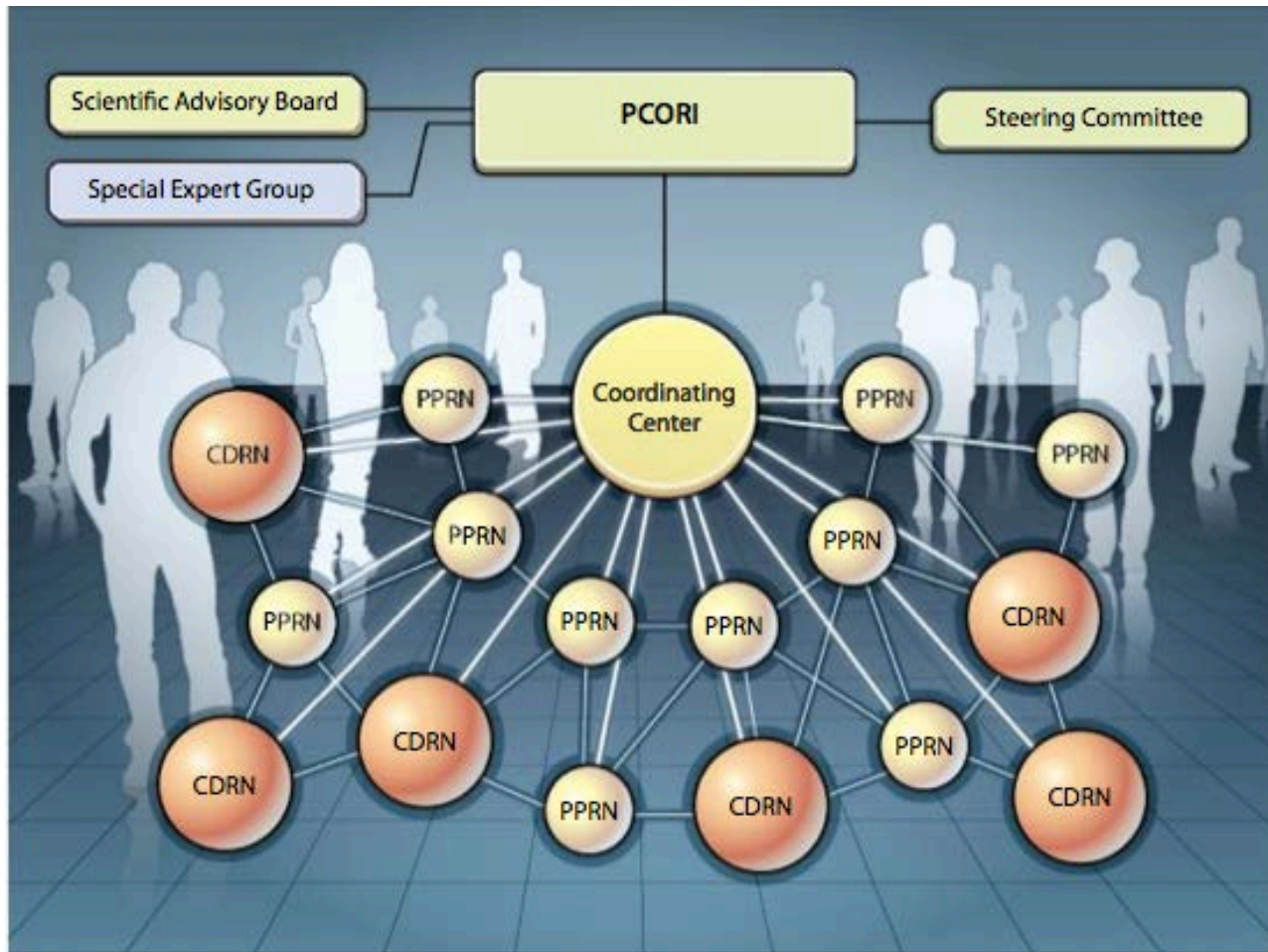
PCORI helps people make informed healthcare decisions and improves healthcare delivery and outcomes by producing and promoting high integrity, evidence-based information that comes from research **guided by patients, caregivers, and the broader healthcare community.**



Vision

Patients and the public have the information they need to **make decisions** that reflect their **desired health outcomes.**

The National Network of Networks



Clinical Data Research Networks

The Clinical Data Research Networks (CDRNs) will develop the capacity to conduct randomized comparative effectiveness studies using data from clinical practice in a large, defined population. These established or newly developed networks involve two or more healthcare systems, with plans to function as integrated research network.

i2b2

- **C1_Scalable Collaborative Infrastructure for a Learning Healthcare System (SCILHS)**
Harvard University

- **C2_Mid-South CDRN**
Vanderbilt University

i2b2

- **C3_Patient-oriented SCALable National Network for Effectiveness Research (pSCANNER)**
University of California, San Diego (UCSD)

i2b2

- **C4_Great Plains Collaborative (GPC)**
University of Kansas Medical Center

- **C5_Kaiser Permanente & Strategic Partners Patient Outcomes Research To Advance Learning (PORTAL) Network**
Kaiser Foundation Research Institute

- **C6_Louisiana CDRN (LACDRN)**
Louisiana Public Health Institute (LPHI)

i2b2

- **C7_National Pediatric Learning Health System (PEDSNet)**
The Children's Hospital of Philadelphia

- **C8_New York City Clinical Data Research Network (NYC-CDRN)**
Weill Medical College of Cornell University

- **C9_Chicago Area Patient Centered Outcomes Research Network (CAPriCORN)**
The Chicago Community Trust

- **C10_Accelerating Data Value Across a National Community Health Center Network (ADVANCE)**
Oregon Community Health Information Network (OCHIN)

i2b2

- **C11_P2aTH: Towards a Learning Health System in the Mid-Atlantic Region**
University of Pittsburgh

PPRNs represent a number of conditions...

Organization	PI	Condition	Population Size
Accelerated Cure Project for Multiple Sclerosis	Robert McBurney	Multiple Sclerosis	20,000
Cincinnati Children's Hospital Medical Center	Peter Margolis	Pediatric Crohn's Disease and Ulcerative Colitis i2b2	15,000
COPD Foundation	Richard Mularski	Chronic Obstructive Pulmonary Disease	50,000
Crohn's and Colitis Foundation of America	R. Balfour Sartor	Inflammatory Bowel Disease (Crohn's disease and ulcerative colitis)	30,000
Global Healthy Living Foundation	Seth Ginsberg	Arthritis (rheumatoid arthritis, spondyloarthritis), musculoskeletal disorders (osteoporosis), and inflammatory conditions (psoriasis)	50,000
Massachusetts General Hospital	Andrew Nierenberg	Major Depressive Disorder and Bipolar Disorder i2b2	50,000
Univ of California, San Francisco	Mark Pletcher	Cardiovascular health	100,000
University of South Florida	Rebecca Sutphen	Hereditary Breast & Ovarian Cancer	17,000

...including rare diseases

Organization	PI	Condition	Population Size
ALD Connect, Inc	Florian Eichler	Adrenoleukodystrophy	3,000
Arbor Research Collaborative for Health	Bruce Robinson	Primary Nephrotic Syndrome, Focal Segmental Glomerulosclerosis, Minimal Change Disease, and Membranous Nephropathy Multiple Sclerosis	1,250
Duke University	Laura Schanberg	Juvenile Rheumatic Disease i2b2	9,000
Epilepsy Foundation	Janice Beulow	Aicardi Syndrome, Lennox-Gastaut Syndrome, Phelan-McDermid Syndrome, Hypothalamic Hamartoma, Dravet Syndrome, Tuberous Sclerosis	1,500
Genetic Alliance, Inc	Sharon Terry	Alström syndrome , Dyskeratosis congenital, Gaucher disease, Hepatitis, Inflammatory breast cancer, Joubert syndrome, Klinefelter syndrome & associated conditions, Psoriasis, Metachromatic leukodystrophy, Pseudoxanthoma elasticum,	50- 50,000
Immune Deficiency Foundation	Kathleen Sullivan	Primary Immunodeficiency Diseases	1,250
Parent Project Muscular Dystrophy	Holly Peay	Duchenne and Becker muscular dystrophy	4,000
Phelan-McDermid Syndrome Foundation	Megan O'Boyle	Phelan-McDermid Syndrome i2b2	737
University of Pennsylvania	Peter Merkel	Vasculitis	500

Key Challenges

- Governance of Data Types
- Sustainability
- Patient Privacy
- Patient Engagement

The Plan

Common data platform (**i2b2**)

+

Federated queries across sites (**SHRINE**)

+

Point of care apps (**SMART**)

+

Patient-facing technologies (**RedCap, SMART, +**)

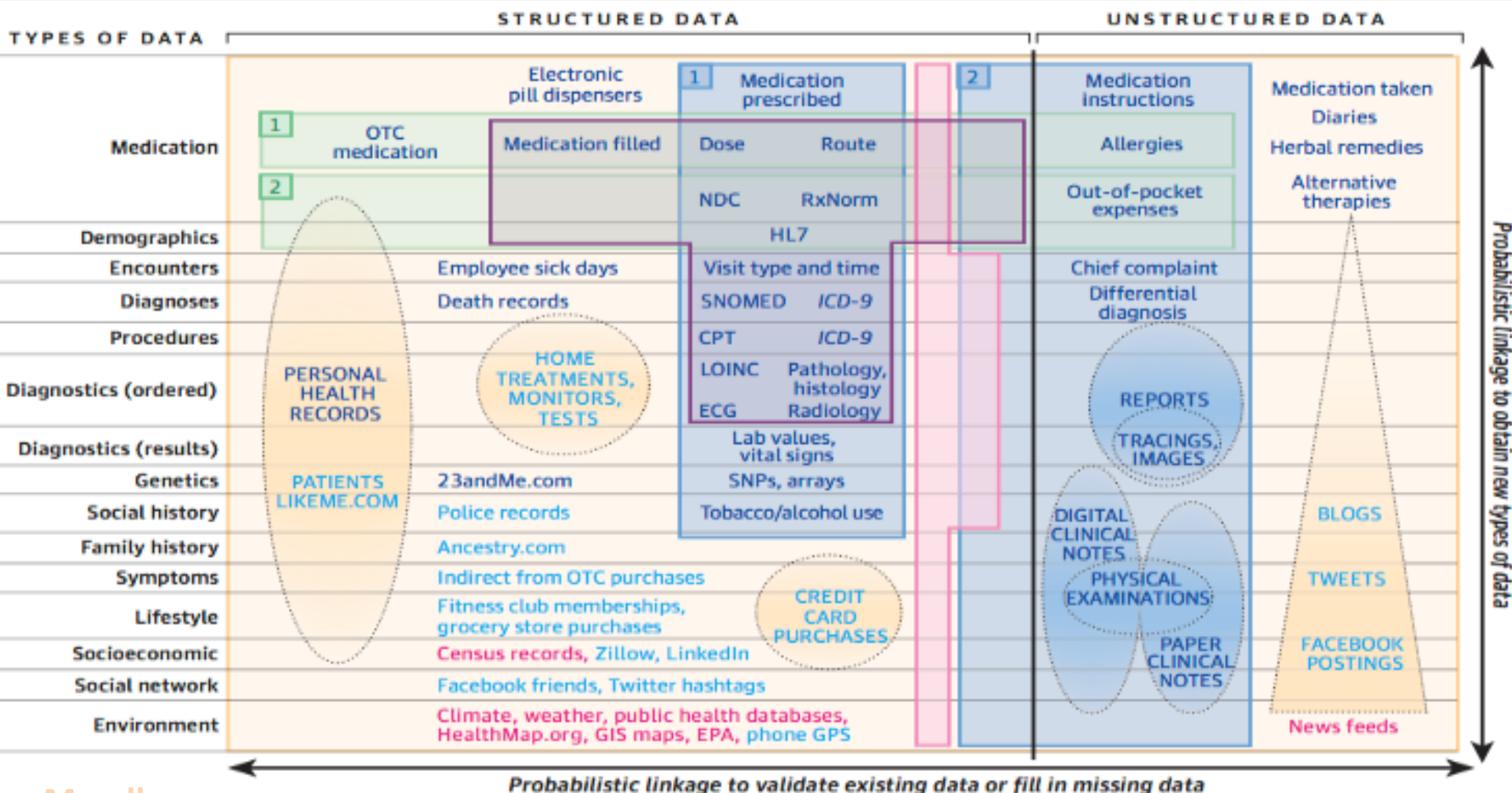
Key Challenges

- Governance of Data Types
- Sustainability
- Patient Privacy
- Patient Engagement

Health Data

- Myriad types
 - Labs, Meds, Diagnoses are dominant in EHRs
 - Payors and some health systems have claims
 - Patient Reported Outcomes are increasingly used in clinical settings
 - Text and notes
 - Genomic, proteomic, medical device device, smart phone social media data critical to some studies

The Tapestry of Potentially High-Value Information Sources That May be Linked to an Individual for Use in Health Care



Probabilistic linkage to validate existing data or fill in missing data

Examples of biomedical data

- 1 2 Pharmacy data
- Claims data
- Data outside of health care system
- 1 2 Health care center (electronic health record) data
- Registry or clinical trial data

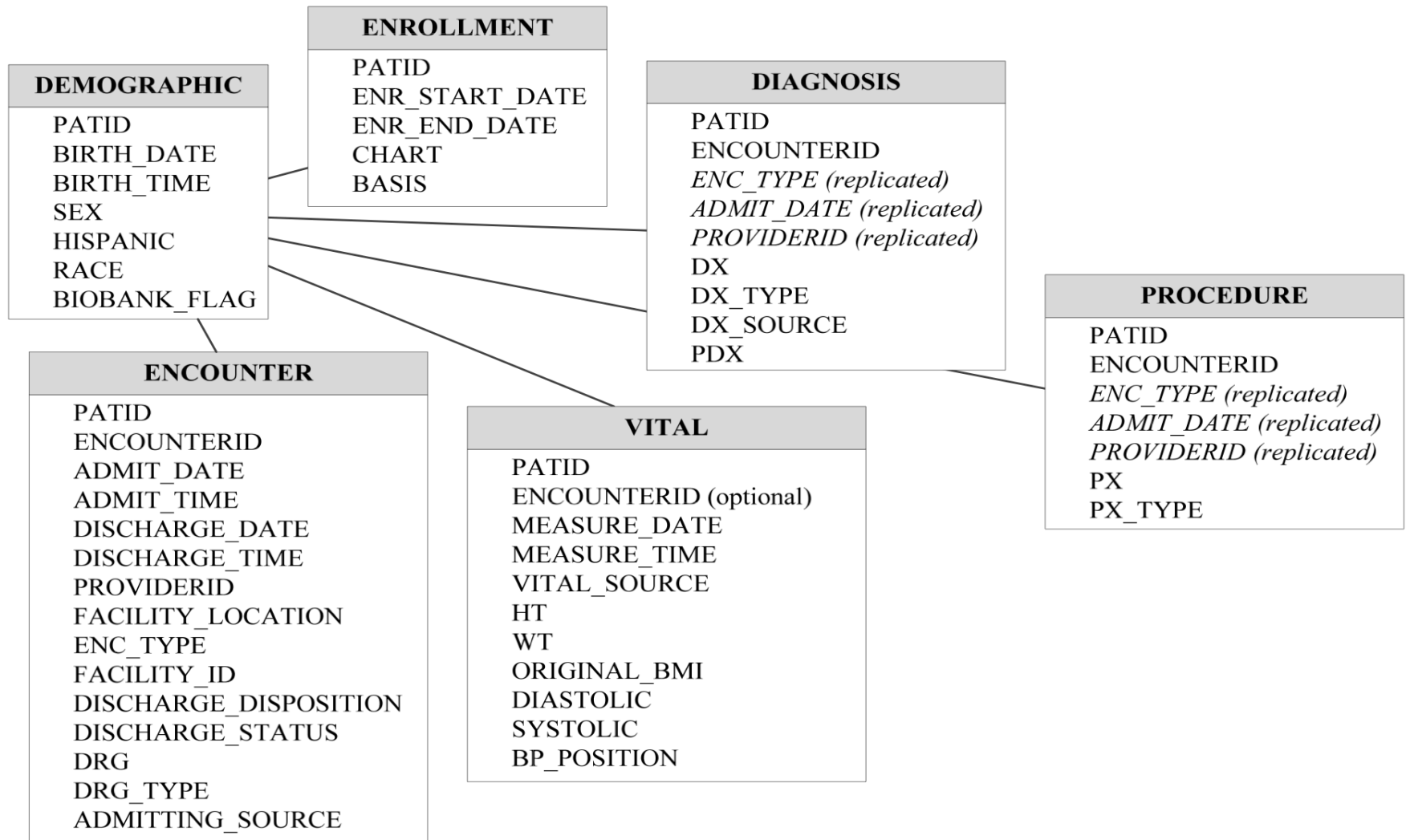
Ability to link data to an individual

- Easier to link to individuals
- Harder to link to individuals
- Only aggregate data exists

Data quantity

- More
- Less

PCORnet CDM v1



PCORnet Common Data Model v3.0

New to v3.0

DEMOGRAPHIC
PATID
BIRTH_DATE
BIRTH_TIME
SEX
HISPANIC
RACE
BIOBANK_FLAG

Fundamental basis

ENROLLMENT
PATID
ENR_START_DATE
ENR_END_DATE
CHART
ENR_BASIS
DISPENSING
DISPENSINGID
PATID
PRESCRIBINGID (optional)
DISPENSE_DATE
NDC
DISPENSE_SUP
DISPENSE_AMT
DEATH
PATID
DEATH_DATE
DEATH_DATE_IMPUTE
DEATH_SOURCE
DEATH_MATCH_CONFIDENCE
DEATH_CONDITION
PATID
DEATH_CAUSE
DEATH_CAUSE_CODE
DEATH_CAUSE_TYPE
DEATH_CAUSE_SOURCE
DEATH_CAUSE_CONFIDENCE

Data captured from processes associated with healthcare delivery

VITAL
VITALID
PATID
ENCOUNTERID (optional)
MEASURE_DATE
MEASURE_TIME
VITAL_SOURCE
HT
WT
DIASTOLIC
SYSTOLIC
ORIGINAL_BMI
BP_POSITION
SMOKING
TOBACCO
TOBACCO_TYPE
CONDITION
CONDITIONID
PATID
ENCOUNTERID (optional)
REPORT_DATE
RESOLVE_DATE
ONSET_DATE
CONDITION_STATUS
CONDITION
CONDITION_TYPE
CONDITION_SOURCE
PRO_CM
PRO_CM_ID
PATID
ENCOUNTERID (optional)
PRO_ITEM
PRO_LOINC
PRO_DATE
PRO_TIME
PRO_RESPONSE
PRO_METHOD
PRO_MODE
PRO_CAT

Data captured within multiple contexts: healthcare delivery, registry activity, or directly from patients

ENCOUNTER
ENCOUNTERID
PATID
ADMIT_DATE
ADMIT_TIME
DISCHARGE_DATE
DISCHARGE_TIME
PROVIDERID
FACILITY_LOCATION
ENC_TYPE
FACILITYID
DISCHARGE_DISPOSITION
DISCHARGE_STATUS
DRG
DRG_TYPE
ADMITTING_SOURCE
DIAGNOSIS
DIAGNOSISID
PATID
ENCOUNTERID
ENC_TYPE (replicated)
ADMIT_DATE (replicated)
PROVIDERID (replicated)
DX
DX_TYPE
DX_SOURCE
PDX
PROCEDURES
PROCEDURESID
PATID
ENCOUNTERID
ENC_TYPE (replicated)
ADMIT_DATE (replicated)
PROVIDERID (replicated)
PX_DATE
PX
PX_TYPE
PX_SOURCE

Data captured from healthcare delivery, direct encounter basis

LAB_RESULT_CM
LAB_RESULT_CM_ID
PATID
ENCOUNTERID (optional)
LAB_NAME
SPECIMEN_SOURCE
LAB_LOINC
PRIORITY
RESULT_LOC
LAB_PX
LAB_PX_TYPE
LAB_ORDER_DATE
SPECIMEN_DATE
SPECIMEN_TIME
RESULT_DATE
RESULT_TIME
RESULT_QUAL
RESULT_NUM
RESULT_MODIFIER
RESULT_UNIT
NORM_RANGE_LOW
NORM_MODIFIER_LOW
NORM_RANGE_HIGH
NORM_MODIFIER_HIGH
ABN_IND
PRESCRIBING
PRESCRIBINGID
PATID
ENCOUNTERID (optional)
RX_PROVIDERID
RX_ORDER_DATE
RX_ORDER_TIME
RX_START_DATE
RX_END_DATE
RX_QUANTITY
RX_REFILLS
RX_DAYS_SUPPLY
RX_FREQUENCY
RX_BASIS
RXNORM_CUI

PCORNET_TRIAL
PATID
TRIALID
PARTICIPANTID
TRIAL_SITEID
TRIAL_ENROLL_DATE
TRIAL_END_DATE
TRIAL_WITHDRAW_DATE
TRIAL_INVITE_CODE

Associations with PCORnet clinical trials

HARVEST
NETWORKID
NETWORK_NAME
DATAMARTID
DATAMART_NAME
DATAMART_PLATFORM
CDM_VERSION
DATAMART_CLAIMS
DATAMART_EHR
BIRTH_DATE_MGMT
ENR_START_DATE_MGMT
ENR_END_DATE_MGMT
ADMIT_DATE_MGMT
DISCHARGE_DATE_MGMT
PX_DATE_MGMT
RX_ORDER_DATE_MGMT
RX_START_DATE_MGMT
RX_END_DATE_MGMT
DISPENSE_DATE_MGMT
LAB_ORDER_DATE_MGMT
SPECIMEN_DATE_MGMT
RESULT_DATE_MGMT
MEASURE_DATE_MGMT
ONSET_DATE_MGMT
REPORT_DATE_MGMT
RESOLVE_DATE_MGMT
PRO_DATE_MGMT
REFRESH_DEMOGRAPHIC_DATE
REFRESH_ENROLLMENT_DATE
REFRESH_ENCOUNTER_DATE
REFRESH_DIAGNOSIS_DATE
REFRESH_PROCEDURES_DATE
REFRESH_VITAL_DATE
REFRESH_DISPENSING_DATE
REFRESH_LAB_RESULT_CM_DATE
REFRESH_CONDITION_DATE
REFRESH_PRO_CM_DATE
REFRESH_PRESCRIBING_DATE
REFRESH_PCORNET_TRIAL_DATE
REFRESH_DEATH_DATE
REFRESH_DEATH_CAUSE_DATE

Process-related data

Bold font indicates fields that cannot be null due to primary key definitions or record-level constraints.

It is a rapidly evolving picture!

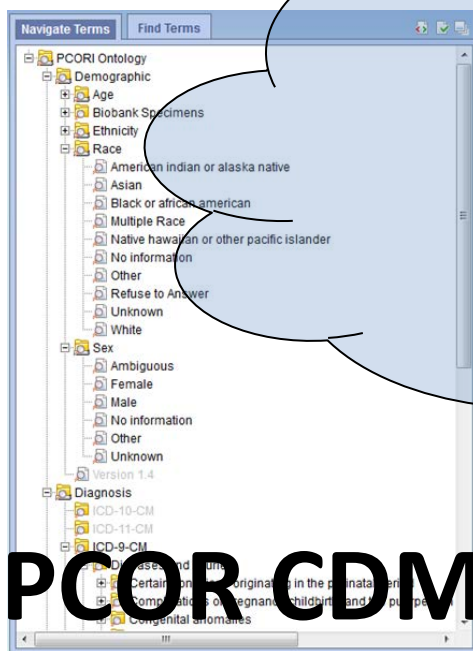
- I2b2 is designed to support health-system-wide data harmonization
- BD2K and beyond

i2b2

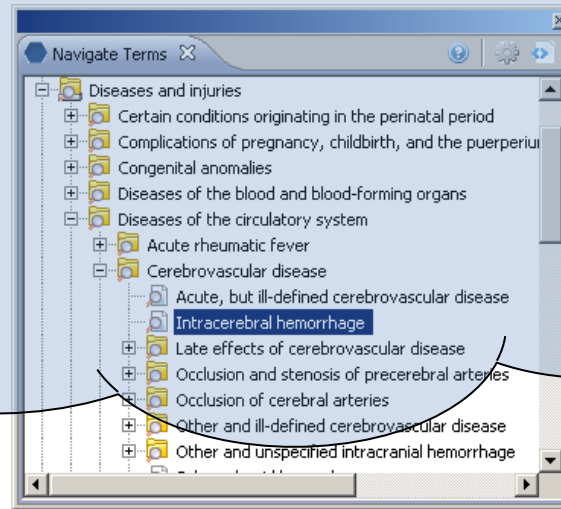
- Explicitly organizes and transforms person-oriented clinical data for clinical and translational research
 - Allows integration of clinical data, trials data, and genotypic data
- A portable and extensible application framework
 - Software is built in a modular pattern that allows additions without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

I2b2 is a Giant Data Sponge

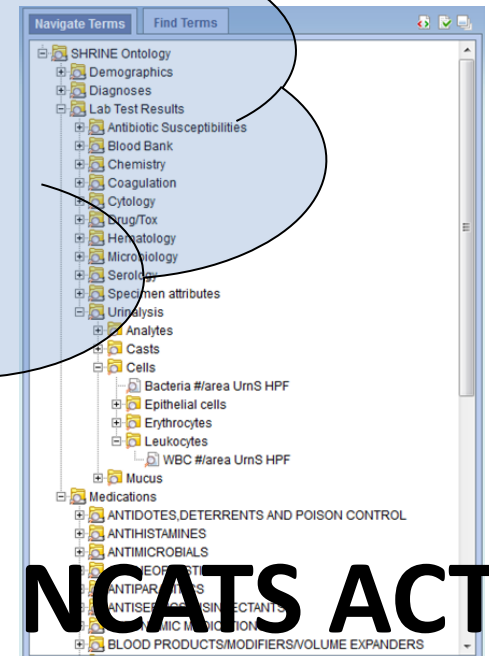
Generally Represents Patient Data



PCOR CDM

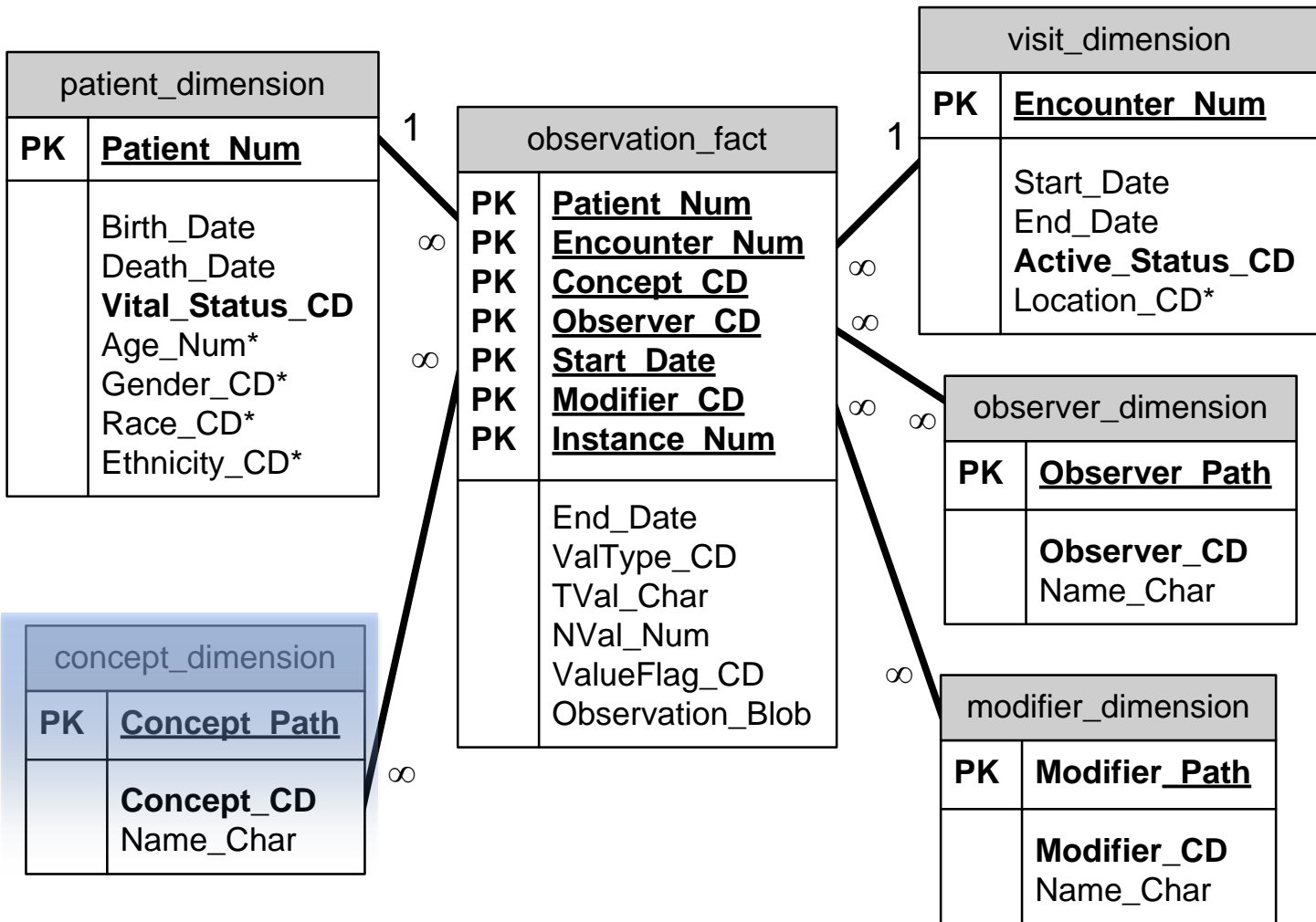


[Other]



NCATS ACT

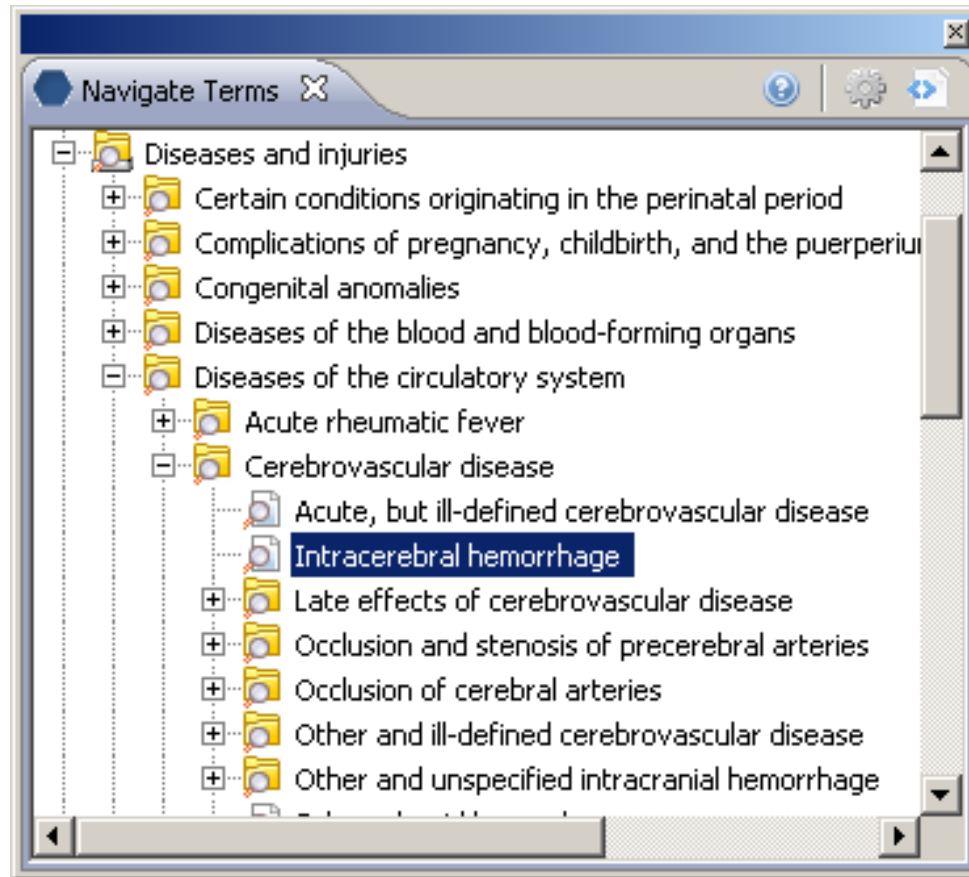
i2b2 Star Schema



Concept_dimension “look” ...

CONCEPT_PATH	CONCEPT_CD	NAME_CHAR
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.4	Toxic myopathy
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.5	Myopathy in endocrine diseases ...
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.6	Symptomatic inflammatory myop...
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.8	Other myopathies
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.81	Critical illness myopathy
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.89	Other myopathies
\i2b2\Diagnoses\Neurologic Disorders (320-389)\Peripheral nerve disorde...	ICD9:359.9	Myopathy, unspecified
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(260) Kwashiorkor\	ICD9:260	Hypoproteinosis
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(261) Nutritional maras...	ICD9:261	Marasmus
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(262) Other severe pr...	ICD9:262	Other severe protein-calorie mal...
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(263) Other and unspe...	ICD9:263	Other and unspecified protein-c...
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(263) Other and unspe...	ICD9:263.0	Malnutrition of moderate degree
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(263) Other and unspe...	ICD9:263.1	Malnutrition of mild degree
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(263) Other and unspe...	ICD9:263.2	Dwarfism, nutritional
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(263) Other and unspe...	ICD9:263.8	Other protein-calorie malnutrition
\i2b2\Diagnoses\Nutritional deficiencies (260-269)\(263) Other and unspe...	ICD9:263.9	Protein-calorie undernutrition

Metadata Presentation



Diseases and injuries \ Diseases of the circulatory system \
Cerebrovascular disease \ Intracerebral hemorrhage

Adapting i2b2 to PCORNet Data Model

1. Ontology-Driven Physical Transformation into PCORNet Common Data Model with a Generalizable approach to adapt to other Common Data Models

Data interchange using i2b2

RECEIVED 31 July 2015
REVISED 26 October 2015
ACCEPTED 31 October 2015

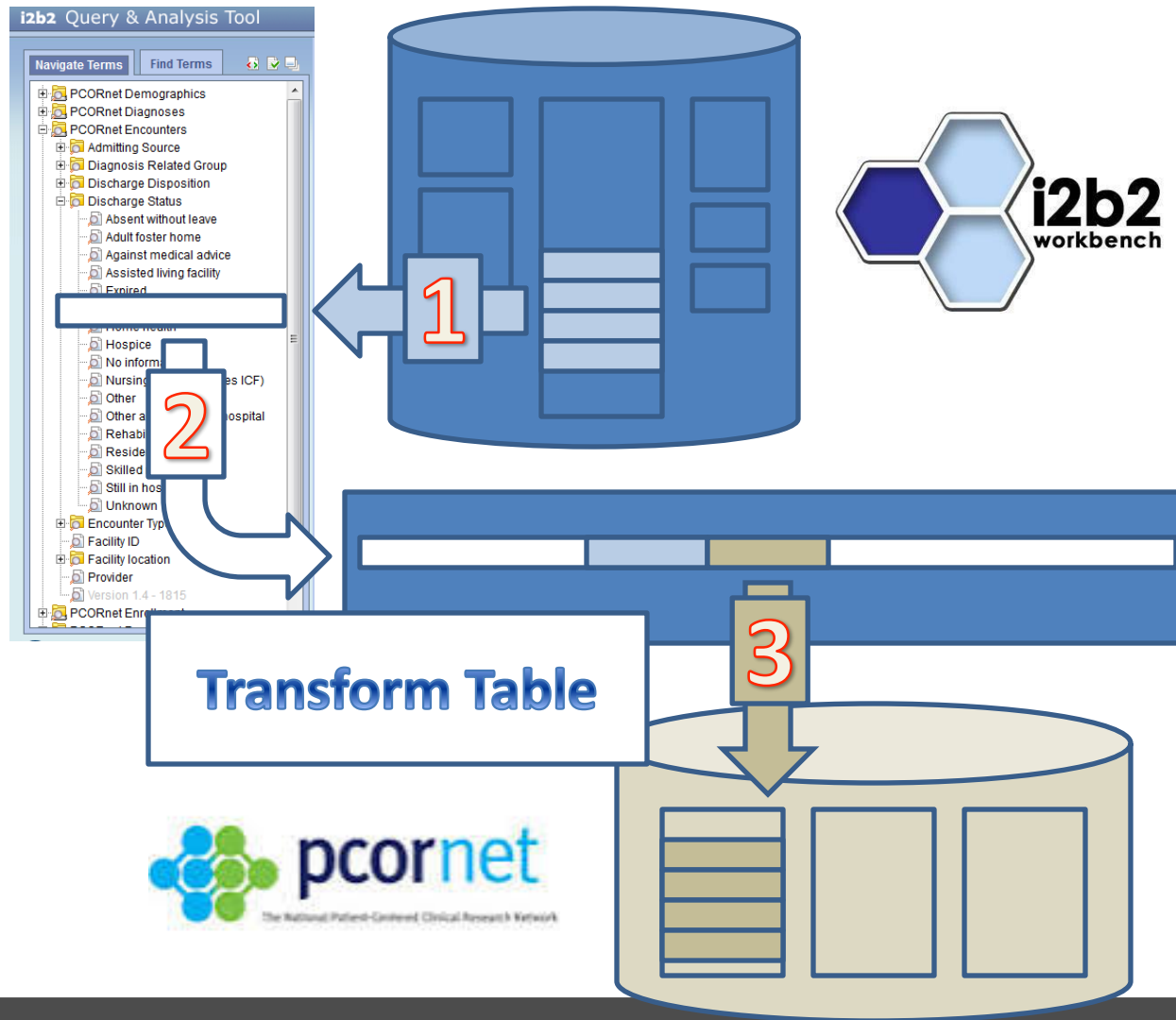
J Am Med Inform Assoc 2016;0:1–8. doi:10.1093/jamia/ocv188, Research and Applications

Jeffrey G Klann^{1,2,3}, Aaron Abend⁴, Vijay A Raghavan², Kenneth D Mandl^{2,5}, and
Shawn N Murphy^{1,2,3}



2. Implement a SHRINE adapter enabling participation in a peer-to-peer network

Physical data transformations to non-i2b2 formats ontology driven



New Information Model Ontology

Consensus Ontology
can live alongside
other ontologies

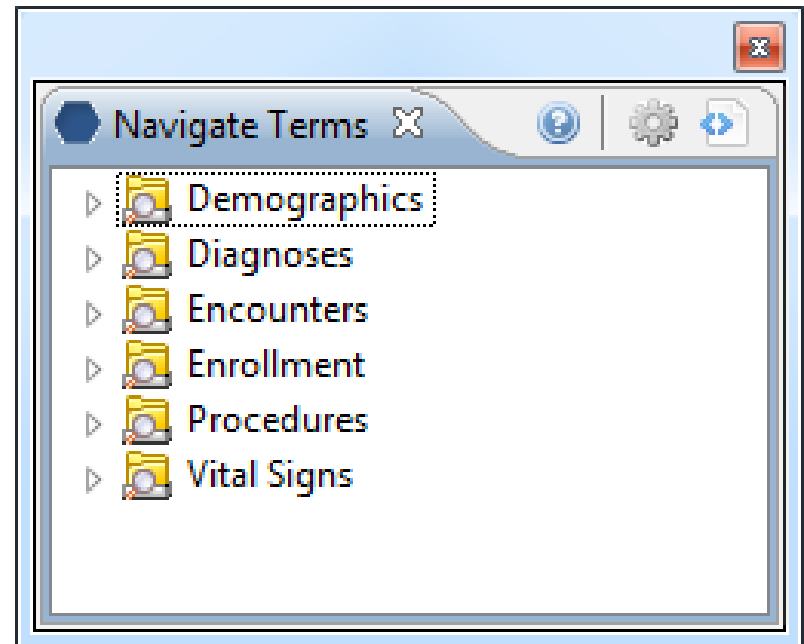
(For example:
PCORNet CDM
ontology and the i2b2
demo ontology in this
case)

- + Clinical Trials
- + Custom Metadata
- + Demographics
- + Diagnoses
- + Diagnoses (ICD10)
- + Expression Profiles Data
- + Laboratory Tests
- + Medications
- + PCORnet Core
- + PCORnet Demographics
- + PCORnet Diagnoses
- + PCORnet Encounters
- + PCORnet Enrollment
- + PCORnet Procedures
- + PCORnet Vital Signs
- + Procedures
- + Providers
- + Reports

PCORI-centric Use Case

A requirement in SCILHS is that all i2b2 instances have to map their ontologies to a central PCORI CDM ontology.

- Local codes need to be integrated into the PCORI hierarchy
- Tools and strategies to assist with this effort



ICD-9 plain_code example

- If your institution uses ICD-9 but in a plain_code format...
 - Change the c_basecode to match your code format.

PCORI_BASECODE	C_BASECODE
ICD9:250.1	2501

```
update pcornt  
set c_basecode = substring(pcori_basecode, 6, 25)  
where pcori_basecode like 'ICD9:%'  
and c_fullname like '\PCORI\DIAGNOSIS\09\%';
```

```
update pcornt  
set c_basecode = replace(c_basecode, '.', '')  
where pcori_basecode like 'ICD9:%'  
and c_fullname like '\PCORI\DIAGNOSIS\09\%';
```

Why does SCILHS not just only use PCORNet Common Data Model for All Queries??

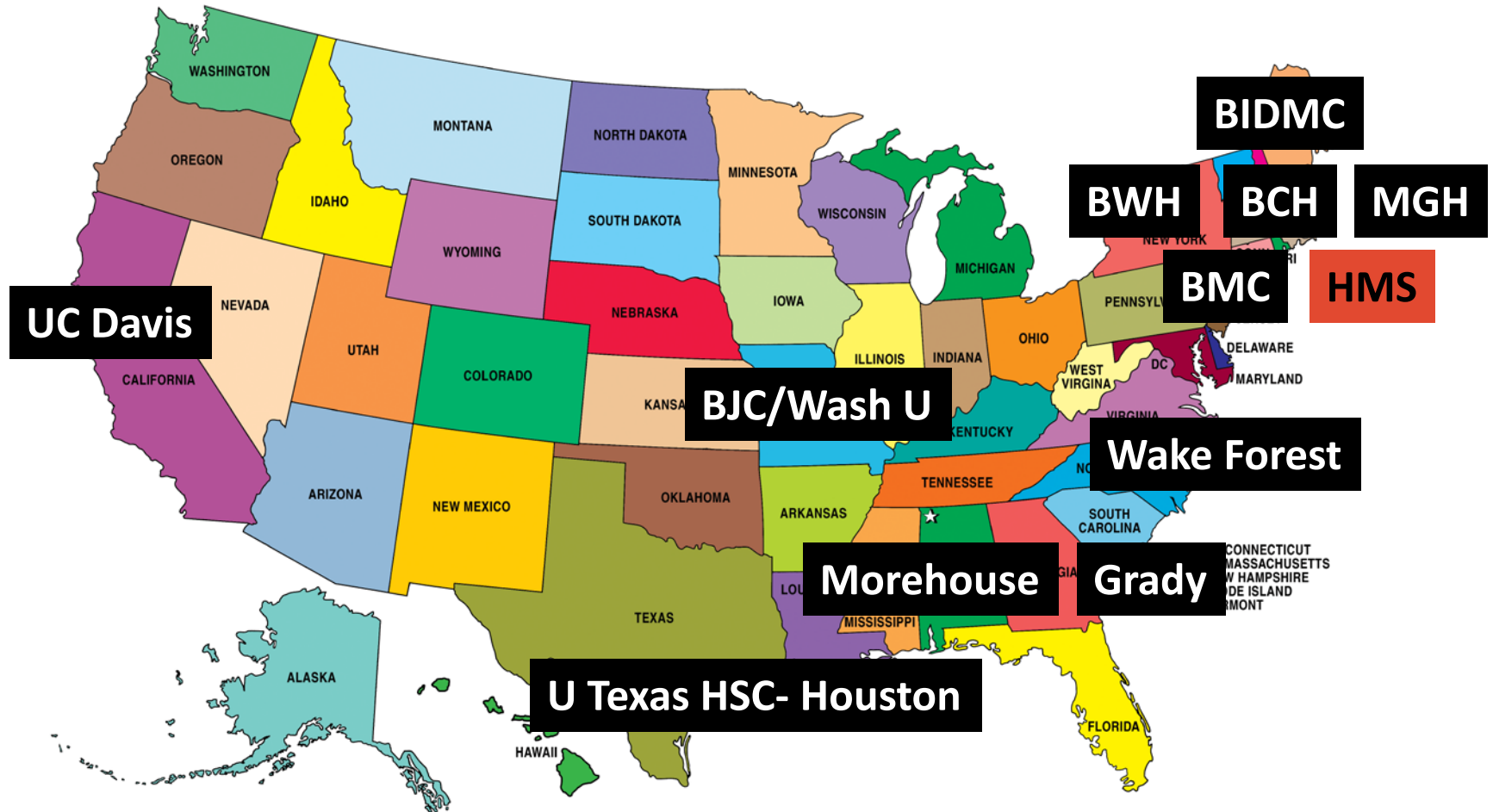
- PCORNet CDM is a measured approach to create a core dataset queryable by a central authority
- But i2b2 is being used as a complete workflow to support clinical trials in a peer-to-peer network

Key Challenges

- Governance of Data Types
- **Sustainability**
- Patient Privacy
- Patient Engagement

SCILHS 11 Participating Sites

Phase II: 10/1/15 – 9/30/18



Agreements/Policy



- Governance Agreement
 - Executed March, 2016
- Master Reliance Agreement
 - Implemented June, 2015
- Contracting & BAAs (MySCILHS IVR)
- PCORNet Data Sharing Agreement v1
 - 6/11 signed
- SCILHS DSA & DUA template (in progress)
- Engagement Policy (in progress)

What is SCILHS?

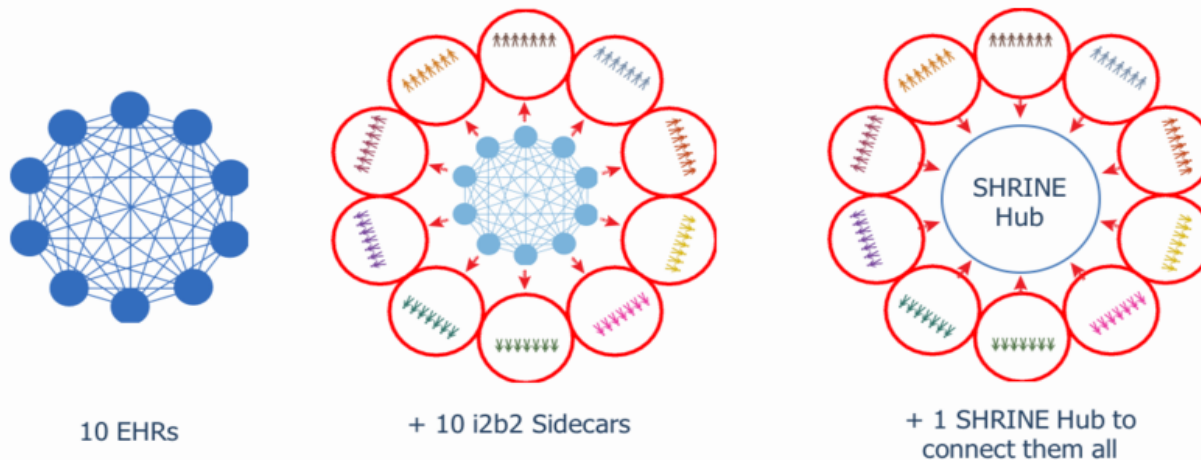
Jun 2, 2016



About the Network

Leveraging an informatics infrastructure that our investigators have developed over the past 15 years, the **Department of Biomedical Informatics at Harvard Medical School** is working across a growing number of **healthcare centers throughout the U.S.** to develop the Scalable Collaborative Infrastructure for a Learning Health System (SCILHS, pronounced "skills"). An **open-source** platform, the system covers more than **10 million patients** and enables clinician and patient participation in research. As of this writing, **a dozen clinical studies are underway** that involve SCILHS.

SCILHS is also a Clinical Data Research Network (CDRN) in the **Patient-Centered Outcomes Research Institute's PCORnet**, a national effort to instantiate a 'network of networks' that supports large-scale comparative effectiveness research. SCILHS uses Informatics for Integrating Biology and the Bedside (i2b2) as its technical backbone and has adopted the PCORnet Common Data Model (CDM) as its foundation for interoperable data exchange. SCILHS uses the Shared Health Research Information Network (SHRINE) platform to enable live distributed querying across sites.



Feasibility / Clinical Study with no contact Information

1

Researcher requests access to the SCILHS Network and define patient cohorts using Shrine Web Client

2

Researcher requests data from SCILHS sites using Shrine Web Client

3

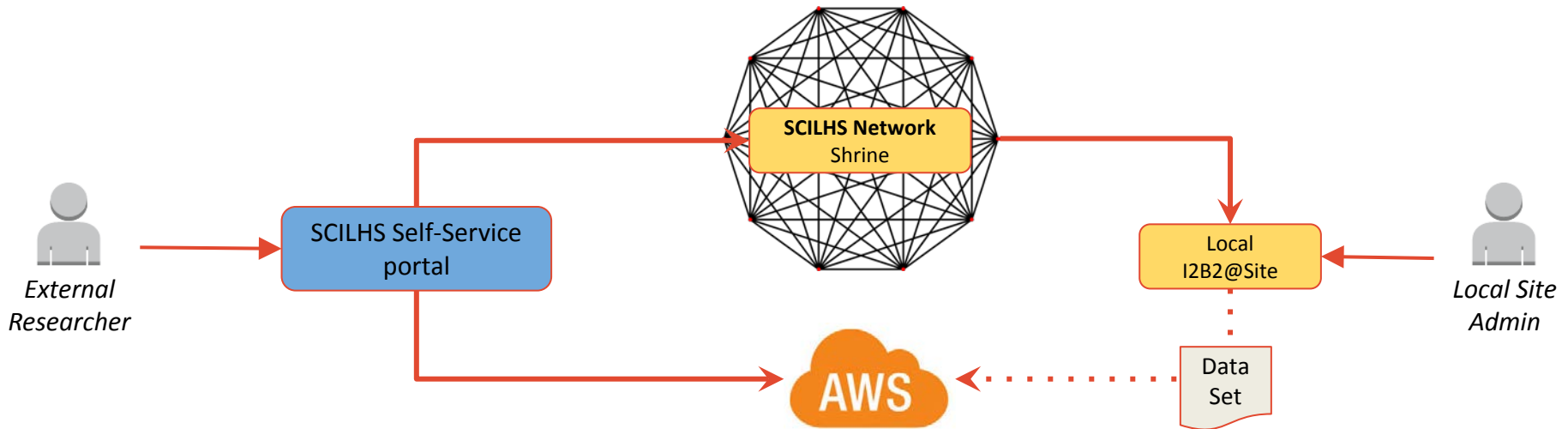
Local Site Admin reviews and approves the query

4

Local Site Admin generates data set file and uploads it into an AWS space

5

Researcher accesses data sets in AWS and performs phenotype computation



SCHILHS SSP Panel
Isaac

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schilhs-ssp/panel/index

SCHILHS
☰

Guest
Request Access

Navigation

Home

The SCILHS Network

Tutorials

Shrine Demo

Analytical Resources Demo

Request Access / Login

What is SCILHS?

Jun 2, 2016

About the Network

Leveraging an informatics infrastructure that our investigators have developed over the past 15 years, the **Department of Biomedical Informatics at Harvard Medical School** is working across a growing number of **healthcare centers throughout the U.S.** to develop the Scalable Collaborative Infrastructure for a Learning Health System (SCILHS, pronounced "skills"). An **open-source** platform, the system covers more than **10 million patients** and enables clinician and patient participation in research. As of this writing, **a dozen clinical studies are underway** that involve SCILHS.

SCILHS is also a Clinical Data Research Network (CDRN) in the **Patient-Centered Outcomes Research Institute's PCORnet**, a national effort to instantiate a

Copyright © 2016 SCILHS All rights reserved.
Demo Version 0.1.5

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schilhs-ssp/panel/index

SCILHS SSP Tutorial

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/tutorial

Isaac

SCILHS Guest

Request Access

Navigation

- Home
- The SCILHS Network
- Tutorials
- Shrine Demo
- Analytical Resources Dem
- Request Access / Login

Shrine tutorial

This page will guide you towards Shrine web client.
Any question? Please, contact info@scilhs.org

SHRINE Basic Tutorial

from Informatics

SHRINE Tutorial for Investigators

04:29 HD :: vimeo

SHRINE Basic TutorialfromInformaticsonVimeo.

Shrine Tutorial

from Informatics

SHRINE Tutorial

04:31 //catalyst.harvard.edu/services/shrine HD :: vimeo

Shrine TutorialfromInformaticsonVimeo.

Copyright © 2016 SCILHS All rights reserved.

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/demo_shrine

Demo Version 0.1.5

Browser window: SCHILS SSP Login
URL: ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/login

Navigation menu:

- SCIILHS Guest
- Request Access
- Home
- The SCILHS Network
- Tutorials
- Shrine Demo
- Analytical Resources Demo
- Request Access / Login

Sign in to the SCILHS self-service portal

researcher

[Password field]

Remember me Log In

[I forgot my password](#)
[New? Request Access](#)

Copyright © 2016 SCILHS All rights reserved. Demo Version 0.1.5

SCILHS SSP Shrine Web

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/new

Isaac

researcher
Online

Navigation

- Home
- The SCILHS Network
- Shrine Tutorials
- Shrine Demo
- Analytical Resources Demo
- Apply for a New Study
- AA32R: First study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources
- WASTS: Diabetes study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources

New Study

Study Name

Short Description

Lead Investigator

Full name (Last, First)

Enter eMail

Institution Name

Department/Division

Title

NIH biosketch for the Lead Investigator

No file chosen

Applying for a new study

Please, use the following form to apply to use the SCILHS network. If you have any questions, please email info@scilhs.org

Fields marked with * are mandatory

Copyright © 2016 SCILHS All rights reserved.

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/.../91271e86-1847-11e6-b6ba-3e1d05dfe78

Demo Version 0.1.5

SCHILHS SSP Shrine Web | ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schilhs-ssp/panel/shrine/91271e86-1847-11e6-b6ba-3e1d05defe78

SHRINE | Project: Schilhs demo | User: Researcher SCILHS user demo | Find Patients | Message Log | Help | Network Help | Logout

researcher Online

Navigation

- Home
- The SCILHS Network
- Shrine Tutorials
- Shrine Demo
- Analytical Resources Demo
- Apply for a New Study
- AA32R: First study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources
- WASTS: Diabetes study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources

Navigate Terms | Find Terms

- >> 00 years old
- >= 85 years old
- Not recorded
- Biobank Specimens
- Ethnicity
- Race
- Sex
- Version 1.5.2 - 354
- PCORnet Diagnoses
 - ICD-11-CM
 - ICD-9-CM
 - DX Type
 - Data Source
 - Primary Diagnosis Flag
 - Diseases and injuries
 - DX Type
 - Data Source
 - Primary Diagnosis Flag
 - Certain conditions originating in the perinatal period
 - Congenital anomalies
 - Diseases of the circulatory system
 - Diseases of the digestive system
 - Diseases of the genitourinary system
 - Diseases of the nervous system and sense organs
 - Diseases of the skin and subcutaneous tissue
 - Endocrine, nutritional and metabolic diseases, and immunity disorders
 - Infectious and parasitic diseases

Query Tool

Query Name: circulatorySystem 18-54

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		
18-34 years old			Diseases of the circulatory system					
35-44 years old								
45-54 years old								

one or more of these **AND** one or more of these **AND** drop a term on here

Run Query | Clear | Print Query | 2 Groups | New Group

Query Status | Graph Results | Query Report | Download Results

Finished Query: "circulatorySystem 18-54" [2.9 secs]
 Compute Time: 2.9 secs
Partners "circulatorySystem 18-54"
 Patient Count: - 43 patients
BCH "circulatorySystem 18-54"
 Patient Count: - 43 patients

Previous Queries

- circulatorySystem 18-54 [6-17-2016] [researcher]
- circulatorySystem 18-54 [6-17-2016] [researcher]
- circulatorySystem 18-54 [6-17-2016] [researcher]
- 18-34-Disease@16:22:11 [6-16-2016] [researcher]
- Occlusion of ce@16:12:53 [6-16-2016] [researcher]
- Congenital anom@16:12:31 [6-16-2016] [researcher]
- 35-44-Chronic@16:12:13 [6-16-2016] [researcher]
- circulatorySystem 18-54 [6-16-2016] [researcher]

Copyright © 2016 SCILHS All rights reserved. Demo Version 0.1.5

SCILHS SSP LDS

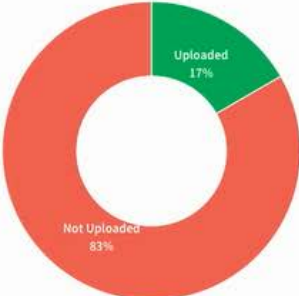
ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/lds/91271e86-1847-11e6-b6ba-3e1d05defe78

researcher Online

Navigation

- Home
 - The SCILHS Network
 - Shrine Tutorials
 - Shrine Demo
 - Analytical Resources Demo
 - Apply for a New Study
- AA32R: First study
 - Scilhs Shrine Client
 - Data sets Compilation
 - Analytical Resources
- WASTS: Diabetes study
 - Scilhs Shrine Client
 - Data sets Compilation
 - Analytical Resources

Data Set Completion Status



Category	Percentage
Uploaded	17%
Not Uploaded	83%

Available Sites

Site Name	File Submitted?	#
Boston Childrens Hospital	Yes	✓
Massachusetts General Hospital	No	✓
Beth Israel Deaconess Medical Center	No	✓
Boston HealthNet	No	✓
University of California Davis	No	✓
Grady Memorial Hospital	No	✓
Brigham and Women's Hospital	No	✓
Morehouse School of Medicine	No	✓
University of Texas Health Science Center at Houston	No	✓
Wake Forest Integrated Health System	No	✓
BJC Healthcare - Washington University St. Louis	No	✓
Partners Healthcare	Yes	✓

Copyright © 2016 SCILHS All rights reserved.

Demo Version 0.1.5

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/compute/91271e86-1847-11e6-b6ba-3e1d05defe78

researcher Online

Navigation

- Home
- The SCILHS Network
- Shrine Tutorials
- Shrine Demo
- Analytical Resources Demo
- Apply for a New Study
- AA32R: First study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources
- WAST5: Diabetes study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources

Environment History

Global Environment

Data

- d 44 obs. of 66 variables
- data 99 obs. of 9 variables
- dataF 42 obs. of 9 variables

X	Patient.ID	Sex	Age.in.Years	Birth.Year	Race	Diseases.of.the.circulatory.system	X	icd	
1	1	1e+09	M	25	1985	hispanic	ICD9:448.9 (2)	NA	ICD9:448.9
2	2	1e+09	F	24	1985	black	ICD9:458.9 (1)	NA	ICD9:458.9
3	3	1e+09	F	24	1985	black	ICD9:458.9 (2)	NA	ICD9:458.9
4	4	1e+09	F	24	1985	black	ICD9:427.9 (2)	NA	ICD9:427.9
5	5	1e+09	M	28	1981	asian	ICD9:401.9 (2)	NA	ICD9:401.9
6	6	1e+09	M	28	1981	asian	ICD9:401.9 (1)	NA	ICD9:401.9
7	7	1e+09	M	29	1981	asian	ICD9:427.89 (2)	NA	ICD9:427.89
8	8	1e+09	M	29	1981	asian	ICD9:427.9 (1)	NA	ICD9:427.9
9	9	1e+09	M	29	1981	asian	ICD9:427.0 (0)	NA	ICD9:427.0
10	10	1e+09	M	29	1981	asian	ICD9:427.0 (1)	NA	ICD9:427.0
11	11	1e+09	M	41	1969	asian	ICD9:427.11 (0)	NA	ICD9:427.11

Showing 1 to 11 of 99 entries

Console

```

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> data <- read.csv('Partners-12b2.csv');
>
> View(data)
> library(ggplot2);
> data <- read.csv('Partners-12b2.csv');
> data["icd"] <- NA;
> data$icd <- substr(data$Diseases.of.the.circulatory.system, 1, 11);
> dataF <- subset(data, subset = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 "));
> #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
> ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
> ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))
> d <- read.csv('surveys/Partners-12b2-survey.csv');
> View(d)
> |

```

Files Plots Packages Help Viewer

- New Folder
- Upload
- Delete
- Rename
- More

Home

Name	Size	Modified
surveys		
script.r	492 B	Jun 17, 2016, 9:08 AM
R		
Partners-12b2.csv	6.1 KB	Jun 15, 2016, 2:17 PM
BCH-12b2.csv	6.1 KB	Jun 16, 2016, 1:38 PM
.Rhistory	9.4 KB	Jun 17, 2016, 9:11 AM

Copyright © 2016 SCILHS All rights reserved. Demo Version 0.1.5

SCHILHS SSP Computation

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schilhs-ssp/panel/compute/91271e86-1847-11e6-b6ba-3e1d05defe78

researcher Online

Navigation

- Home
- The SCILHS Network
- Shrine Tutorials
- Shrine Demo
- Analytical Resources Demo
- Apply for a New Study
- AA32R: First study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources
- WASTS: Diabetes study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources

File Edit Code View Plots Session Build Debug Tools Help

script.r Partners-i2b2.csv data Partners-i2b2-survey.csv d

```

1 library(ggplot2);
2 data <- read.csv('Partners-i2b2.csv');
3 data["lcd"] <- NA;
4 data$icd <- substr(data$Diseases.of.the.circulatory.system, 1, 11);
5 dataF <- subset(data, subset = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 "));
6 #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
7 ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
8 ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))

```

Environment History

Global Environment

Object	Size	Variables
Data	44 obs. of 66 variables	
d	99 obs. of 9 variables	
dataF	42 obs. of 9 variables	

Files Plots Packages Help Viewer

New Folder Upload Delete Rename More

Name	Size	Modified
surveys		
script.r	492 B	Jun 17, 2016, 9:08 AM
R		
Partners-i2b2.csv	6.1 KB	Jun 15, 2016, 2:17 PM
BCH-i2b2.csv	6.1 KB	Jun 16, 2016, 1:38 PM
.Rhistory	9.4 KB	Jun 17, 2016, 9:11 AM

Console

```

Type 'q()' to quit R.
> data <- read.csv('Partners-i2b2.csv');
> View(data)
> library(ggplot2);
> data <- read.csv('Partners-i2b2.csv');
> data["lcd"] <- NA;
> data$icd <- substr(data$Diseases.of.the.circulatory.system, 1, 11);
> dataF <- subset(data, subset = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 "));
> #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
> ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
> ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))
> d <- read.csv('surveys/Partners-i2b2-survey.csv');
> View(d)
> data <- read.csv('Partners-i2b2.csv');

```

Copyright © 2016 SCILHS All rights reserved. Demo Version 0.1.5

SCHILHS SSP Computation x

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schilhs-ssp/panel/compute/91271e86-1847-11e6-b6ba-3e1d05defe78

researcher Online

Navigation

- Home
- The SCILHS Network
- Shrine Tutorials
- Shrine Demo
- Analytical Resources Demo
- Apply for a New Study
- AA32R: First study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources
- WASTS: Diabetes study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources

File Edit Code View Plots Session Build Debug Tools Help

script.r Partners-i2b2.csv data Partners-i2b2-survey.csv d

```

1
2 load('Partners-i2b2.csv');
3
4 seases.of.the.circulatory.system, 1, 11);
5 et = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 ");
6 n.Years, fill=icd)) + geom_histogram(binwidth=5);
7 i.Years, fill=icd)) + geom_density(alpha=.3);
8 i.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))

```

Environment History

Global Environment

Variable	Observations	Variables
d	44 obs.	of 66 variables
data	99 obs.	of 9 variables
dataF	42 obs.	of 9 variables

Files Plots Packages Help Viewer

Age.in.Years

ICD9:401.9 ICD9:424.0 ICD9:427.9 icd

Console

```

> #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
> ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
> ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))
> d <- read.csv('surveys/Partners-i2b2-survey.csv');
> View(d)
> data <- read.csv('Partners-i2b2.csv');
> View(data)
> library(ggplot2);
> data <- read.csv('Partners-i2b2.csv');
> data["icd"] <- NA;
> data$icd <- substr(data$Diseases.of.the.circulatory.system, 1, 11);
> dataF <- subset(data, subset = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 "));
> #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
> ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
> ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))
>

```

8:110 (Top Level) R Script

Copyright © 2016 SCILHS All rights reserved. Demo Version 0.1.5

SCILHS SSP Computation x

ec2-52-41-13-166.us-west-2.compute.amazonaws.com/schils-ssp/panel/compute/91271e86-1847-11e6-b6ba-3e1d05defe78

researcher Online

Navigation

- Home
- The SCILHS Network
- Shrine Tutorials
- Shrine Demo
- Analytical Resources Demo
- Apply for a New Study
- AA32R: First study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources
- WASTS: Diabetes study
- Scilhs Shrine Client
- Data sets Compilation
- Analytical Resources

File Edit Code View Plots Session Build Debug Tools Help

```

1 {2b2.csv}';
2
3
4 seases.of.the.circulatory.system, 1, 11);
5 let = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 ");
6 n.Years, fill=icd)) + geom_histogram(binwidth=5);
7 i.Years, fill=icd)) + geom_density(alpha=.3);
8 i.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))

```

Environment History

Global Environment

Data	Observations	Variables
d	44 obs.	of 66 variables
data	99 obs.	of 9 variables
dataF	42 obs.	of 9 variables

Files Plots Packages Help Viewer

```

> #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
> ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
> ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))
> d <- read.csv('surveys/Partners-i2b2-survey.csv');
> View(d)
> data <- read.csv('Partners-i2b2.csv');
> View(data)
> library(ggplot2);
> data <- read.csv('Partners-i2b2.csv');
> data["icd"] <- NA;
> data$icd <- substr(data$Diseases.of.the.circulatory.system, 1, 11);
> dataF <- subset(data, subset = data$icd %in% c("ICD9:401.9 ", "ICD9:424.0 ", "ICD9:427.9 "));
> #ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_histogram(binwidth=5);
> ggplot(dataF, aes(x=Age.in.Years, fill=icd)) + geom_density(alpha=.3);
> ggplot(dataF, aes(y=Age.in.Years, x=icd))+geom_point(shape=1, position=position_jitter(width=.05,height=.05))
>

```

8:110 (Top Level) R Script

Console

Copyright © 2016 SCILHS All rights reserved. Demo Version 0.1.5

Active SCILHS Studies

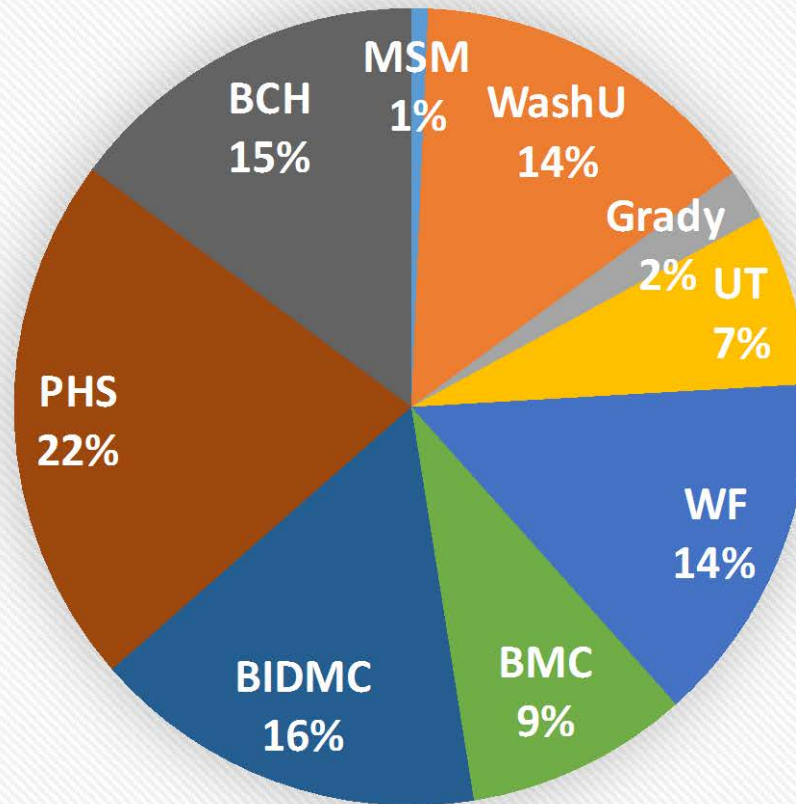
- Bariatric Surgery Obesity Demonstration Project (Tavakkoli)
- Pediatric Antibiotic Obesity Demonstration Project (Taveras)
- Suicide Prediction – Fuss Proposal (Smoller and Reis)
- Parkinson’s Clinical Trial – Data Recruitment (Schwarzschild)
- INVESTED Clinical Trial (Solomon)
- RELIANCE Clinical Trial (Pending SSRC Review)

SCILHS Studies Pending Funding Decision

- PCORnet Health System Demonstration Project, Effectiveness of Case Management Programs in ACOs (Metlay) - **FUNDED**
- Methods to collect longer-term PROs for Trauma Registries (Haider)
- Population-Scale Predictive Modeling for Suicide Risk (Reis)

Characteristics of Network

Patients in SCILHS



**15,856,539
patients!**

■ MSM ■ WashU ■ Grady ■ UT ■ WF ■ BMC ■ BIDMC ■ PHS ■ BCH

Key Challenges

- Governance of Data Types
- Sustainability
- **Patient Privacy**
- Patient Engagement

SHRINE and i2b2

- Rapid, real time query
- Increasingly complex queries possible
- PHI completely protected

Researchers formulate queries across hospitals

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

Navigate Terms Find Terms

- [-] Diseases of the circulatory system
 - [+] Acute rheumatic fever
 - [+] Cerebrovascular disease
 - [+] Chronic rheumatic heart disease
 - [+] Diseases of arteries, arterioles, a
 - [-] Diseases of pulmonary circulation
 - [+] Acute pulmonary heart disease
 - [-] Chronic pulmonary heart disease
 - [-] Chronic pulmonary embolism
 - [-] Chronic pulmonary heart disease
 - [-] Kyphoscoliotic heart disease
 - [-] Other chronic pulmonary heart disease
 - [-] Primary pulmonary hypertension
 - [+] Other diseases of pulmonary circulation
 - [+] Diseases of veins and lymphatics
 - [+] Unresponsive disease

Query Tool

Query Name: Prima-18-34-Black@13:16:55

Group 1			Group 2			Group 3				
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude		
Treat Independently			Treat Independently			Treat Independently				
Primary pulmonary hypertension			18-34 years old			Black or african american				
one or more of these			AND	one or more of these			AND	one or more of these		

Run Query Clear Print Query 3 Groups New Group

Previous Queries

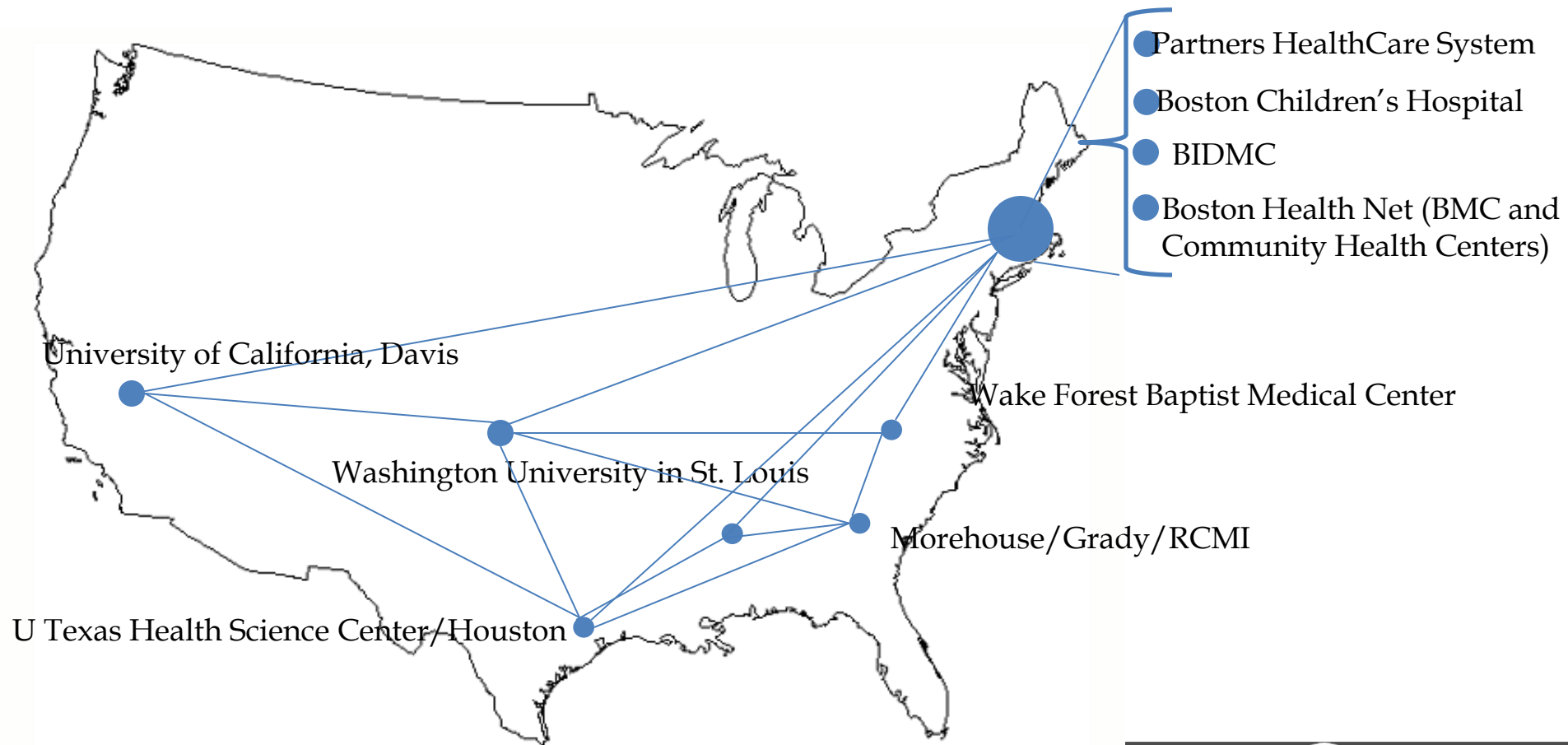
- [+] Height-Weight@09:30:13 [1-9-2015] [smurphy]
- [+] >= 85-Heigh-Weigh@09:21:06 [1-9-2015] [smurp
- [+] Weight@09:57:04 [1-9-2015] [smurphy]
- [+] >= 85-Heigh-Weigh@09:53:27 [1-9-2015] [smurp
- [+] Height-Weight@09:40:58 [1-9-2015] [smurphy]
- [+] 18-34-Heigh-Weigh@13:44:36 [1-6-2015] [smurp
- [+] Asian-Appendi@13:37:29 [1-6-2015] [smurphy]

Query Status

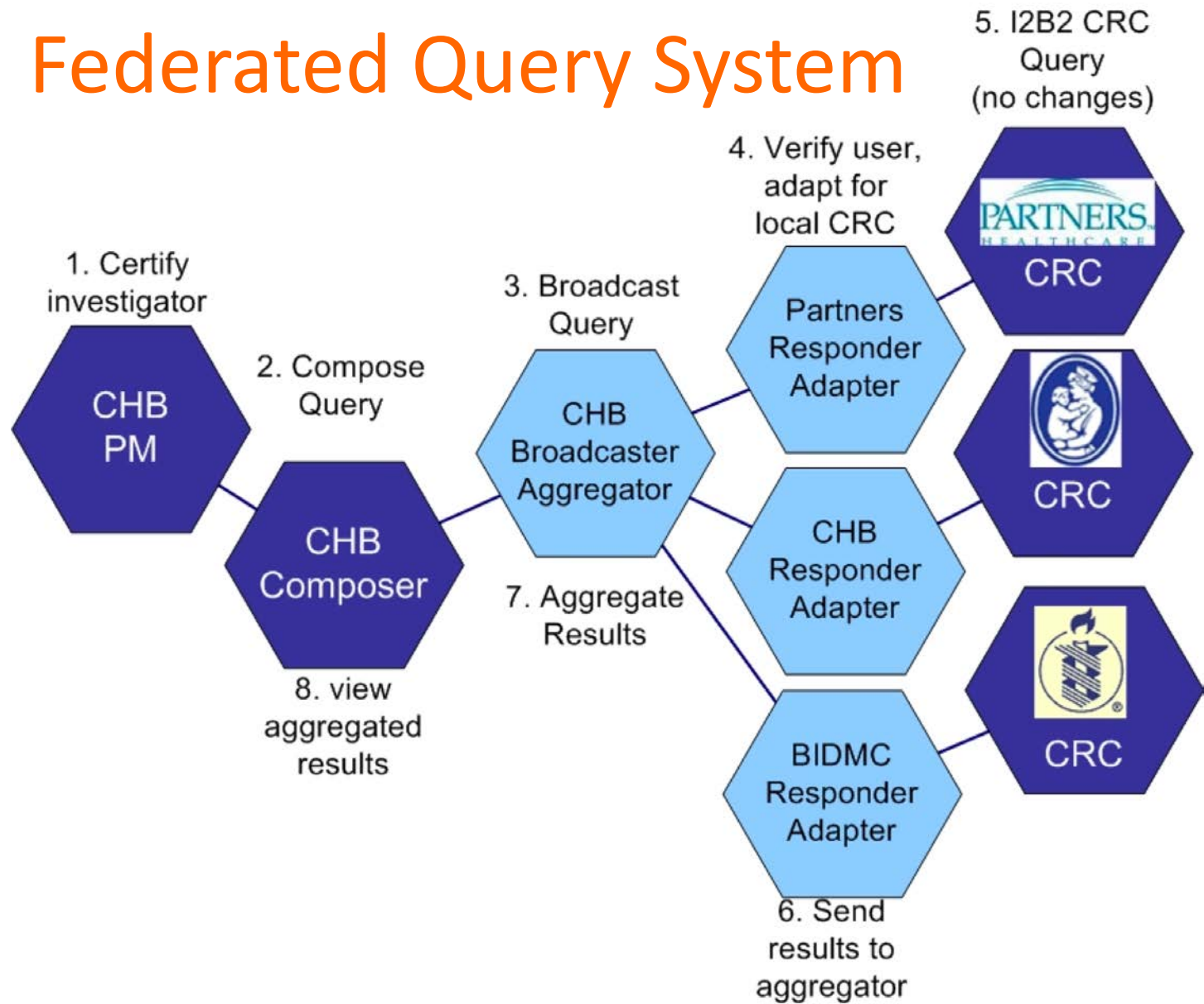
Finished Query: "Prima-18-34-Black@13:16:55" [19.4 secs]

CCHMC - 10 patients or fewer	FINISHED [2.0 secs]
UT - 96 ±3 patients	FINISHED [4.7 secs]
Partners HealthCare - 30 ±3 patients	FINISHED [3.5 secs]
Wake - 15 ±3 patients	FINISHED [10.0 secs]
BMC - 14 ±3 patients	FINISHED [13.2 secs]
BCH - 125 ±3 patients	FINISHED [18.4 secs]

SCILHS Clinical Data Research Network



Federated Query System



- **Measuring Quality of Decisions About Treatment of Depression**
ClinicalTrials.gov Identifier: NCT01152307

- Eligibility

Ages Eligible for Study: 18 Years and older
Genders Eligible for Study: Both
Accepts Healthy Volunteers: No

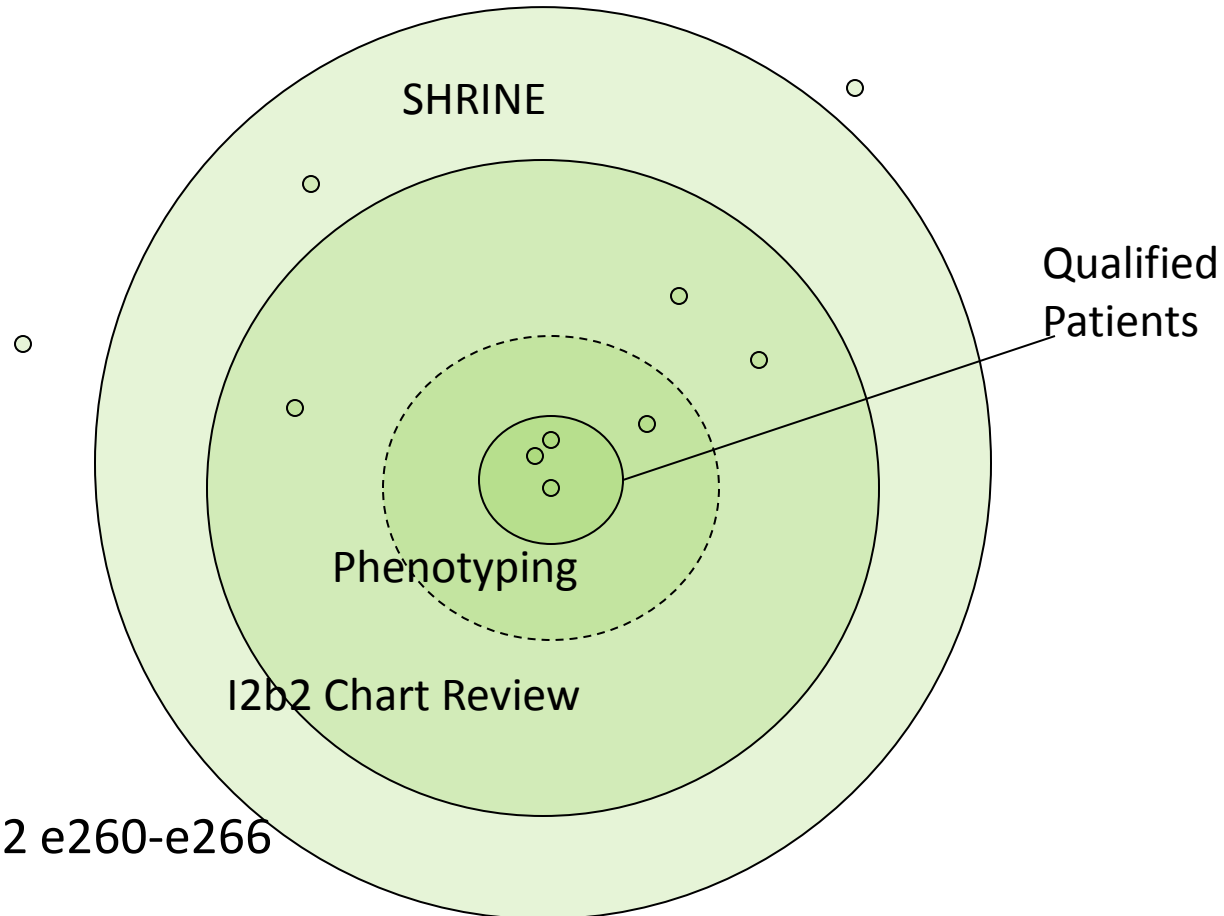
- Inclusion Criteria:

- Aged 18+
- In the last 12 months, **talked to a health care provider about starting or stopping a treatment** (prescription medicine for depression or counseling)

- Exclusion Criteria:

- none

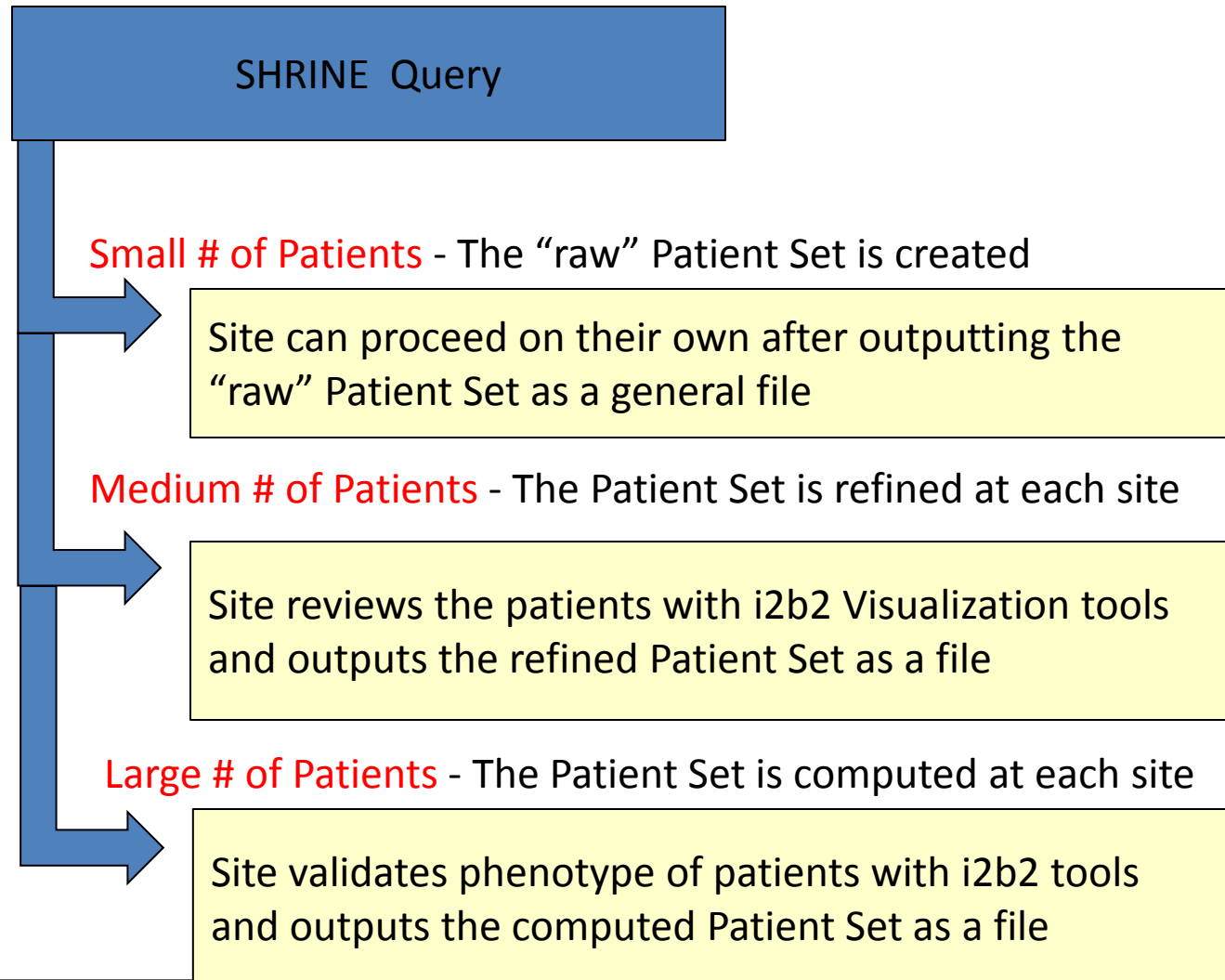
Strategy to Find Qualified Patients for Clinical Trials



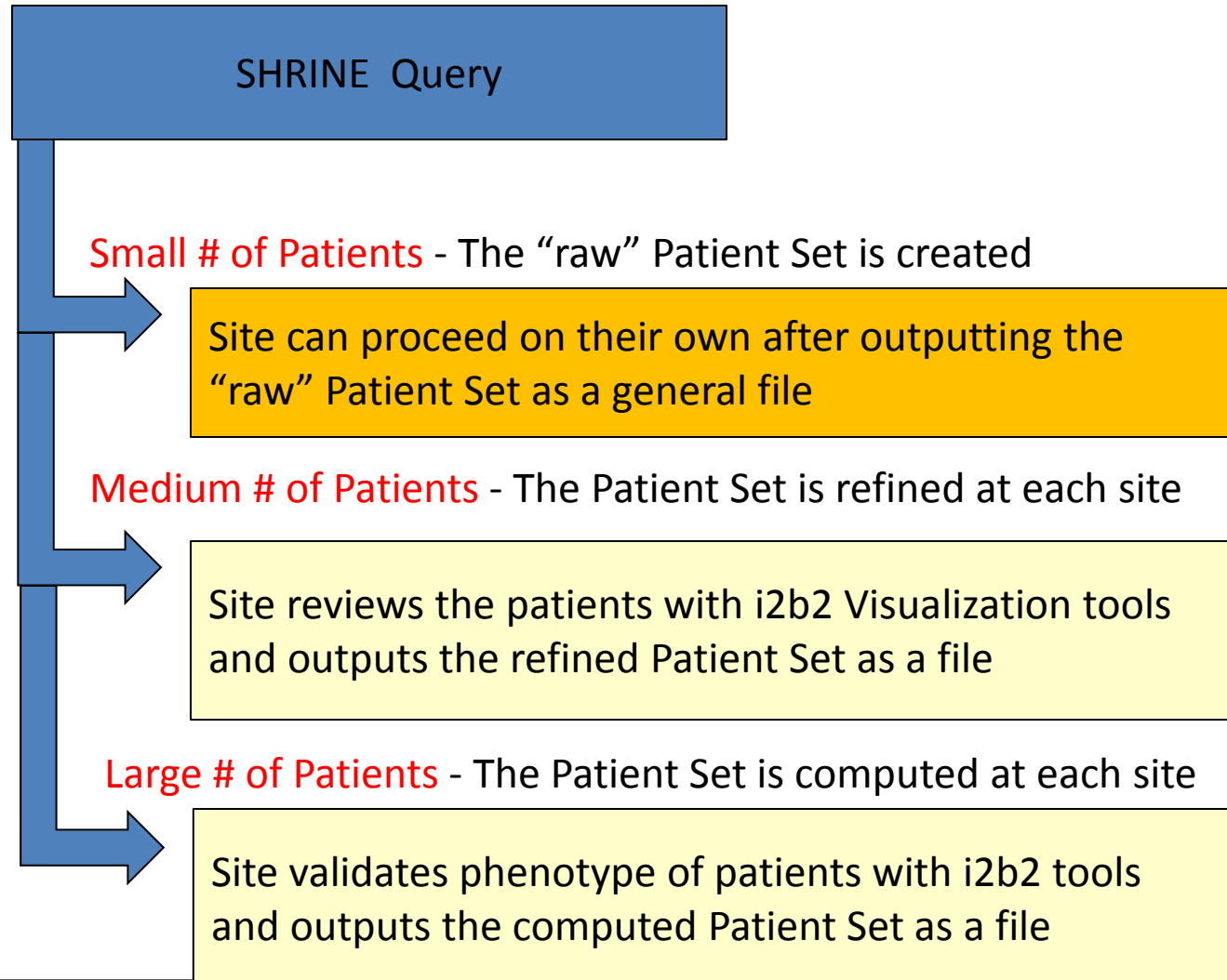
Jack London
JAMIA 2013;20:e2 e260-e266

Design-phase prediction of potential cancer clinical trial accrual success using a research data mart

Handling Clinical Trials in an i2b2-SHRINE Network depends on Number of Patients found



If a Small Number of Patients are Identified in SHRINE Query



Clinician Sensitivities in the Recruitment Step – why the detailed data stays at the sites

- Control of Protected Health Information
 - PHI remains at the hospitals when collected under a waiver of consent
- Control of recruitment policies
- Patients as a scarce and valuable resource
 - Revenue and scientific recognition are tied to recruitment for clinical trials

Run Query Using SHRINE

The screenshot displays the SHRINE web application interface. At the top, the header includes the application name "SHRINE", the project name "Project: SHRINE", the user name "User: Shawn Murphy", and navigation links for "Find Patients", "Message Log", "Help", and "Logout".

The main interface is divided into several sections:

- Navigate Terms / Find Terms:** A tree view on the left showing medical categories such as "Diseases of the circulatory system", "Diseases of pulmonary circulation", and "Diseases of veins and lymphatics".
- Query Tool:** A central area for building queries. It shows a "Query Name" field with the value "Prima-18-34-Black@13:16:55". Below this, three query groups are defined:
 - Group 1:** Contains the term "Primary pulmonary hypertension".
 - Group 2:** Contains the term "18-34 years old".
 - Group 3:** Contains the term "Black or african american".The groups are connected by "AND" operators. Each group has a green box labeled "one or more of these".
- Run Query / Clear / Print Query:** Buttons for executing the query, clearing the tool, and printing the query.
- Query Status:** A section at the bottom right showing the results of the query. It indicates the query is "Finished" and provides a breakdown of results by site:
 - Finished Query: "Prima-18-34-Black@13:16:55"** [19.4 secs]
 - CCHMC - 10 patients or fewer [FINISHED [2.0 secs]]
 - UT - 96 ±3 patients [FINISHED [4.7 secs]]
 - Partners HealthCare - 30 ±3 patients [FINISHED [3.5 secs]]
 - Wake - 15 ±3 patients [FINISHED [10.0 secs]]
 - BMC - 14 ±3 patients [FINISHED [13.2 secs]]
 - BCH - 125 ±3 patients [FINISHED [18.4 secs]]
- Previous Queries:** A list on the bottom left showing recent queries, such as "Height-Weight@09:30:13 [1-9-2015] [smurphy]" and "18-34-Heigh-Weigh@13:44:36 [1-6-2015] [smurphy]".

Local Site Admin starts up i2b2 workbench

SHRINE Workbench for Shrinserver2001

Shrine Admin Status: ● i2b2

SHRINE Previous Queries

Search By Name

Containing

Find eCommons ID: All Users

- Heartbeat Query [04-18-2014] [heartbeat][25871]
- Heartbeat Query [04-18-2014] [heartbeat][25870]
- Heartbeat Query [04-18-2014] [heartbeat][25869]
- SHRINE Asthma Query @18:15:32 [04-18-2014] [nww22] ✓
- Heartbeat Query [04-18-2014] [heartbeat][25867]
- Male-Hyperte@18:12:37 [04-18-2014] [kz31][25866]
- Heartbeat Query [04-18-2014] [heartbeat][25865]
- Heartbeat Query [04-18-2014] [heartbeat][25864]
- Heartbeat Query [04-18-2014] [heartbeat][25863]
- Heartbeat Query [04-18-2014] [heartbeat][25862]
- Heartbeat Query [04-18-2014] [heartbeat][25861]
- Heartbeat Query [04-18-2014] [heartbeat][25860]
- Male-Hyperte@17:41:30 [04-18-2014] [kz31][25859] ✓
- Heartbeat Query [04-18-2014] [heartbeat][25858]
- Heartbeat Query [04-18-2014] [heartbeat][25857]
- Heartbeat Query [04-18-2014] [heartbeat][25856]
- breast cancer test long run query [04-18-2014] [prk2][25854] ✓
- Heartbeat Query [04-18-2014] [heartbeat][25855]
- Heartbeat Query [04-18-2014] [heartbeat][25853]
- Heartbeat Query [04-18-2014] [heartbeat][25852]

Query Name:

SHRINE workbench BETA

Analysis Types

- PATIENT SET
- PATIENT COUNT XML
- Gender patient breakdown
- Vital Status patient breakdown
- Race patient breakdown
- Age patient breakdown
- TimeLine

Query Timing

- Treat all groups independent
- Selected groups occur in the
- Items instance will be same

Patient(s) returned:

Begin: 04-18-2014 13:20:01

© Copyright i2b2 contributors and others 2005-2013. All rights reserved. Build 00023B

Local Site admin reviews query

The screenshot displays the SHRINE Workbench interface for Shrinserver2001. The main window is titled "SHRINE Workbench for Shrinserver2001" and includes a menu bar (File, Window, Help) and a status bar (Shrine Admin Status: ● i2b2).

The interface is divided into several panels:

- SHRINE Previous Queries:** A list of queries with search filters. A blue callout box points to this list with the text "SHRINE Queries are shown in this panel".
- Query Tool:** A panel for configuring a query. It features a "Query Name" field, three groups (Group 1, Group 2, Group 3), and an "Add Group" button. Each group has options for "Dates", "Occurs > 0x", and "Exclude".
- Analysis Types:** A sidebar on the right showing various analysis options like "PATIENT SET", "PATIENT COUNT XML", "Gender patient breakdown", etc.
- Query Timing:** A section with radio buttons for "Treat all groups independent", "Selected groups occur in the", and "Items instance will be same".
- Buttons:** "Get Everyone" and "Run Query Above" buttons are located at the bottom of the Query Tool panel.

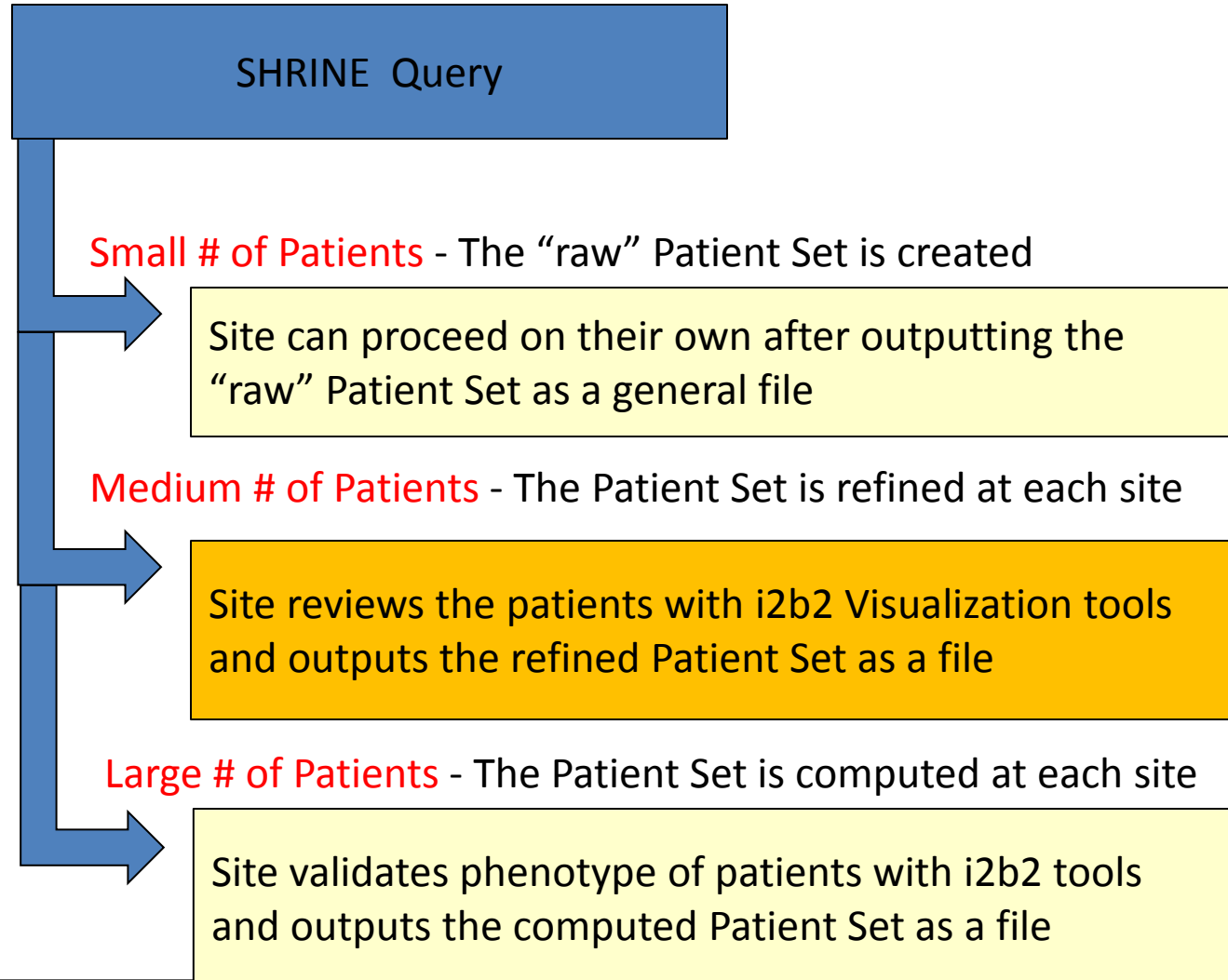
The bottom status bar shows "Begin: 04-18-2014 13:20:01".

Can output “raw” Patient Set

The screenshot displays the i2b2 Workbench interface. The main window shows a query named "Circulatory sys@03:09:37" with three groups. The "Patient Set" analysis type is selected, and the "Number of patients" is highlighted. A blue callout box points to the patient set results, stating: "The Patient Set can be dragged to this panel which shows patient MRNs and allows set to be output." The results panel shows a list of patient MRNs and demographic information for 25 patients.

HIVE	BWH	MGH	PW1	FH	DFC				
1000000001	2000001961	300000182							
1000000003	2000001963	300000182							
1000000007	2000001967	300000182							
1000000009	2000001969	300000182							
1000000011	11489929	3000001831	400002011	S500003061	U500004035	0915	01164887		
1000000012	11489986	3000001832	400002012	S500003062	U500004035	790926	01164897		
1000000013	2000001973	3000001833	400002013	U500004035					
1000000014	11490505	3000001834	400002014	S500003064	U500004035	100791247	01165476		
1000000016	2000001976	3000001834	400002016	S500003066	U500004035				
1000000018	00000117	3000001838	400002018	S500003068	U500004035	101164949	01612917		
1000000020	2000001980	3000001840	400002020	S500003070	U500004030				
1000000023	2000001983	3000001843	400002023	U500004033					
1000000024	2000001984	3000001844	400002024	S500003074					
1000000025	15783376	3000001845	400002025	S500003075	U500004035	101809330	01800290		
1000000026	17028580	3000001845	400002026	S500003076	U500004036	102344360	01954309	00001003	252304
1000000028	17028598	3000001848	400002028	S500003078	U500004038	102344362	01954310		252305

Handling Clinical Trials in an i2b2-SHRINE Network



Or can go on to Review Patients in i2b2

The screenshot displays the i2b2 Workbench interface. The main window shows a patient set named "Asthma-Albuter@01:06:16". The "Render Tables" tab is active, displaying a table of patient data. The table has columns for SMART, Patient ID, PSet #, Patient Name, Gender, Race, Date of Birth, Age, Obesity, Acute Myocardial..., and HMG-CoA reductase inhibitors. The first row is highlighted in blue.

SMART	Patient ID	PSet #	Patient Name	Gender	Race	Date of Birth	Age	Obesity	Acute Myocardial...	HMG-CoA reduct
+	100000001	1+1	xxxxx, xxxxx	F	black	1985-11-17T...	21	✓	✓	✓
*	100000002	1+1	xxxxx, xxxxx	F	white	1966-08-29T...	40	✗	✓	✗
*	100000003	1	xxxxx, xxxxx	M	asian	1969-03-07T...	38	✓	✗	✗
*	100000004	1	xxxxx, xxxxx	M	black	1976-08-13T...	30	✓	✓	✓
*	100000005	1	xxxxx, xxxxx	F	hispanic	1973-06-25T...	33	✓	✓	✓
*	100000006	1	xxxxx, xxxxx	F	black	1981-08-05T...	25	✓	✗	✓
*	100000007	1	xxxxx, xxxxx	M	asian	1981-07-10T...	25	✓	✓	✗
*	100000008	1	xxxxx, xxxxx	M	hispanic	1986-12-16T...	20	✓	✓	✓
*	100000009	1	xxxxx, xxxxx	F	hispanic	1989-01-11T...	18	✗	✓	✓
*	100000010	1	xxxxx, xxxxx	F	hispanic	1976-11-23T...	30	✗	✓	✗
*	100000011	1	xxxxx, xxxxx	F	white	2049-10-19T...	57	✓	✗	✓
*	100000012	1	xxxxx, xxxxx	F	black	1991-03-29T...	16	✓	✗	✓
*	100000013	1	xxxxx, xxxxx	F	black	2027-02-27T...	80	✗	✗	✓
*	100000014	1	xxxxx, xxxxx	M	white	1956-05-22T...	50	✗	✗	✓
*	100000015	1	xxxxx, xxxxx	M	hispanic	1978-09-11T...	28	✓	✗	✓
*	100000016	1	xxxxx, xxxxx	F	indian	1983-11-19T...	23	✓	✓	✓

The interface also shows a search panel on the left with "reductase inhibitors" selected, and a patient set list on the bottom left. The top right shows the user "i2b2 User" and status "i2b2".

Review Patients with SMART tools

The screenshot displays the SMART i2b2 Demo Site interface for Patient 100000001. The patient's demographic information is as follows:

- Num: 100000001
- Sex: Male
- Birth Date: 1986-12-16
- Race: Hispanic
- Age: 26 years (current)

The interface features several key components:

- Navigation Pane (Left):** Includes 'Navigate Term', 'Search by Names', and 'Previous Query' sections.
- Main Content Area:** Contains four primary views:
 - Medications:** A table listing drugs such as Quar 80 mcg/inh aerosol, Atrovent 18 mcg/inh aerosol, and ibuprofen 200mg tablet.
 - BP Centiles:** A graph showing blood pressure trends over time with markers for 'OK' and 'A'.
 - Cardiac Risk:** A risk assessment tool showing LDL and HDL cholesterol levels and a 'Your Risk' of 1% for cardiovascular disease.
 - Problems List:** A table of medical conditions including 'Abdominal pain, other specified site, multiple sites' and 'Acute pharyngitis (disorder)'.
- Right-Hand Sidebar:** A 'Workplace' section with a file explorer showing folders like 'Clinical Trial 'Asthma'', 'Maybe', 'NO', 'YES', and 'demo'.

A “Criteria Matcher” App to review the “list”

View Patient with SMART Apps

Patient Centric View Patient Number: 1000000001

Patient 1000000001
 Num: 1000000001 Sex: Female
 Birth Date: 1985-11-17 Race: Black
 Age: 26 years (current)

Click an icon to add a new SMART app

Meds List CT Selector Problems List Cardiac Risk Labs List

Select View

- Medications
- Problems
- Lab Results
- Vital Signs
- Notes
- Clinical Trials

Clinical Trials

Save Layout Rename Delete View

CT Selector

Your active clinical trial is currently set to:

Pilot Study of Pioglitazone for the Treatment of Moderate to Severe Asthma in Obese Asthmatics (v2)

History Load Previous Clinical Trial Specify A New Clinical Trial

CT Matcher

	100%	Searched	Matched
Demographics	✓	Age between 18 and 60 Both	Age 26; Both
Medications	✗	actos; pioglitazone	
Notes	✗	asthma; wheezing asthmatics	wheezing [1],
Problems	✗	asthma; wheezing; obesity	Asthma, unspecified without mention of status asthmaticus; Asthma, unspecified type, with acute exacerbation; Extrinsic asthma without status asthmaticus; Cough

Drop a new patient record here to load it.

scalable collaborative infrastructure
 learning health system

Classify Patients

i2b2 Workbench for i2b2 Demo (Oracle)

i2b2 User Status: ● i2b2

Timeline View Analysis View Patient Mapping View CT Viewer

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 reduct
✖	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	✓	✓	✓

Workplace

- SHARED
 - Clinical Trial 'Asthma'
 - Maybe
 - NO
 - YES
 - 1000000001 [21 y...
- demo

Output the “refined” Patient Set

The screenshot displays the i2b2 Workbench for i2b2 Demo (Oracle) interface. The main window is titled "Excel Plugin" and contains the "Spreadsheet Exporter Plugin" header. Below the header is a table with two columns: "Set #" and "Patient Set Name". The table is currently empty.

On the left side, there are two search panels. The top panel, "Find Terms", has "Search by Names" and "Search by Codes" tabs. Under "Search by Names", the search criteria are "Containing" and "reductase inhibitors". The search results list "HMG-CoA reductase inhibitors" with sub-items: Atorvastatin, Cerivastatin, Fluvastatin, Lovastatin, Lovastatin-niacin, Pravastatin, Rosuvastatin, and Simvastatin. The bottom panel, "Patient Sets", has "Search By Name" and "Containing" tabs. The search criteria are "Containing" and an empty field. The search results list "Patient Set for 'Asthma-Albuter @01:06:16'" with sub-items: 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], and 1000000010 [33 y/o F hispanic].

On the right side, there is a "Workplace" panel showing a tree view of folders and files. The folders are "Clinical Trial 'Asthma'", "Maybe", "NO", and "YES". The files are listed with their IDs and ages: 1000000010 [33 y/o], 1000000039 [27 y/o], 1000000007 [29 y/o], 1000000017 [49 y/o], 1000000018 [71 y/o], 1000000028 [41 y/o], 1000000001 [21 y/o], 1000000004 [34 y/o], 1000000008 [23 y/o], and 1000000005 [37 y/o].

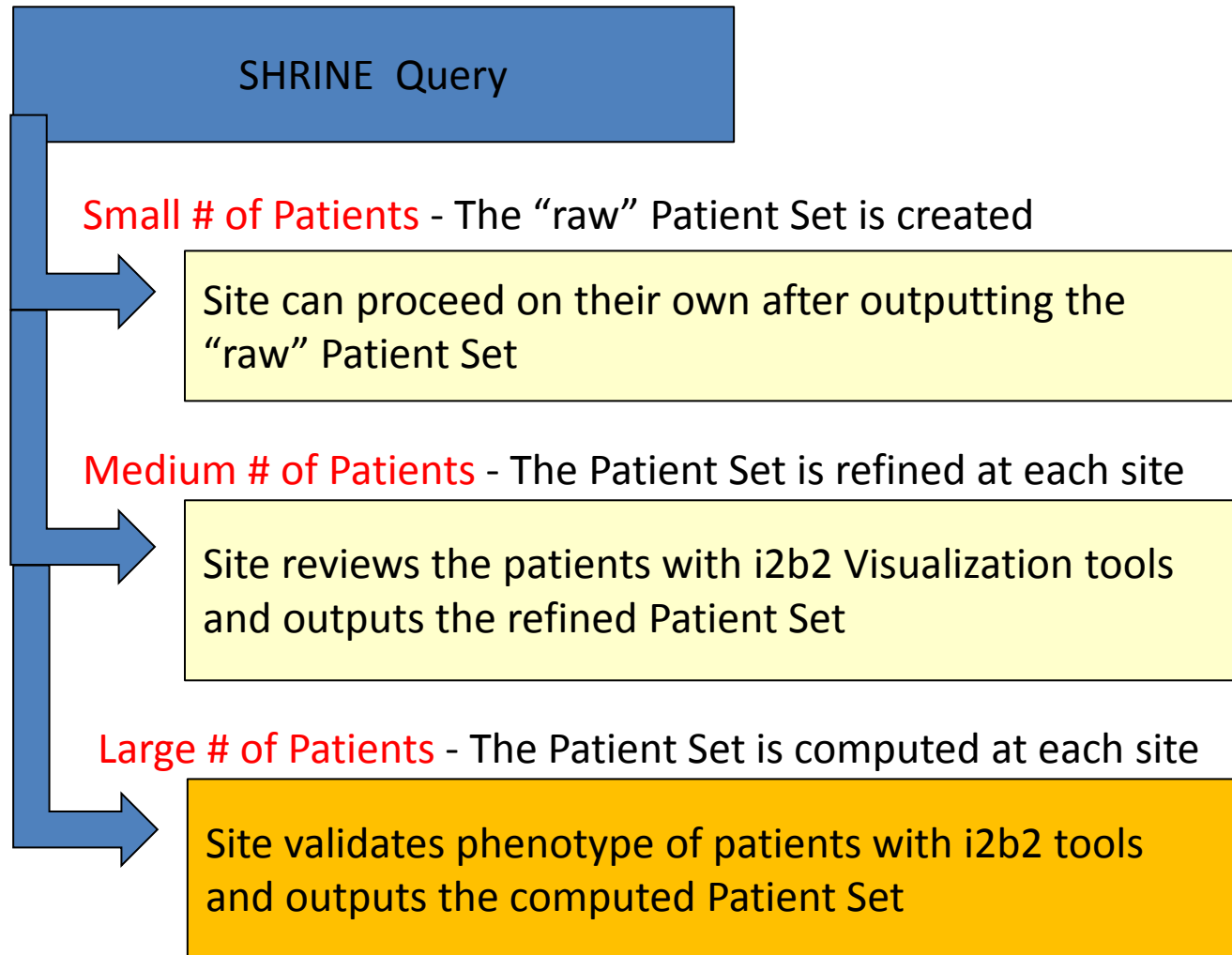
Often Output to REDCap

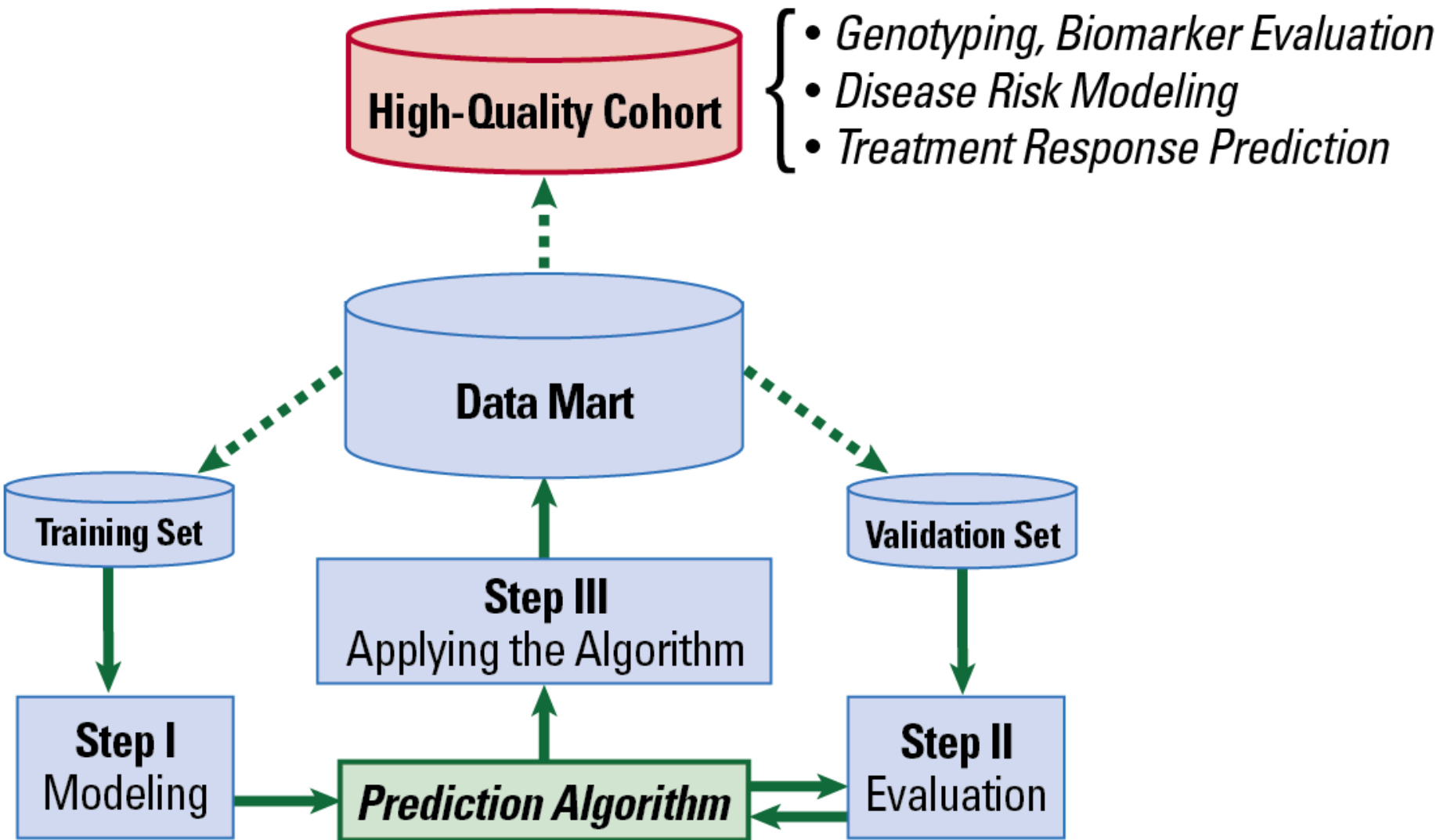
The screenshot displays the i2b2 Workbench interface. The main window is titled "i2b2 Workbench for i2b2 Demo (Oracle)" and shows the "REDCap Plugin" window. The interface is divided into several sections:

- Search by Names:** A search filter set to "Containing" with the text "reductase inhibitors". Below it, a list of HMG-CoA reductase inhibitors is shown, including Atorvastatin, Cerivastatin, Fluvastatin, Lovastatin, Lovastatin-niacin, Pravastatin, Rosuvastatin, and Simvastatin.
- Search By Name:** A search filter set to "Containing" with an empty text box.
- Patient Sets:** A list of patient sets for "Asthma-Albuter @01:06:16", including:
 - 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]
- REDCap Exporter Plugin:** A table with the following structure:

Set #	Patient Set Name
- Workplace:** A tree view showing a "Clinical Trial 'Asthma'" folder with sub-folders "Maybe", "NO", and "YES", each containing patient sets with their IDs and demographic information.

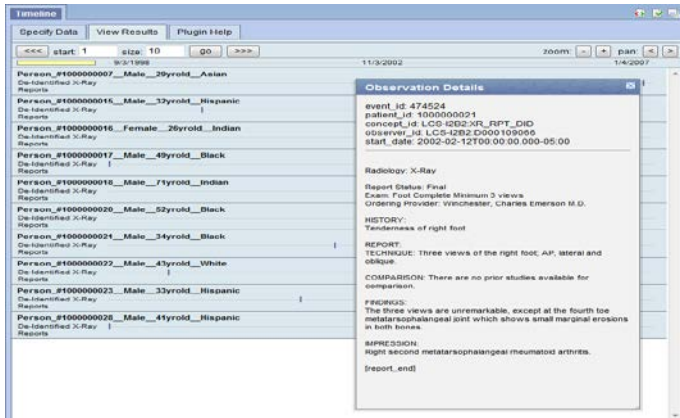
When Large numbers of patients found use Machine Learning Approach



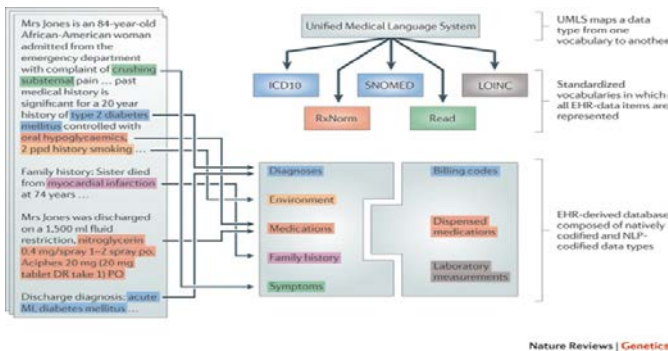


Creating Quality Data with Supervised Machine Learning

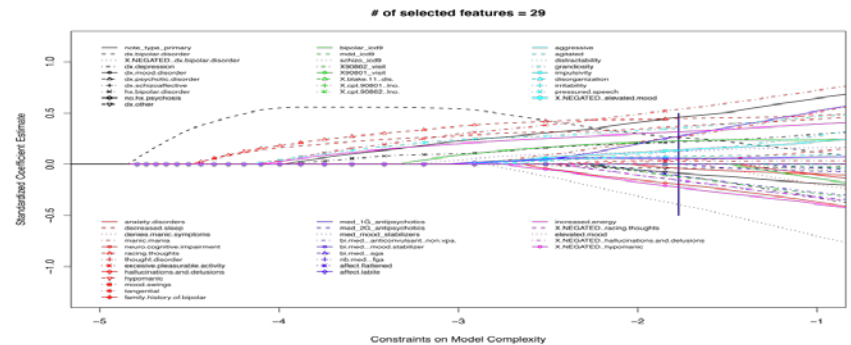
1. Create a gold standard training set.



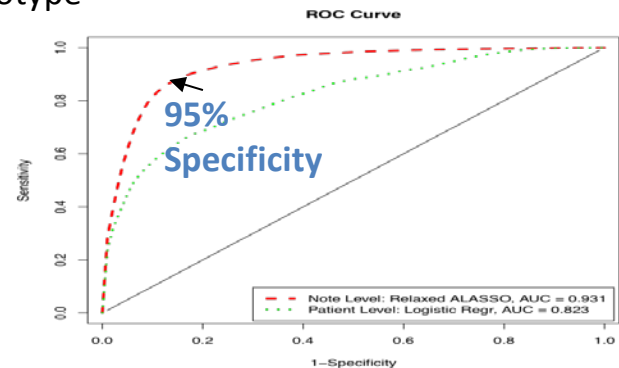
2. Create a comprehensive list of features (concepts/variables) that describe the phenotype of interest



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



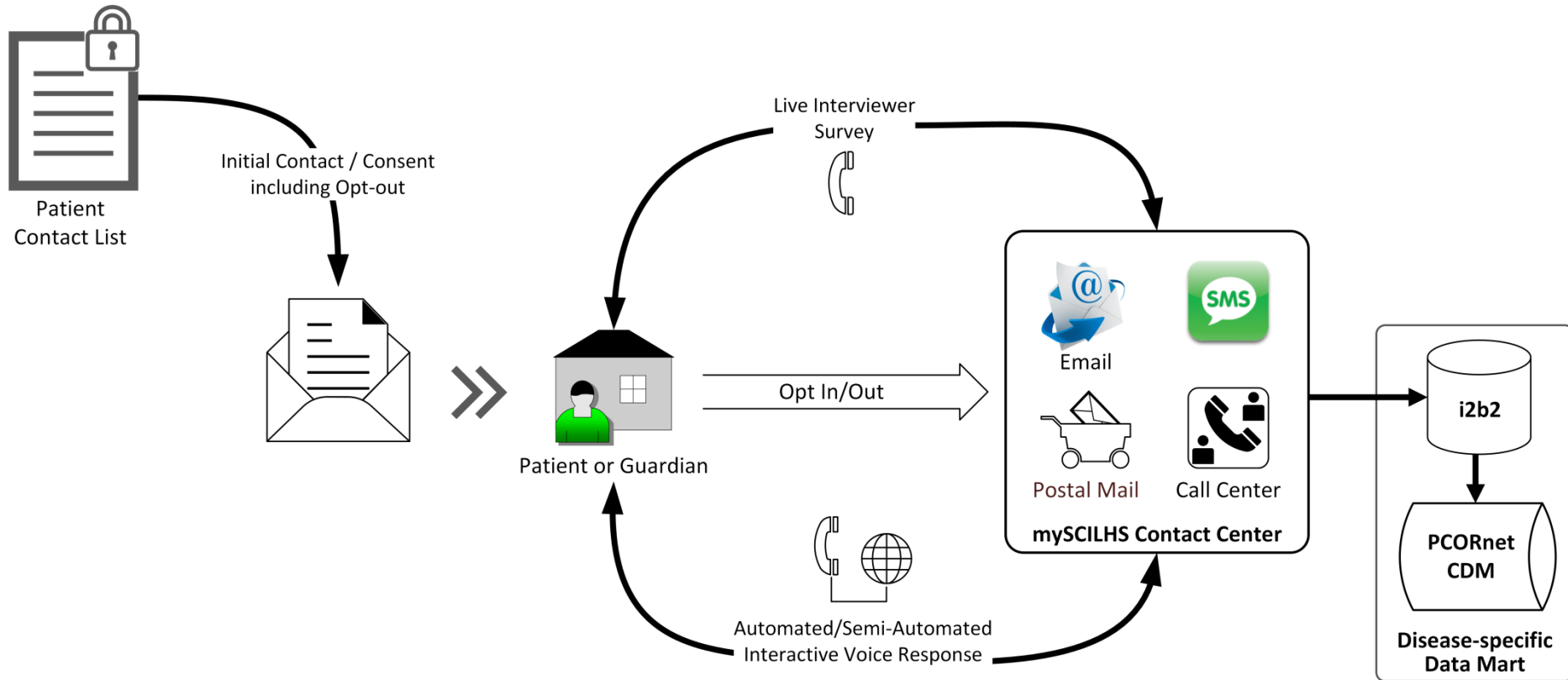
4. Apply the algorithm to all subjects in the superset and assign each subject a probability of having the phenotype



Key Challenges

- Governance of Data Types
- Sustainability
- Patient Privacy
- Patient Engagement

A full workflow from query to contact



Clinic

Firewall

Home



Multichannel Outreach

Multichannel Outreach

EHR

Care Delivery System

- Medical Records
- Claims
- Medications

Data

SCILHS

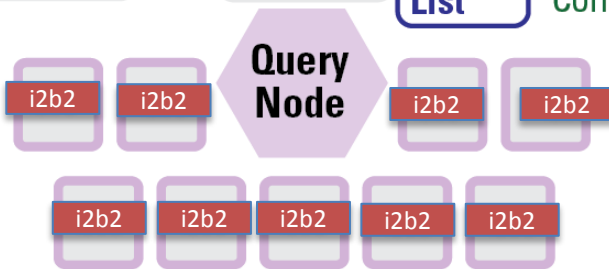
i2b2

Patient Reported Data,
Point of Care App Data

Patient
Cohort
List

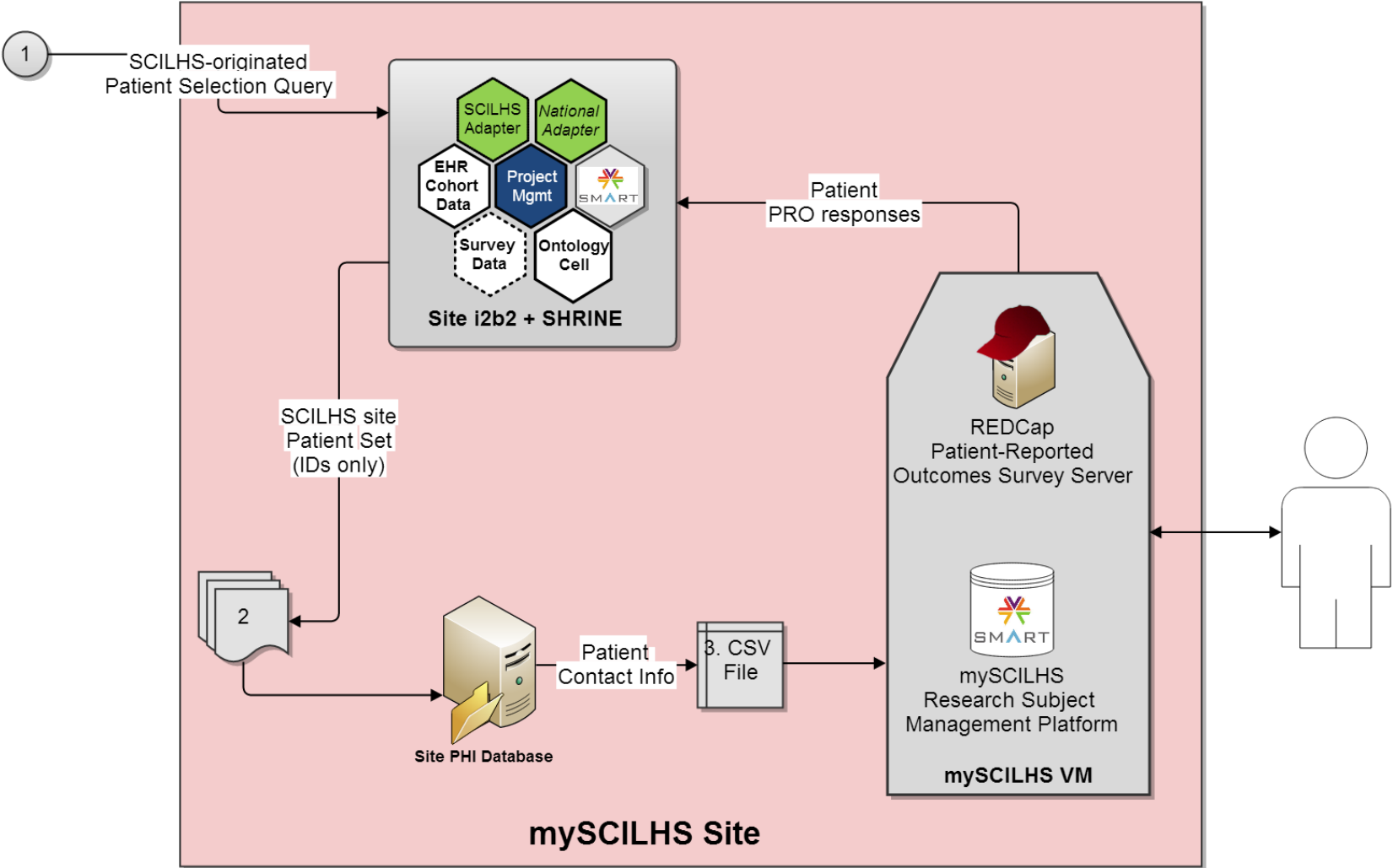
Linkage for
Contact Data

My.SCILHS



Other SCILHS Sites

mySCILHS



Designing the App Store for Health



SMART



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 designed 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-



Inspired by a

WIRED

18.12 Issue

Design Challenge



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

Patient info

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

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.org

4 What now?



Diet & exercise - can improve your cholesterol levels



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

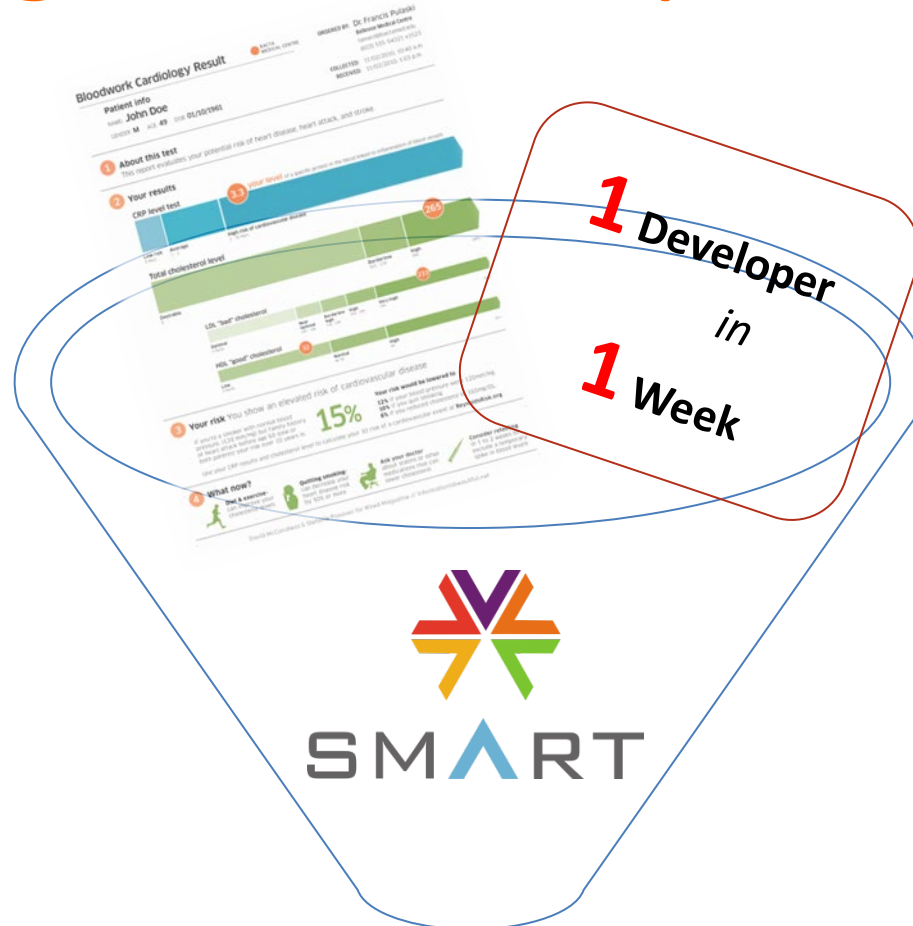


Ask your doctor about statins or other medications that can lower cholesterol



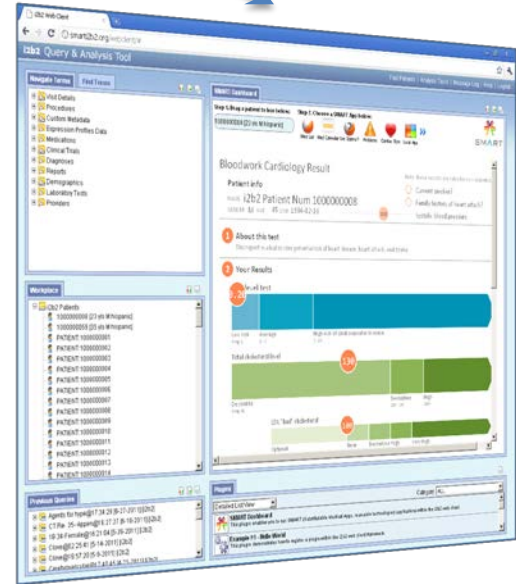
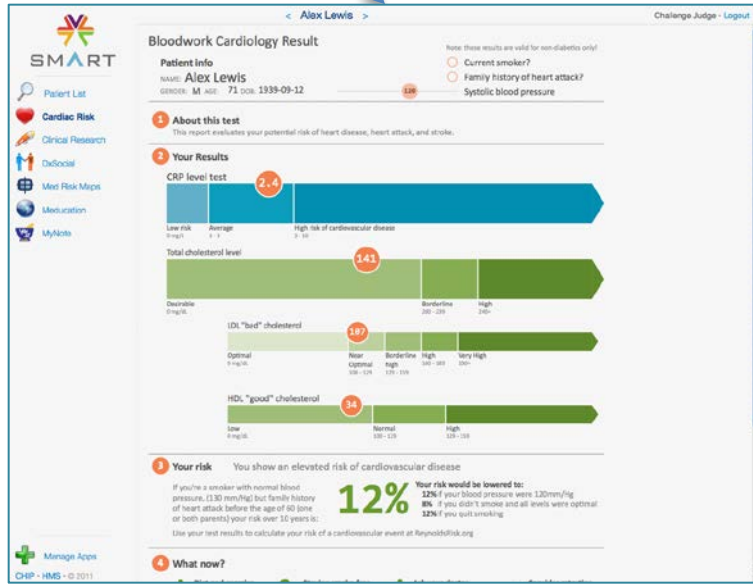
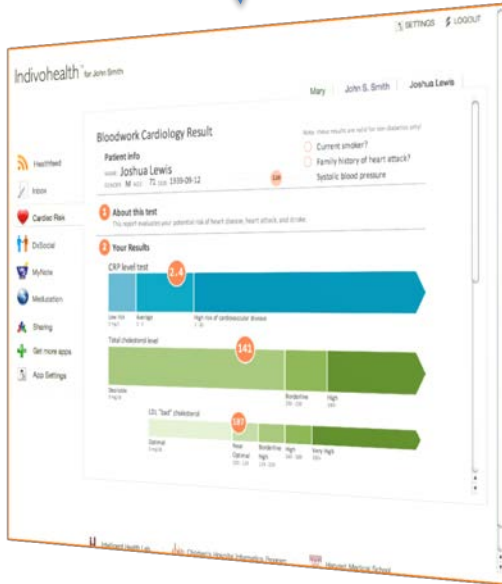
Consider retesting in 1 to 2 weeks to exclude a temporary spike in blood levels

1 Design + 1 Developer + 1 Week



SMART
Cardiac
Risk
App

1 SMART App in 3 SMART Systems



Trial Eligibility



Female **Male**

43 y

Boston, MA

rheumatoid arthritis +methotrexate

Cancel

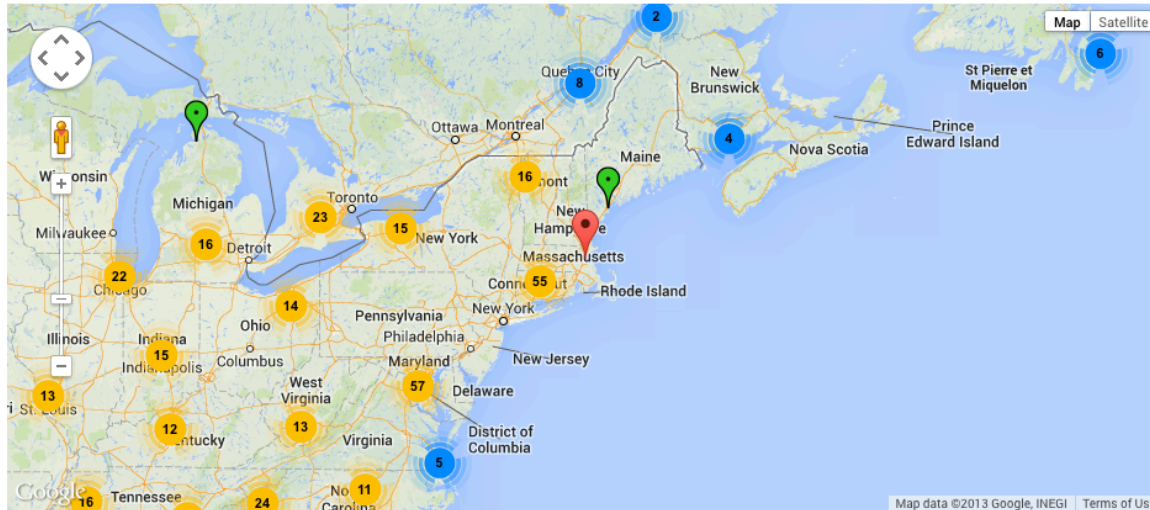
Refresh

Intervention / Observation

- Biological (14)**
- Device (1)
- Dietary Supplement (1)
- Drug (69)**
- Observational (14)
- Other (4)
- Procedure (1)

Trial phase

- N/A (8)**
- Phase 1 (8)**
- Phase 2 (23)**
- Phase 3 (17)**
- Phase 4 (26)**



Key Challenges

- Governance of Data Types
- Sustainability
- Patient Privacy
- Patient Engagement

The Plan

Common data platform (**i2b2**)

+

Federated queries across sites (**SHRINE**)

+

Point of care apps (**SMART**)

+

Patient-facing technologies (**RedCap, SMART, +**)

Tribute to...

■ I2b2 Core Team

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

■ SMART i2b2 Team

- Nich Wattanasin
- Kenneth Mandl
- Joshua Mandel

■ i2b2 – SHRINE Team

- William Simons
- Douglas MacFadden

■ tranSMART Team

- Paul Avillach
- Michael McDuffy

■ mySCILHS Team

- Marc Natter
- Isaac Pinol Catadau
- Stanley Boykin

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>)

SHRINE at Harvard (<http://shrine.catalyst.harvard.edu>)

SHRINE Software:

(<https://open.med.harvard.edu/display/SHRINE/Software>)

SMART i2b2 Homepage (<http://smarti2b2.org>)

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

THANK YOU