

Deploying SNOMED CT longitudinal history data for clinical research

James R. Campbell MD

W. Scott Campbell PhD

James McClay MD

University of Nebraska Medical Center

Omaha, Nebraska, USA

SNOMED CT Exposition

October 29, 2015

Disclosure

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

Learning Objectives

- Understand and recount an interoperable information model for past medical history, family history and social history data employing post-coordinated SNOMED CT attributes and value sets
- Appreciate the mapping and data extraction challenges to populating the research data warehouse with longitudinal data from one common EHR vendor
- Review and understand decision analytical and clinical research use cases for employing longitudinal history data in support of enterprise and national objectives

Agenda

- Meaningful Use case for structured longitudinal health data in the electronic health record
- Research use cases for longitudinal clinical data
- Situations: Past Medical History
- Clinical observables: Smoking and alcohol history
- Situations: Surgical-procedural History
- Situations: Family History

Meaningful Use Stage 2 Vocabulary Requirements for Semantic Interoperability

- Demographics (80% stage 2): LOINC, HL7/OMB code set
- Longitudinal data; 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, SNOMED CT
- Medication orders: RxNORM, SNOMED CT
- Orders: LOINC, RxNORM
- Immunizations (Immunization registries): CVX, MVX
- Procedures: CPT, HCPCS
- Documents: LOINC

Office National Coordinator 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
- Enhance clinical research within Learning Healthcare Environment

Interoperable Learning Health Ecosystem



How to merge data from the EHR into integrated, interoperable research and public health datasets?

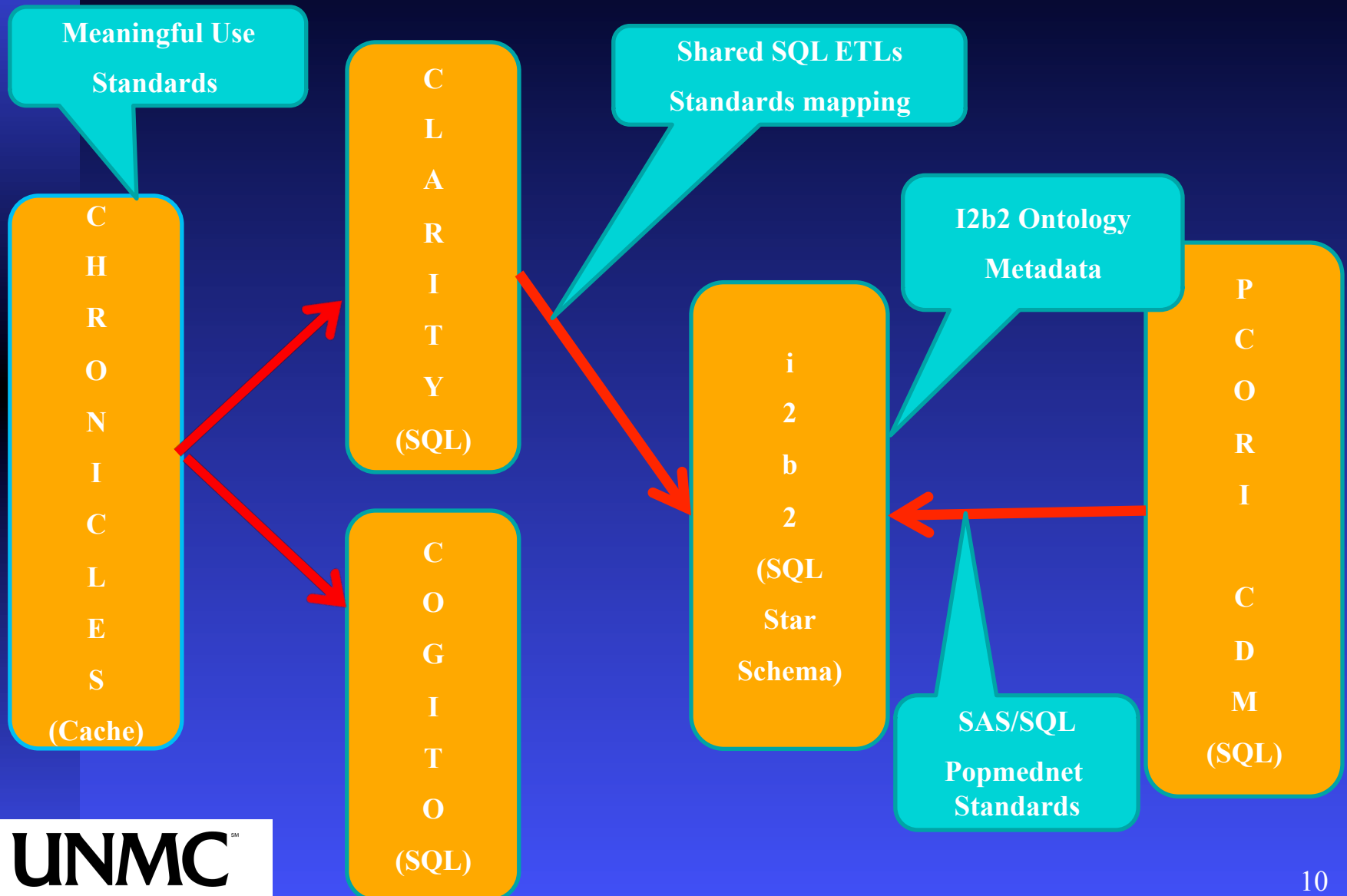
PCORI CDRN 4: Greater Plains Collaborative

- 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
- Indiana University
- University of Missouri
Columbia

GPC Architecture for Research Interoperation

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

UNMC Research Data Development Model



UNMC Research Standards Transformation Architecture

	Epic EHR	Clarity ETL Maps	i2b2 Ontology Metadata
Demographics		LOINC(site) US ed SNOMED CT	US OMB LOINC
Problem List	US ed SNOMED CT(IMO) ClinFind, Events, Situations, Social context		(SNOMED CT)
Encounter diagnoses			
Billing diagnoses	ICD-9-CM(IMO)		ICD-(*)-CM
Social history		LOINC US ed SNOMED CT	LOINC, SNOMED CT Observables
Laboratory results	LOINC(site)		LOINC
Clinical findings		LOINC US ed SNOMED CT	SNOMED CT observables LOINC
Medication orders			RxNORM
Prescriptions	RxNORM(FDB)		
Medications Administered			RxNORM NDC
Medications Dispensed	NDC (Surescripts)		NDC
Procedures	CPT(Epic) ICD-9-PCS HCPCS		CPT ICD-9-PCS HCPCS(Colorado)
Immunizations	CVX MVX(site)		CVX MVX
Documents		LOINC	LOINC



U
M

U
M

(Some) Research Data Warehouse Requirements

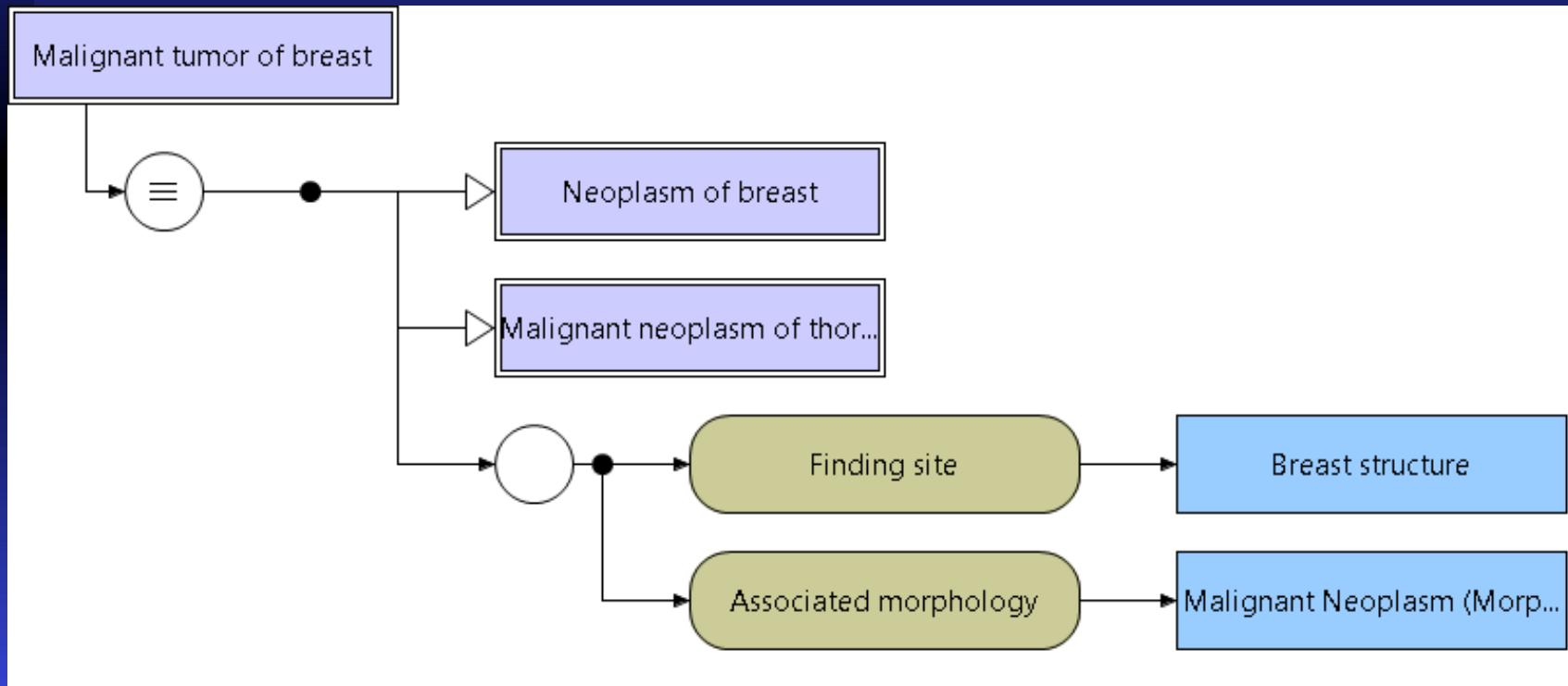
- Organization, presentation and query data access understandable to the clinical researcher
- Data aggregation queries easy to do; relevant to research questions; standards driven but not requiring knowledge of the standard
- Query of quantitative results and conversion of data units must be supported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

Research Data Warehouse: Query use cases

- How many patients with COPD have smoked more than 30 pack-years by age and gender?
- How many patients are at risk of breast cancer due to first degree family history or BRCA syndrome?
- How many patients have had a splenectomy, have functional asplenia or sickle cell anemia and should be vaccinated for pneumococcus?

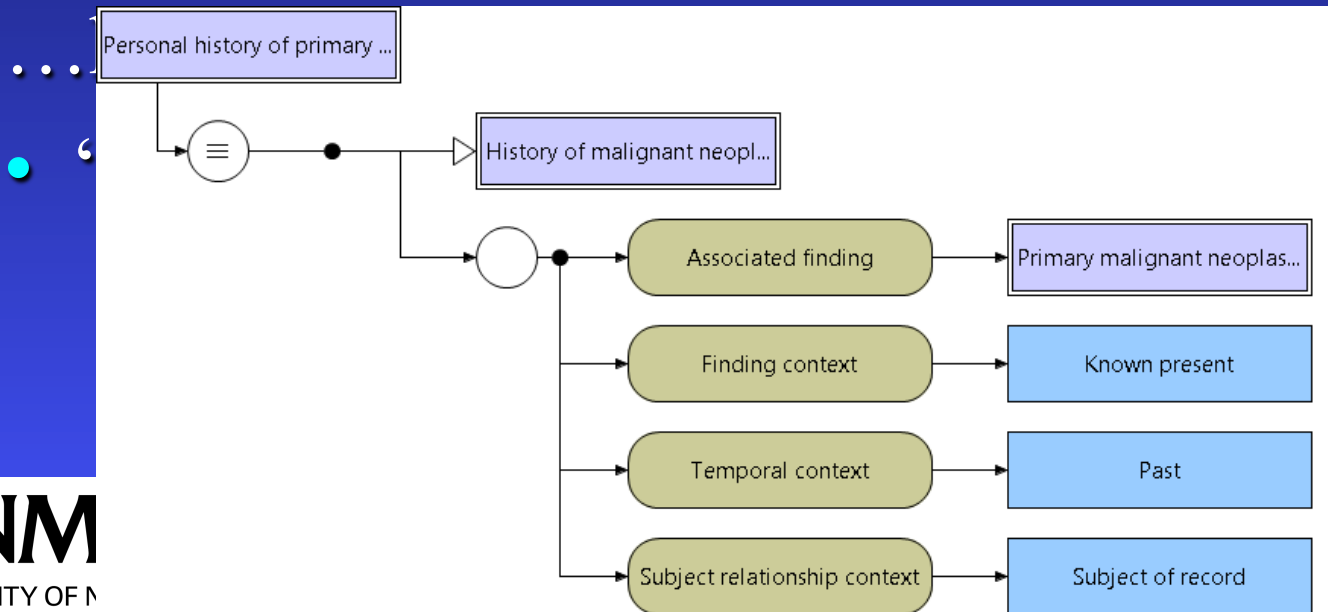
Historical data: Situation with explicit context

- Problem: “Breast cancer”



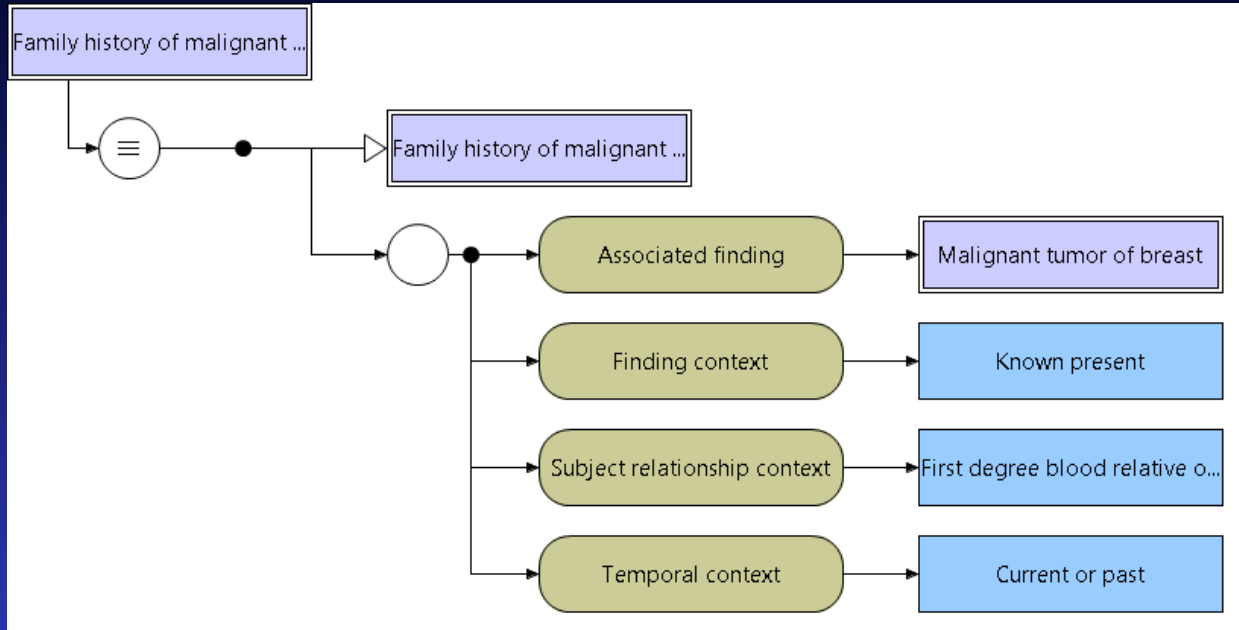
Historical data: Situation with explicit context

- Problem: “Breast cancer”
- ...time passes
- “History of breast cancer”



“other”

Historical data:



text

- “Family history of breast cancer in mother”

Epic Data Model: Past Medical History

Problem list record

- IMO maps:
- <<417662000|History of clinical finding|
- <<|413350009|Finding with explicit context|

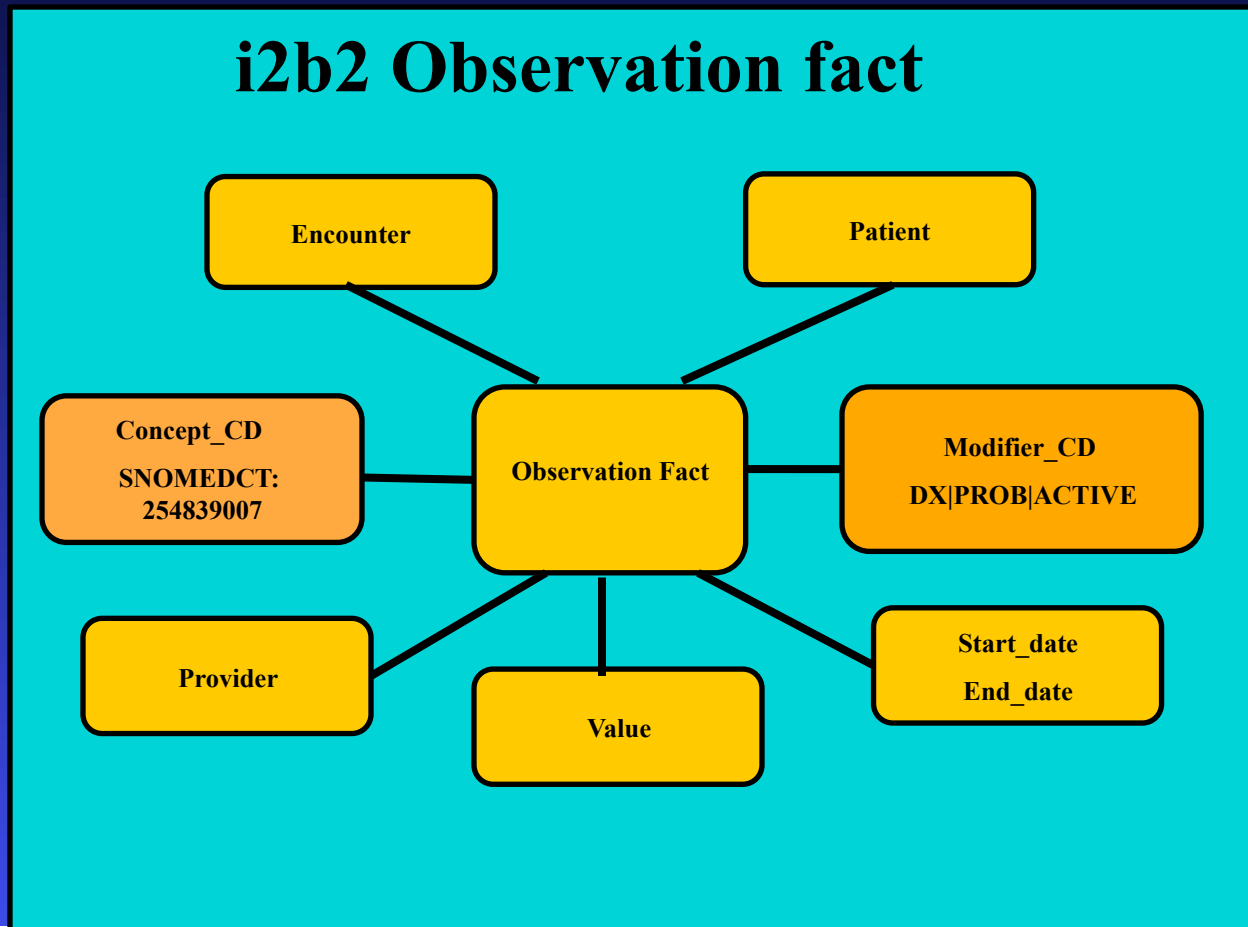
Past history record

- UNMC maps:
<<404684003|Clinical finding|
- Age or date of onset
- Comment
- Source of history

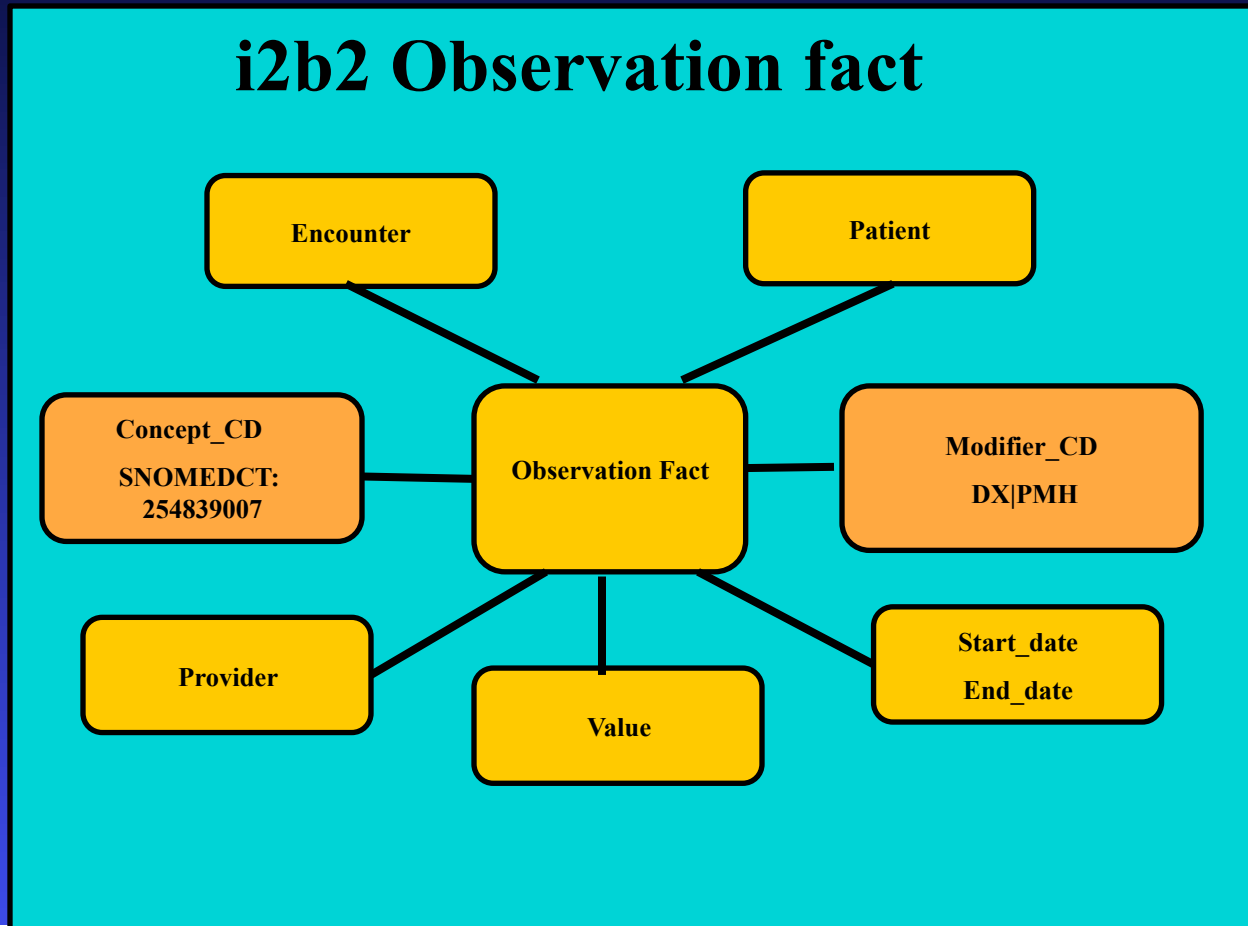
Pertinent negative history

- <<404684003|Clinical finding|
- Source of history

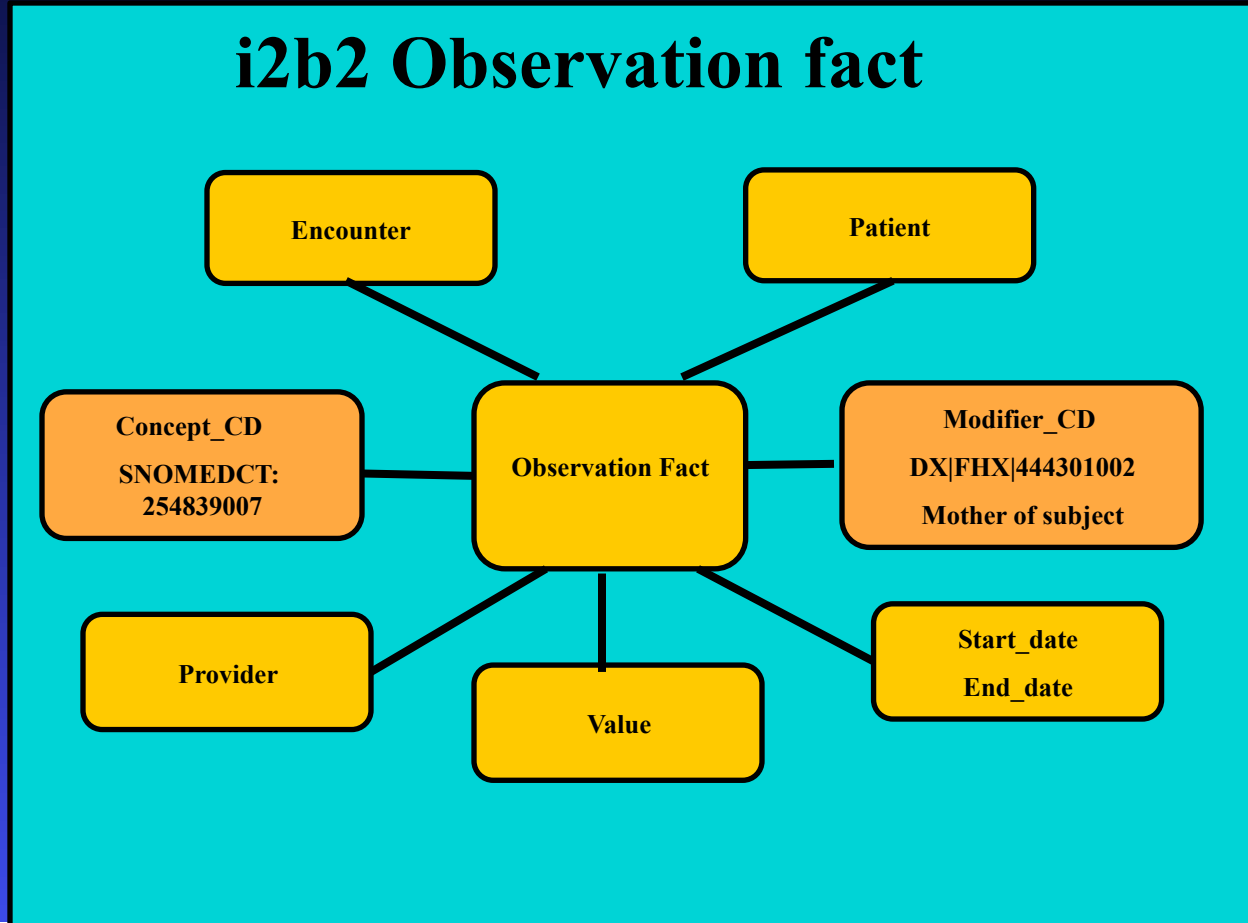
Current problem: “Breast cancer”



“Subject history of breast cancer”



“Mother’s history of breast cancer”



Social History Observables

- Social history data is extracted routinely by Epic to Clarity tables
- Tables contain encounter records for observables about patient habits and behaviors: smoking, alcohol use, sexual preference, OB history and others
- Some of these clinical observables can be found in LOINC OR SNOMED CT.....
- ONC specified value set for ‘Tobacco use status’ as SNOMED CT; mapped for Meaningful Use by Epic and published by NLM

Epic Data Model: Social history

Social history

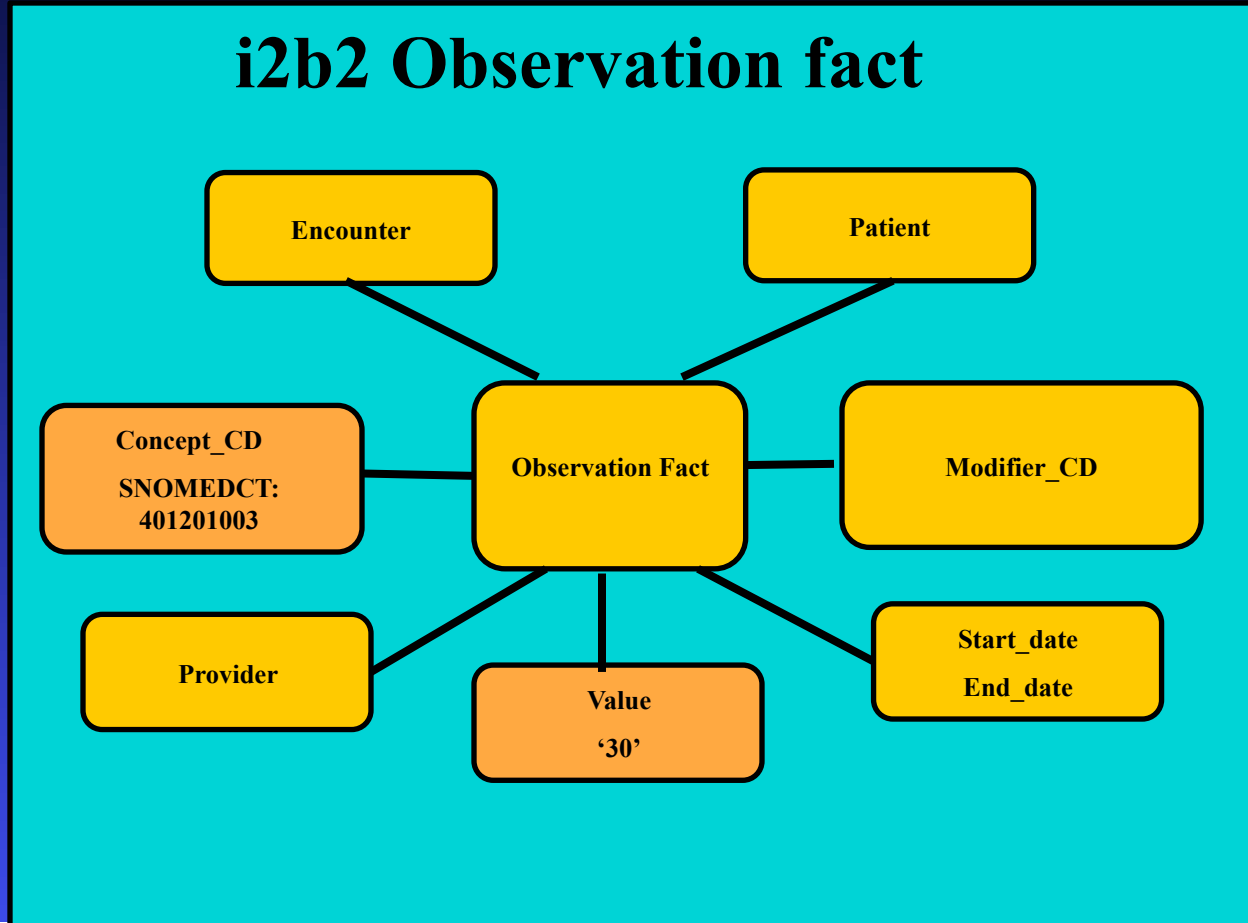
- Uses tobacco
- Packs per day
- Tobacco years
- Quit date
- Cigarettes
- Pipes
- Cigars
- Snuff
- Chews
- Uses alcohol
- Alcohol oz per week
- Uses IV drugs
- Frequency of drug use
- Sexually active
- Female sexual partner
- Male sexual partner
- Uses a condom
- ...

Epic Data Model: Social history

Social history

- UNMC maps:
- Uses tobacco = LOINC 72166-2 | Pattern of tobacco use |
- Packs per day X Tobacco years = 401201003 | Cigarette pack years (observable) |
- Quit date
- Cigarettes
- Pipes
- Cigars
- Uses alcohol
- UNMC maps: Alcohol oz per week = 228311005 | Pattern weekly alcohol consumption (observable) |
- Uses IV drugs
- Frequency of drug use
- Sexually active
- Female sexual partner
- ...
- Uses a condom

“Subject pack-year history = 30”



“How many COPD patients do we have?”

i2b2 Query & Analysis Tool Project: i2b2 Demo User: Jim Campbell Find Patients | Analysis Tools | Message Log | Help | Change Password | Logout

Navigate Terms

Find Terms

Search by Names

Search by Codes

Containing chronic obstructive lung

Find

Clinical findings (SNOMEDCT) (TC) [2,683,176 facts; 461,851 patients]

- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Acute exacerbation of chronic obstructive airways disease (disorder) [1,598 facts; 1,598 patients]
- Chronic obstructive airway disease with asthma (disorder) [28 facts; 28 patients]
- Chronic obstructive pulmonary disease with acute lower respiratory infection (disorder) [1 fact; 1 patient]
- End stage chronic obstructive airways disease (disorder) [<10 facts]
- Mild chronic obstructive pulmonary disease (disorder) [18 facts; 18 patients]
- Moderate chronic obstructive pulmonary disease (disorder) [21 facts; 21 patients]
- Obliterative bronchiolitis (disorder) [66 facts; 64 patients]
- Pulmonary emphysema (disorder) [1,456 facts; 1,436 patients]
- Severe chronic obstructive pulmonary disease (disorder) [53 facts; 50 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]
- Chronic obstructive lung disease (disorder) [13,533 facts; 11,507 patients]

Query Tool

Query Name:

Temporal Constraint:

Treat all groups independently

Group 1			Group 2			Group 3		
Dates	Occur	Exclude	Dates	Occur	Exclude	Dates	Occur	Exclude
Treat Independently			Treat Independently			Treat Independently		
Chronic obstructiv								

Run Query

Clear

Print Query

Query Status

Query Status area containing a large empty text box for displaying query results or status messages.

“How many COPD patients have smoked ≥ 30 pack-years?”

The screenshot displays a clinical data query tool interface. On the left, a 'Find Terms' pane lists various medical terms with their respective fact and patient counts. On the right, the 'Query Tool' pane shows a query configuration window for 'Cigarette pack-years'. A dialog box is open, allowing the user to specify the value and operator for the query.

Find Terms List:

- Finding by method (finding) [17,363 facts; 16,714 patients]
- Finding by site (finding) [1,873,206 facts; 410,000 patients]
- Finding of grade (finding) [<10 facts]
- Finding related to physiologic substance (finding) [<10 facts]
- Finding reported by subject or history provider (finding) [4,563 facts; 4,419 patients]
- General clinical state finding (finding) [31,887 facts; 27,131 patients]
- Jaundice (finding) [397 facts; 385 patients]
- Neurological finding (finding) [158,925 facts; 86,059 patients]
- Papule (finding) [719 facts; 717 patients]
- Prognosis/outlook finding (finding)
- Stiffness (finding) [597 facts; 582 patients]
- Swelling (finding) [4,535 facts; 3,417 patients]
- Wound finding (finding) [12,013 facts; 8,603 patients]
- Clinical Measurements [138,315,971 facts; 1,766,250 patients]
- Demographics [56,135,786 facts; 1,767,089 patients]
- Diagnoses (ICD9CM) [8,730,179 facts; 538,292 patients]
- DRG [46,435 facts; 29,993 patients]
- Encounters [36,833,446 facts; 990,442 patients]
- Event (event) (SNOMEDCT) (TC) [6,617 facts; 6,020 patients]
- GPC [88,690,137 facts; 331,959 patients]
- Laboratory Measurements [172,320,116 facts; 1,766,130 patients]
- Medications [81,072,933 facts; 257,288 patients]
- Medications (UNMC)
- Observable entity (SNOMEDCT observables) (TC) [413,748 facts; 241,043 patients]
- Organism (organism) (SNOMEDCT) (TC)
- PCORI [118,184,821 facts; 1,768,243 patients]
- Procedures [7,222,044 facts; 278,050 patients]
- Procedures (SNOMEDCT) (TC) [4,340 facts; 3,810 patients]
- Situations (SNOMEDCT) (TC) [69,211 facts; 46,168 patients]
- UNMC One Chart [158,470,785 facts; 609,277 patients]
- Schmid [7,518,074 facts; 72,909 patients]
- Social history [1,780,120 facts; 265,702 patients]
- Cigarette pack-years [1,478 facts; 291 patients]
- Pattern of weekly alcohol consumption [168,812 facts; 22,911 patients]
- Tobacco consumption [1,609,830 facts; 265,702 patients]
- UNMC Laboratory Results [158,690,665 facts; 585,501 patients]

Query Tool Configuration:

- Query Name: [Empty]
- Temporal Constraint: Treat all groups independently
- Group 1: Chronic disease of respiratory system (disorder) [3]
- Group 2: Cigarette pack-years [1,478 facts; 291 patients]
- Group 3: [Empty]

Choose value of Cigarette pack-years (Test:SNOMED CT:401201003)

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

No Value

By Value

Please select operator: **GREATER THAN OR EQUAL (\geq)**

Please enter a value: **30**

Click on a bar segment to help specify a value or range:
Range in pk-yrs

Units = **pk-yrs**

Buttons: OK, Cancel

Drop a term on here

Social History Observables

- Social history data is extracted routinely by Epic to Clarity tables
- Tables contain encounter records for observables about patient habits and behaviors: smoking, alcohol use, sexual preference, OB history and others
- Some of these clinical observables can be found in LOINC OR SNOMED CT.....
- ONC specified the value set for 'Tobacco use status' as SNOMED CT; mapped for Meaningful Use by Epic and published by NLM

Social History Observables: LOINC (72166-2:Tobacco use status)



Value Set Authority Center
U.S. National Library of Medicine



Value Set Name	Current Cigarette Smoker	
OID	2.16.840.1.113883.3.600.2390	
Type	Extensional	
Definition ID (release/update date)	20140430	
Steward	Quality Insights of Pennsylvania	
	This update was generated by VSAC and has not yet been endorsed by Quality Insights of Pennsylvania	
Expansion ID	20130614	
Code	Description	Code System
266920004	Trivial cigarette smoker (less than one cigarette/day) (finding)	SNOMEDCT
365981007	Finding of tobacco smoking behavior (finding)	SNOMEDCT
365982000	Finding of tobacco smoking consumption (finding)	SNOMEDCT
428041000124106	Occasional tobacco smoker (finding)	SNOMEDCT
428061000124105	Light tobacco smoker (finding)	SNOMEDCT
428071000124103	Heavy tobacco smoker (finding)	SNOMEDCT
449868002	Smokes tobacco daily (finding)	SNOMEDCT
56578002	Moderate smoker (20 or less per day) (finding)	SNOMEDCT
56771006	Heavy smoker (over 20 per day) (finding)	SNOMEDCT
65568007	Cigarette smoker (finding)	SNOMEDCT
77176002	Smoker (finding)	SNOMEDCT



UNMCSM

UNIVERSITY OF NEBRASKA
MEDICAL CENTER

ONC: “Pattern of tobacco use”

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

Navigate Terms Find Terms

- Diagnoses (ICD9CM) [8,700,118 facts; 888,202 patients]
- DRG [46,435 facts; 29,993 patients]
- Encounters [36,833,446 facts; 990,442 patients]
- Event (event) (SNOMEDCT) (TC) [6,617 facts; 6,020 patients]
- GPC [88,690,137 facts; 331,959 patients]
- Laboratory Measurements [172,500 facts; 1,000 patients]
- Location
- Medications (UNMC)
- Medications [81,072,933 facts; 2,000 patients]
- Observable entity (SNOMEDCT)
- Organism (organism) (SNOMEDCT)
- PCORI [118,184,821 facts; 1,760 patients]
- Procedures (SNOMEDCT) (TC)
- Procedures [7,222,044 facts; 270 patients]
- Situations (SNOMEDCT) (TC) [6,000 facts; 60 patients]
- UNMC One Chart [158,470,785 facts; 1,000 patients]
- Schmid [7,518,074 facts; 72,000 patients]
- Social history [1,780,120 facts; 1,000 patients]
- Cigarette pack-years [1,400 facts; 1,000 patients]
- Pattern of weekly alcohol [1,000 facts; 1,000 patients]
- Tobacco consumption [1,000 facts; 1,000 patients]
- Type of Tobacco - 22911 [1,000 facts; 1,000 patients]
- UNMC Laboratory Results [1,000 facts; 1,000 patients]

Query Tool

Query Name: _____

Temporal Constraint: Treat all groups independently

Group 1	Group 2	Group 3

Choose value of Tobacco consumption (Test:LOINC:72166-2)

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

No Value

By Value

Please select a value:

- 449868002 Current everyday smoker
- 428041000124106 Current some day smoker
- 77176002 Current smoker status unknown
- 8517006 Former smoker
- 266919005 Never a smoker

OK Cancel

drop a term on here

New Group

Epic Data Model: Surgical (Procedural) History

Problem list record

- IMO maps:
- <<443938003|Procedure carried out on subject|
- <<416940007|Past history of procedure|
- <<387713003|Surgical procedure|

Surgical history record

- UNMC Maps:
- <<71388002|Procedure|
- Laterality code <<182353008|Side|
- Procedure date
- Comment
- Source of history

Epic Data Model: Family History

Problem list record

- IMO maps:
- <<416471007|Family History of clinical finding|
- Status (active, resolved..)
- Comment

Family history record

- UNMC maps:
- Relationship to patient <<303071001|Person in the family|
- <<404684003|Clinical finding|
- Age of onset
- Vital status
- Age at death

Questions?
Comments?