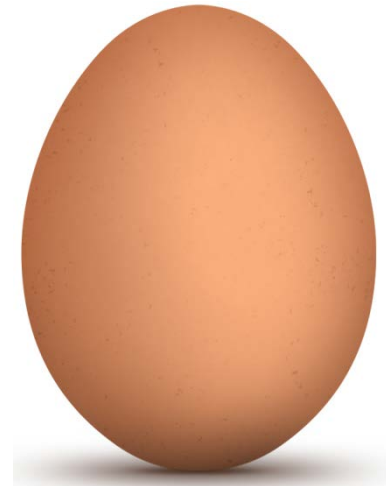




I2b2 Update

Data Managers Work Group

June 18, 2013



As Is



- I can't access the data we want for cohort discovery
- Wow it takes a long time to get an answer
- Do I really need an IRB to find out how many YY diagnosis or procedures we had at JHH?
- When can I have access to the Epic data?

Pilot Data



Name

Row Count

Observations

191,106,517

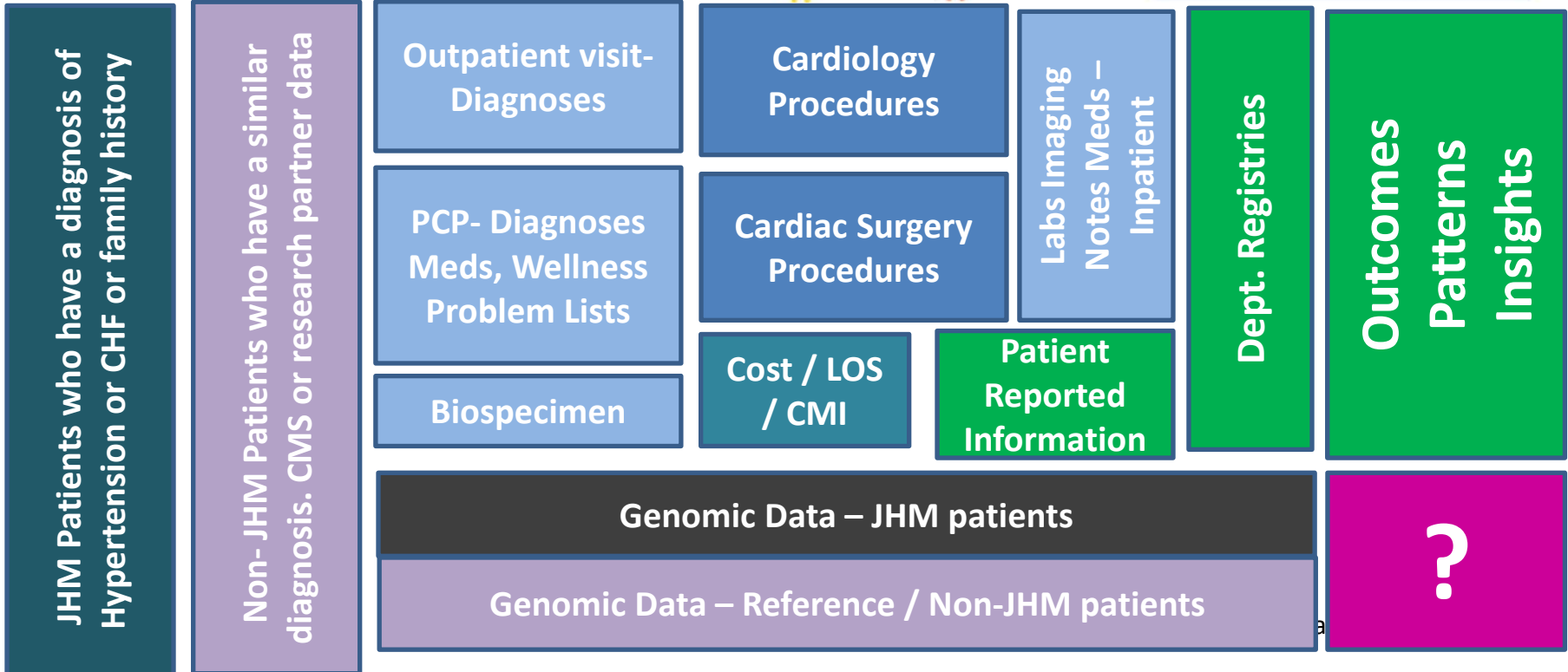
Visits

44,718,052

Patients

2,552,110

Data / Insight



Intelligence



Asthma

Average yearly health care cost
of a 50-year-old with Asthma: \$4,415

Personal Cost: \$872

Insurer Cost: \$3,542

Total yearly health care cost for
the 442 patients with Asthma:

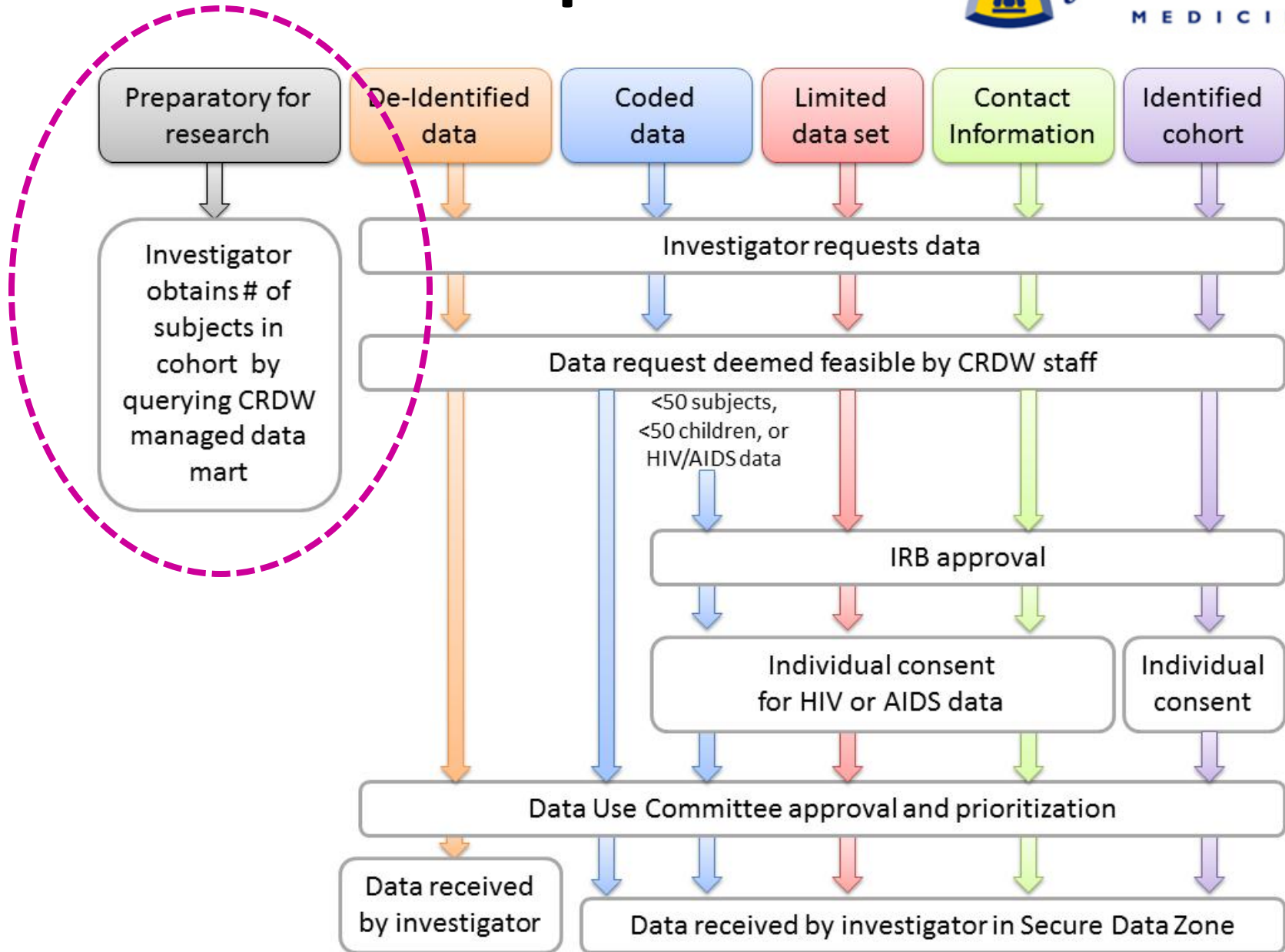
\$1,951,433



http://www.pearltrees.com/#/N-s=1_5949664&N-u=1_780832&N-p=56056080&N-play=1&N-f=1_5949664&N-fa=5949655

<http://vimeo.com/36354086>

Access - Concept

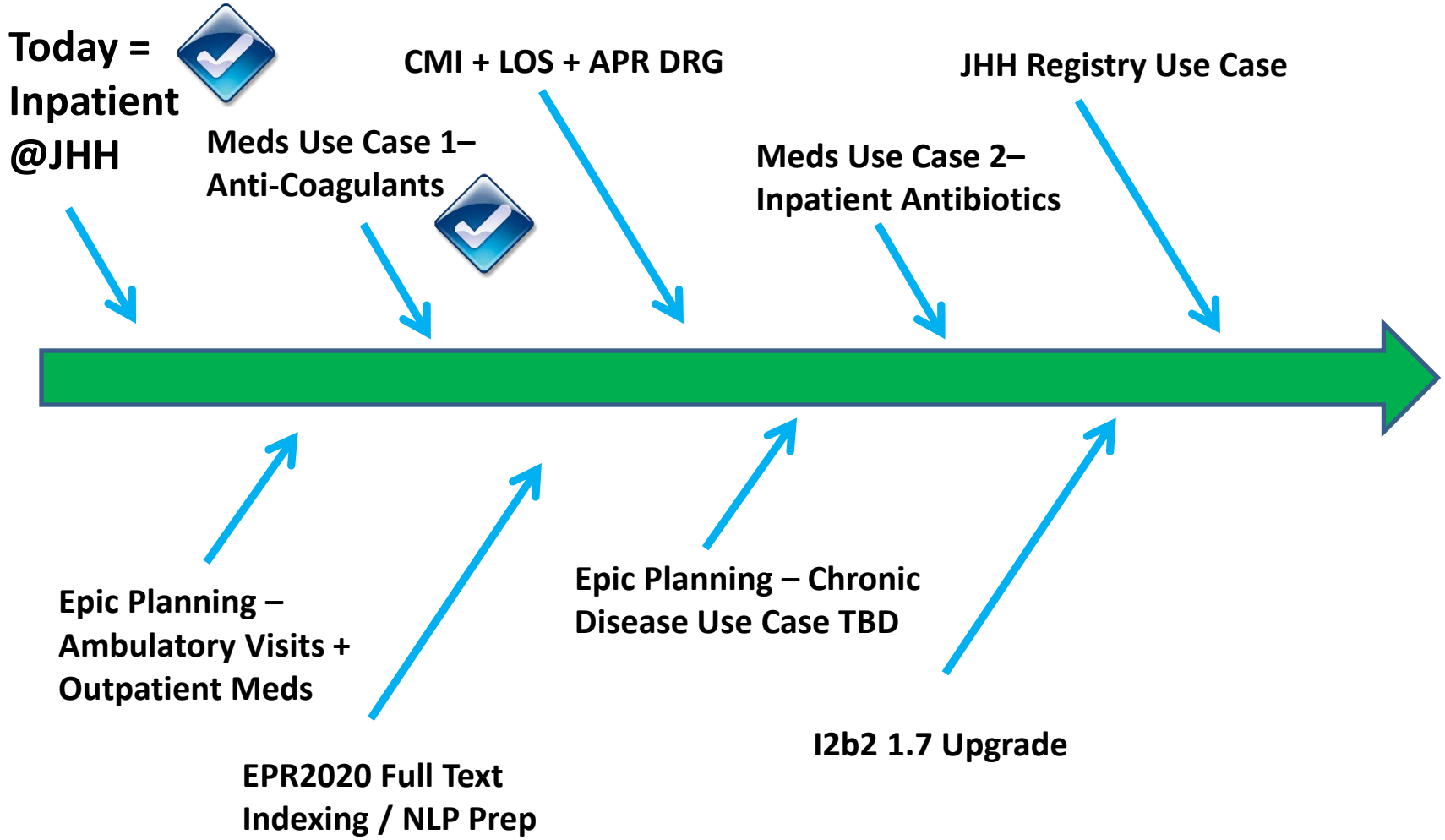


Draft Data



Data	Status	Q	Notes
Patient Mapping and Dimension	Completed	2.5M	Patients and visits based on EPR2020 identity backbone Including multi-institution and multi-MRN references.
Visit mapping and dimension	Completed	42M	Needs validation of inpatient v. outpatient flags
Provider dim.	Completed	0.35M	Contents of EPR/EPR2020 (CDT provider)
Provider Ontology	In progress		JHH and Bayview providers
Demographics	Completed	2.5M	Includes: Age, zipcode, race, gender
Diagnosis	Completed	40M	Based on ICD9 codes from datamart and IDX via EPR2020. Some gaps identified
Inpatient Procedures	Completed	2M	Based on ICD9 codes from datamart and IDX via EPR2020.
Outpatient Procedures	In Progress		Requires a CPT lookup table and procedure fact records
Lab Data	In Progress	3Years	Ontology and Numeric results complete. Standard Text results, complex test results in progress
Meds	Not Started	TBD	Dispensed meds from SCM (JHH only for now). Current plan is to use Multim ontology until rxNorm can be added.

Concept Road Map



EMR Data – Chronic Disease – Planned Business Objects Universes



Diabetes	Chronic Disease Management
Snapshot Date - Diabetes	Diabetes
Alcohol & Tobacco Use - Diabetes	Diabetes
Comorbidities - Diabetes	Diabetes
Numeric - Comorbidities - Diabetes	Comorbidities - Diabetes
Immunizations - Diabetes	Diabetes
Medications - Diabetes	Diabetes
Numeric - Diabetes	Medications - Diabetes
Most Recent Labs - Diabetes	Diabetes
Pregnancy - Diabetes	Diabetes
Visit History - Diabetes	Diabetes
Vitals - Diabetes	Diabetes

628 Objects / 47 tables / 1528 Rows across CHF, Diabetes, Hypertension Registries

Epic will be a phenomenal tool, and . . .

- **Set up required:** We need to check for Clarity, BOE, and future Data Warehouse connections between tables
- **Alignment:** We need to align our data access policies with JHM-available tools
- **Deidentified Access:** I2b2 can act as a interim bridge to access deidentified Epic data if interface connections are approved
- **Interim Data Warehouse:** I2b2 is currently acting as a synthetic data warehouse until the Epic EDW is funded and brought on line.

Barrett's Esophagus Requirements

– i2b2 and EPR2020



Provide a crosstab report approximately along these lines. Each cell would be a count value of *people* who had a particular finding. In the case of someone with multiple findings, they would appear in both detail cells, but not summed in the summary cells. For example, someone with a ICD9 finding of both 530.85 AND 150.1 would appear in the count in both the values in cell in rownumber 2 and 4, but only once in the rownumber 1.

Along top row are:

1. Count - number of Patients per diag grouping
2. Timepoints = unique patient-eventdates per diag grouping (eidforsoid-eventdate) not time
3. Specimen Types Associated via PathLab Text Findings as follows

Look For any of	Results Column Header
Normal OR squamous	Normal
Barrett and ('No Dysplasia' OR 'NonDysplastic')	BE-ND
Barrett and ('Indefinite%')	BE-IND
Barrett and ('Low%Grade')	BE-LGD
Barrett and ('High%Grade')	BE-HGD
'Adeno%'	AdenoCA
'Squamous Cell Carcinoma'	SQ-CA

2. Specimen types associated via MRN and in CATissue

S. Melzer ACS Review: Barrett's Esophagus



- 4.4 M patients and 26M visits
- 12.8 M documents / 7.2 M Radiology Reports
- 28M Lab Results

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

CARRA Registry and i2b2



The 56-site Childhood Arthritis & Rheumatology Research Alliance (CARRA) Registry and 3-site Harvard Inflammatory Bowel Diseases Longitudinal Data Repository now utilize i2b2 self-scaling registry technology (i2b2-SSR).

Table 1 Diagnosis at baseline visit, Childhood Arthritis & Rheumatology Research Alliance (CARRA) registry population (as of February 2012)

Diagnosis	N (%)
Juvenile idiopathic arthritis	4510 (72%)
Pediatric systemic lupus erythematosus	618 (10%)
Juvenile dermatomyositis	433 (7%)
Localized scleroderma	236 (4%)
Juvenile primary fibromyalgia	122 (2%)
Vasculitis	117 (2%)
Mixed connective tissue disease	112 (2%)
Sarcoidosis	38 (1%)
Systemic sclerosis	36 (1%)

CARRA Interface

12B2 SSR Self Scaling Registry Workbench Site: Hospital ABC
Group: CARRAnet Welcome: User123 (Logout) [Leave Feedback](#)

Ontology

- CARRA Studies
 - CARRAnet National Registry_1.0
 - History
 - Demographics And Disease History
 - Subject Disease History
 - 6. Antinuclear antibody test (ANA)
 - 7. What is subject's primary rheumatological disease diagnosis
 - 8. Date of Onset of disease symptoms (best estimate of month)
 - 9. Zip code of primary residence at onset of disease symptom
 - 10. Date first seen by pediatric rheumatologist
 - Demographics
 - ACR functional class
 - Family Medical History
 - Baseline
 - Visit
 - MEDICATION HISTORY
 - Other drugs of interest
 - Source
 - Medication review - by category
 - 3. Has subject ever taken any non-biologic immune modulators?
 - 4. Has subject ever taken any biologics?
 - Yes
 - No
 - Unknown
 - 5. Has subject ever taken any glucocorticoids (other than inhaled)
 - 2. Has subject been in any blinded drug trial(s) where treatment
 - SECTION 1: Biologics
 - SECTION 1: Non-biologic immune modulators and DMARDs
 - SECTION 1: Glucocorticoids
 - BASELINE SLE/MCTD Part 1
 - BASELINE SLE/MCTD Part 2
 - BASELINE Systemic Sclerosis
 - BASELINE Juvenile Dermatomyositis Disease
 - BASELINE Localized Scleroderma
 - Baseline Juvenile Idiopathic Arthritis
 - BASELINE Vasculitis
 - BASELINE Sarcoid
 - BASELINE JPPS
 - Status

Query Builder

New Query Add Query Group

1 Clear Exclude (X) 2 Clear Exclude (X) 3 Clear Exclude (X)

7. What is subject's primary rheumatological disease diagnosis Yes

OR AND OR AND OR

Drag terms here Drag terms here Drag terms here

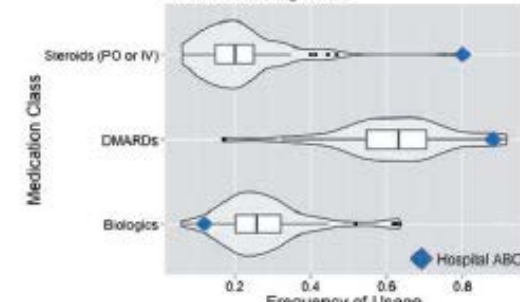
Run Query

Query Counts **Query Results & Data**

Subject List Reports & Graphs Export

My Dashboard: Hospital ABC

Immunosuppressive Usage Summary:
Hospital ABC - CARRAnet Aggregate Cohort: All Diagnoses



Medication Class

Steroids (PO or IV)

DMARDs

Biologics

Frequency of Usage

Immunosuppressive Usage Summary:
Hospital ABC - CARRAnet Aggregate Cohort: All SLE/MCTD

Query History

CARRA Design Principles i2b2



This scenario illustrates our objectives for a modular, collaborative, self-scaling registry providing minimal barriers to participation. We have focused our i2b2-SSR development efforts around five design principles:

1. Provide data contributors with full ownership of and access to their own data
2. Minimize barriers for data owners to collaboratively contribute their data to new or existing datasets
3. Support a tiered sharing model which provides a granular, permissioned, and audit-capable data sharing framework
4. Enable near real-time access to data, supporting a virtuous cycle in which immediate data access promotes further data contribution and collaboration
5. Encourage ongoing incorporation of outside datasets from multiple sources.

Ask



- Try it. If not you, who could this benefit?
- What is the problem you need solved?
- Genomics and bio specimen data
- Volunteer resources for micro projects

BACKUP MATERIAL

Obesity and Labs



Query Name: Diabetes Type2,Obesity,with HBa1c <=8.5%		
Temporal Constraint: Treat all groups independently		
Group 1		
Date From: none Date To: none Excluded? false Occurs X times: > 0 Relevance %: 100 Temporal Constraint: Treat Independently		
Path	Concept/Term	Other Information
Demographics \ Age \ 18-34 years old	18-34 years old	
Demographics \ Age \ 35-44 years old	35-44 years old	
Demographics \ Age \ 45-54 years old	45-54 years old	
Demographics \ Age \ 55-64 years old	55-64 years old	
Group 2		
Date From: none Date To: none Excluded? false Occurs X times: > 0 Relevance %: 100 Temporal Constraint: Treat Independently		
Path	Concept/Term	Other Information
Diagnoses \ ... \ Diabetes mellitus, type II [non-insulin dependent {...}] or unspecified type with hyperosmolarity, not stated as uncontrolled	Diabetes mellitus, type II [non-insulin dependent type] [NIDDM type] [adult-onset type] or unspecified type with hyperosmolarity, not stated as uncontrolled	
Diagnoses \ Metabolic and immunity disorders \ Obesity and other hyperalimentation \ Obesity	Obesity	
Group 3		
Date From: 04/01/2012 Date To: 12/5/2012 Excluded? false Occurs X times: > 0 Relevance %: 100 Temporal Constraint: Treat Independently		
Path	Concept/Term	Other Information
Labtests \ Chemistry \ Hemoglobin \ GHBA1C (Group:GHBA1C) \ Hgb a1c (LOINC:4548-4)	HGB A1C (LOINC:4548-4)	LE : 8.5 %

The screenshot shows a query builder interface. At the top, the query name is "Diabetes Type2,Obesity,with HBa1c <=8.5%". Below it, there's a tree view showing the selected concepts: "Diabetes mellitus, type II" and "Obesity". A logic diagram below the tree shows two boxes connected by "AND" operators, with a central box labeled "one or more of these". At the bottom, there's a "Print Query" button circled in red, and a label "3 Groups".

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

Compute Time: 18 secs

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

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

patient_count: 189

i2b2 Driving Biology projects



JOHNS HOPKINS
MEDICINE

<https://www.i2b2.org/disease/index.html>

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

I2b2 Pilot

- I2b2 pilot is a proven path –
 - <https://www.i2b2.org/>
 - Cohort Finding Tool
 - Some basic data visualization tools
 - Flexible and simple Data Mart Schema
 - Customizable Ontologies
 - Many open source additional modules
 - SHRINE as a possible link to non-JHM institution data sets
- Shorten cycle time from idea to data for analysis

Self Service + Collaboration