Cerner Research: Health Facts

- Rob Taylor, Scientist, Research & Analytics

Health Facts® Content

In this section we will…

- …identify and describe the data elements available in *Health Facts*®
- …review an example *Health Facts*® query

Health Facts® Content: Facility Characteristics

- US Census Region and Division (all represented)
- Bed size category
- Teaching status
- Urban/rural community setting
- Part of hospital system
- Acute care status
- Cardiac cath lab status
  - Full cath lab
  - Diagnostic cath lab only

Health Facts® Content: Registration/Admin Data

- Time-stamped admission and discharge
- Age
- Gender
- Race
- Payer, where available
- Admission source
- Discharge caresetting, where available
- Discharge disposition
  - Includes home, SNF/NH/LTC, died, etc.

Health Facts® Content: Diagnosis and Procedure Codes

- Codes
  - ICD-9/ICD-10 diagnosis and procedure codes from billing data
    - UB-92 (pre-2008): Up to nine diagnosis codes, one designated as primary
    - UB-92 (pre-2008): Up to six procedure codes, one designated as primary
    - UB-04 (post-2008): Up to 18 diagnosis codes, one designated as primary
    - UB-04 (post-2008): Up to seven procedure codes, one designated as primary
  - Subset of facilities in new architecture (2009 to present) contributing directly from Millennium
  - Present on Admission (POA) fields, when available, post 2008

Health Facts® Content: Dates/Times

- Admission and discharge are date and time stamped
- Medication orders are date and time stamped
- Gen. lab orders, draw times, and test completion are date and time stamped
- Microbiology orders, specimen collection times, and completion are date and time stamped
  - Typically updated to the most recent result
- Clinical events/assessments are date and time stamped
- Procedures codes have calendar dates only
  - Orders for IV sedation has been used to refine time for PCI
  - Pharmacy order/dispensing caresetting could be used to refine (e.g. pre-op, OR, post-anesthesia recovery)

Health Facts® Content: Laboratory

- Ordered test and time
- The detail lab procedure name, character or numeric result, unit of measurement
  - All mapped lab procedures (>2500)
- Dates and times captured on each laboratory record may include: ordered, drawn, received, completed, cancelled, performed, verified (HIPAA blinded)
- Care setting
- The medical specialty for the ordering physician, per order
- Reference range values, where applicable
  - ULN values have been leveraged in multiple projects to define abnormal and severely abnormal values (e.g. 3X ULN) for that specific lab at that point in time

Health Facts® Content: Microbiology

- Same basic information as gen lab (order time, specimen collection time, specialty, caresetting)
- Specific test, source and site used to define qualifying cultures
  - Blood, urine, respiratory and sterile sites such as CSF, joint fluid straightforward to identify
    - For blood cultures, have the ability to exclude common contaminants or require two positive cultures for those organisms on the list
  - Others such as cSSSI and cIAI require some decisions as to what to include due to risk of contaminants
    - For example culture from J-P drain may be acceptable if immediately post-op
    - cSSSI swab of wound may be acceptable if no skin contaminant isolated
- The organism name if there was a positive finding or NULL if not
- Type of result, for example, preliminary, final, amended final
- The antimicrobial used for the susceptibility testing

Health Facts® Content: Pharmacy

- Medical specialty for the ordering physician, per order
- Type of clinical care provided at the locations where the medication order was requested and dispensed
  - Many of these are specialty specific in addition to caresetting such as ICU
- Route of administration
- Frequency of the order
- Dates/times for the order: start, stop, discontinued (HIPAA blinded)
- Total quantity of doses dispensed per order
- Internally, Cerner uses Multum to roll up drugs into therapeutic

Health Facts® Content: Leveraging Care Settings

- Critical care is currently inferred indirectly:
  - Care setting on pharmacy/lab orders
    - For analysis, the most sensitive variable is binary
      - For example, ICU exposure within defined timeframe Y or N?
    - Calculation of length of exposure (ICU LOS) can be attempted, but is challenging
  - Treatment has also been used to infer critical care setting
    - Mechanical ventilation or vasopressor use, regardless of care setting
  - Combination of above (either)

- Care setting descriptions on pharmacy orders and dispensing can give insight into how a given drug or class is being used. Examp

ICU - Neurology	Medical/Surgical	Transplant	Emergency Room
ICU - Surgical	**Neurology**	**Oncology**	**Dialysis**
Internal Medicine	**Step-Down Unit - Cardiac**	**Operating Room**	**Cardiac Cath**

Health Facts® Content: Clinical Assessments

- Vital Signs and Physiologic Measures
  - Height, Weight, Blood Pressure, Heart Rate, Pulse, Temperature, BSA, BMI } **APACHE** now feasible
  - LOC and Glasgow Coma Scale
  - The dates and times for the event start, end (HIPAA blinded)
  - The result value captured in several possible forms such as text, number, date depending on the type of event, unit of measurement
  - A critical low or high, normal low or high, and normalcy description when applicable
  - The medical specialty for the performed personnel per event if available
- Patient Symptoms
  - Pain assessments
  - Vomiting, diarrhea
- Health History
  - Previous hospitalizations, Smoking cessation, Influenza vaccination status

Health Facts® Content: Leading Outcomes Used

- Length of stay
- Length of stay, survivors
- In-hospital mortality
- 30 day readmission rate
- Total charges
- Estimated costs (using HCUP cost-to-charge ratios)

Query Example: Alternate Definitions for Juvenile Arthritis

All Encounters with any diagnosis of Juvenile Arthritis (ICD-9 Code 714.30, 714.31, 714.32, 714.33) who are 17 or younger

	# Hospitals	# Patients	# Encounters
Overall	74	1,719	6,290

Year	# Hospitals	# Patients	# Encounters
2000	4	34	84
2001	13	64	106
2002	13	63	122
2003	16	55	104
2004	17	62	112
2005	20	74	121
2006	24	256	450
2007	33	454	1,103
2008	38	483	1,106
2009	40	584	1,359
2010	42	566	1,239
YTD 2011	24	249	384
		Total	6,290

Patient Type	# Hospitals	# Patients	# Encounters
ER	38	142	170
Inpatient	31	213	305
Other	66	1,474	5,815
		Total	6,290

Number of Encounters per Patient		# Patients	
1		841	
2		281	
3 or 4		244	
5 or more		353	

All Encounters with any diagnosis of Juvenile Arthritis (ICD-9 Code 714.30) who are 17 or younger

	# Hospitals	# Patients	# Encounters
Overall	74	1,493	4,511

Year	# Hospitals	# Patients	# Encounters
2000	4	32	78
2001	11	49	80
2002	12	44	70
2003	16	39	58
2004	16	46	74
2005	18	54	76
2006	24	202	299
2007	33	358	750
2008	38	398	790
2009	39	505	1,041
2010	41	481	917
YTD 2011	23	198	278
		Total	4,511

Patient Type	# Hospitals	# Patients	# Encounters
ER	38	137	163
Inpatient	29	198	283
Other	64	1,248	4,065
		Total	4,511

Number of Encounters per Patient		# Patients	
1		775	
2		261	
3 or 4		210	
5 or more		247	

- Majority of encounters use 714.30 Polyarticular juvenile rheumatoid arthritis, chronic or unspecified
- Most patients have multiple encounters captured

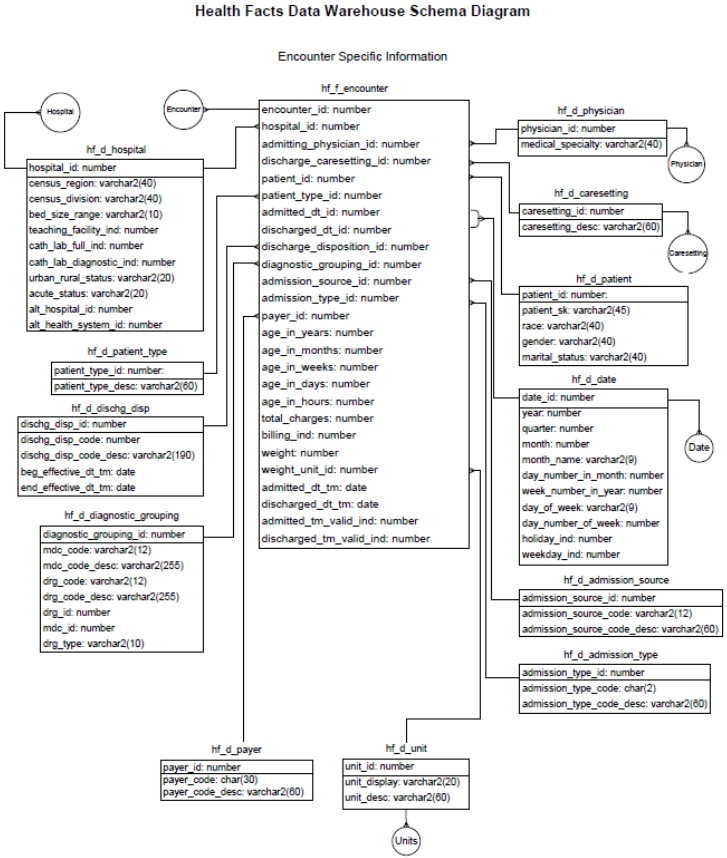
14

Working with *Health Facts*® Data

In this section we will discuss…

- …the *Health Facts*® schema
- … *Health Facts*® fact and dimension tables
- …key data relationships
- …techniques for working with the data

Health Facts® - Schema



Health Facts® - Schema

- Star schema
  - “Fact” tables
    - Naming convention: begins with hf_f_{*}
    - Contained detailed information about the “encounter” (i.e., “visit”, “admission” level)
  - “Dimension” (or “Dim”) tables
    - Naming convention: begins with hf_d_{*}
    - These are lookup tables for code values – descriptions of attributes

Health Facts®– Dimension Table Values

- Values from clients' source systems are mapped to values in the "Dim" tables
  - Example – hospital locations:

	Raw data provided by hospital		Information in hf_d_caresetting table in Health Facts	
Hospital_id	**Code**	**Description**	**Code**	**Description**
1	27389	4 North	13	Cardiology
1	4309	8 West	146	Surgery
1	748821	2 South	143	Step-Down Unit – Surgical
2	38755	Green 3	143	Step-Down Unit – Surgical
2	232898	Blue 6	146	Surgery
2	102393	Orange 7	13	Cardiology

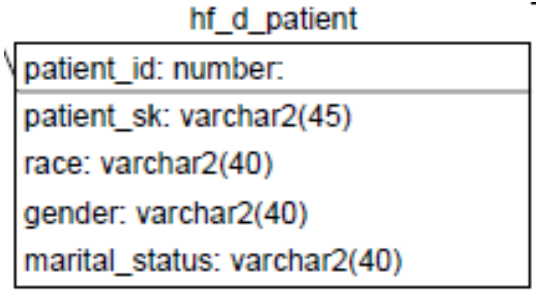
Health Facts®– Dim Table Values

- NULL, Unknown, Not Mapped – all equate to missing

PAYER_ID	PAYER_DESC
1	Blue Cross/Blue Shield
2	CHAMPUS (Military dependents)
3	Free, Research
4	HMO/Managed Care (undesignated)
5	International Patient
6	Medicare
7	Medicare Managed Care (undesignated)
8	Medicare Psychiatric
9	Medicare Rehab.
...etc	...etc
20	Unknown / Missing / Invalid
21	Worker's Compensation
22	NULL
23	Not Mapped

Patient Table – hf_d_patient

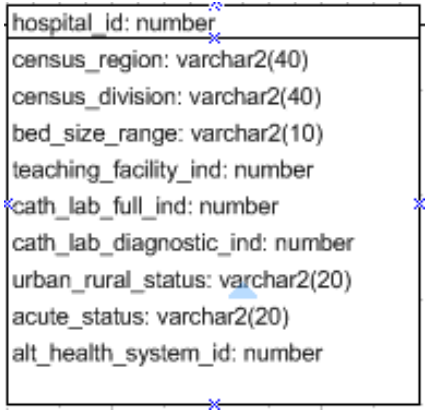
- A new row is inserted in this table each time an attribute changes



PATIENT_ID	PATIENT_SK	RACE	GENDER	MARITAL_STATUS
38882	7748	Unknown	Female	Married
47749	7748	Caucasian	Female	Married
57693	7748	Caucasian	Female	Divorced
48390	49303	Asian	Male	Married

Hospital Table – hf_d_hospital

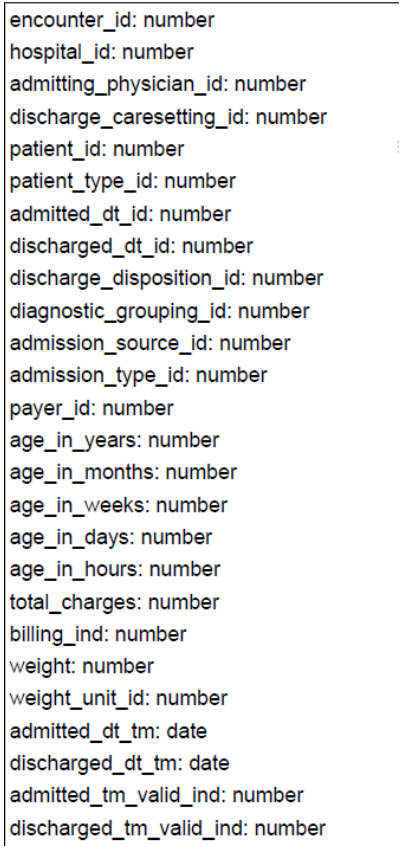
Hospital characteristics



Fact Tables

- “Fact” tables
  - Hf_f_encounter
  - Hf_f_diagnosis
  - Hf_f_procedure
  - Hf_f_lab_procedure
  - Hf_f_medication
  - Hf_f_microbiology
  - Hf_f_micro_susceptibility
  - Hf_f_clinical_event
  - Hf_f_surgical_case

Encounter Table – hf_f_encounter



Encounter/Patient Relationship

Data in hf_f_encounter

ENCOUNTER_ID	PATIENT_ID	ADMITTED_DT_TM
2188	38882	10/19/04 15:35
34447	47749	3/23/05 19:57
50009	57693	4/18/06 12:34
78933	57693	6/8/08 16:12
48293	48390	4/2/07 10:17
93888	48390	8/25/08 19:54

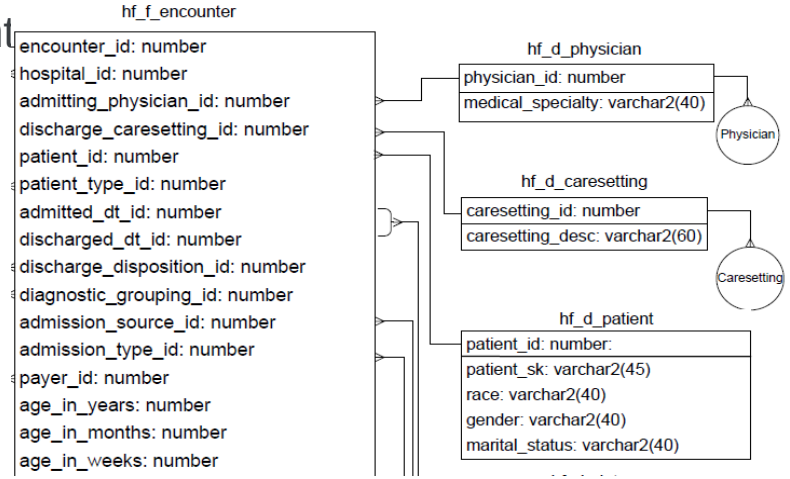
Two unique patients

Data in hf_d_patient

PATIENT_ID	PATIENT_SK	RACE	GENDER	MARITAL_STATUS
38882	7748	Unknown	Female	Married
47749	7748	Caucasian	Female	Married
57693	7748	Caucasian	Female	Divorced
48390	49303	Asian	Male	Married

Longitudinal Linkage

1. Obtain encounter_ids of interest
2. Obtain patient_ids for those encounter_ids
3. Obtain unique patient_sks for those patients
4. Obtain all patient_ids for those unique patient_sks
5. Obtain all encount



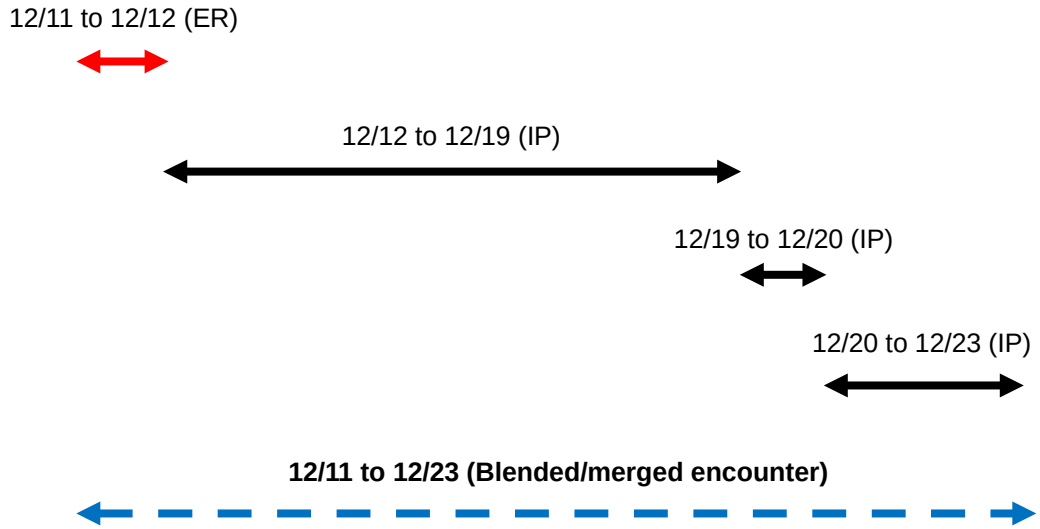
ER and Inpatient Encounters

- ER encounters – patient_type_id = 84
- Inpatient (IP) encounters – patient_type_id = 87
- Other patient_type_ids correspond to outpatient encounters (clinic visits, outpatient visits to physicians, outpatient surgery, etc…)

- In Health Facts, there are two ways that related ER and IP encounters appear
  - Two separate encounters, one with patient_type_id = 84, followed soon after by a second encounter with patient_type_id = 87
  - A single inpatient encounter with an admission_source_id = 7 (‘Emergency’)

Adjacent Encounters

Consider blending adjacent encounters into a single episode of care



Health Facts®– hf_f_diagnosis

- Diagnosis data
  - UB-92 (pre-2008): Up to nine ICD-9 diagnosis codes, one designated as primary
  - UB-04 (post-2008): Up to 18 ICD-9/ICD-10 diagnosis codes, one designated as primary
  - Use diagnosis_type_id = 83 for billing diagnosis
  - Some clients' systems have additional types, such as reason for visit and working diagnosis.
    - For some clients, up to 143 diagnoses have been extracted from the clinical system
  - Primary diagnosis: DIAGNOSIS_PRIORITY = 1
  - Non-primary diagnoses: DIAGNOSIS_PRIORITY > 1
  - Comorbidities are not necessarily coded on every encounter
    - Look back to previous encounters for evidence of the comorbidity
    - Infer using lab and/or medication data if warranted

Health Facts®– hf_f_diagnosis

- UB-04 has Present on Admission (POA) field, potentially useful for identifying comorbidities:
  - Y = Present at time of inpatient admission
  - N = Not present at time of inpatient admission
  - U = Documentation insufficient to determine if condition present on admission
  - W = Provider unable to clinically determine if condition present on admission
  - NULL – not provided, or did not apply (i.e., pre – UB-04)

Health Facts®– hf_f_procedure

- Codes from billing - Up to six procedure codes, one designated as primary
- For some clients, up to 61 procedures have been extracted from the clinical systems.
- Procedures have a date stamp (not date-time)
- Primary procedure: PROCEDURE_PRIORITY = 1
- Non-primary procedure: PROCEDURE_PRIORITY > 1

Health Facts®– hf_f_lab_procedure

- Health Facts contains labs that go through the hospital's general lab system
  - Contains any fingerstick glucoses entered into the lab system.
  - If bedside test results (including FSBG) are recorded only on paper, these will not appear in Health Facts.
  - Specialty tests sent to outside labs (e.g., insulin levels) will not appear in Health Facts unless the results were entered into the general lab system.

Health Facts®– hf_f_lab_procedure

- Order_lab_procedure_id – panel name (if any)
  - Example: Chem-7, Metabolic panel, CBC with differential
- Detail_lab_procedure_id – individual test name
  - Example: Na, K, Cl, CO2, BUN, SCR, glucose
- Upper and lower normal ranges have already been adjusted for age and gender.

Health Facts®– hf_f_lab_procedure

- Example:

ORDER_LAB_PROCEDURE_DESC	DETAIL_LAB_PROCEDURE_DESC	LAB_DRAWN_DT_TM	NUMERIC_ RESULT	UNIT_DESC	NORMAL_ RANGE_ LOW	NORMAL_ RANGE_ HIGH
Metabolic Panel Basic	Blood Urea Nitrogen	2/16/2009 3:30	16	Milligrams per Deciliter	5	20
Metabolic Panel Basic	Creatinine, Serum	2/16/2009 3:30	0.85	Milligrams per Deciliter	0.6	1.4
Metabolic Panel Basic	Potassium, Serum	2/16/2009 3:30	4.6	Millimoles per Liter	3.5	5.5
Metabolic Panel Basic	Calcium, Serum	2/16/2009 3:30	9.1	Milligrams per Deciliter	8.5	10.5
Metabolic Panel Basic	Chloride, Serum	2/16/2009 3:30	99	Millimoles per Liter	98	108
Metabolic Panel Basic	Sodium, Serum	2/16/2009 3:30	137	Millimoles per Liter	135	145
Metabolic Panel Basic	Glucose, Serum	2/16/2009 3:30	115	Milligrams per Deciliter	70	109
Metabolic Panel Basic	Anion Gap	2/16/2009 3:30	7	Millimoles per Liter	0	0
Metabolic Panel Basic	Carbon Dioxide CO2 Total, Serum	2/16/2009 3:30	31	Millimoles per Liter	22	29
Hemoglobin A1C	Hemoglobin A1C	6/3/2005 17:40	8	Percent	0	6.9
PT & PTT/APTT	Partial Thromboplastin Time	6/3/2005 17:40		Second	20	32
PT & PTT/APTT	Prothrombin Time	6/3/2005 17:40	12.3	Second	9	12.6
PT & PTT/APTT	INR	6/3/2005 17:40	1.12	NULL		

Health Facts®– Labs

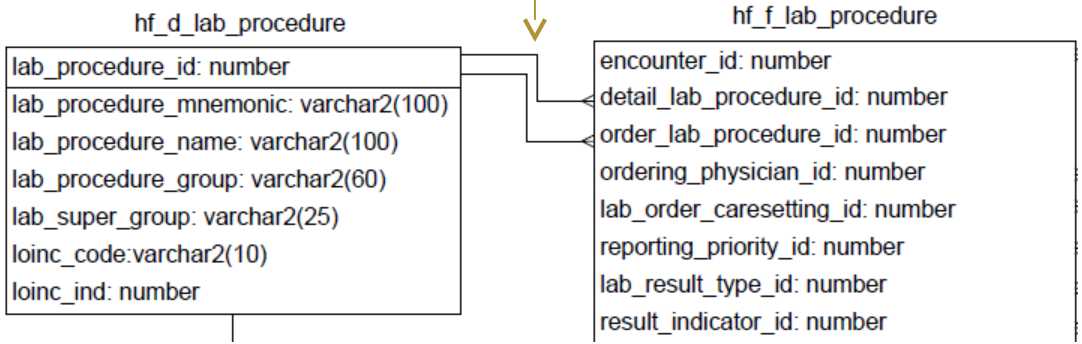
Obtain lab_procedure_ids of interest – Example:

```
Select *
From hf_d_lab_procedure
Where lower(lab_procedure_name) like '%glucose%'
```

LAB_PROCEDURE_ID	LAB_PROCEDURE_NAME	LAB_PROCEDURE_GROUP
888	Chem Panel OP3 (K,Glucose,Creatinine)	Chem Panel
1025	Glucose CSF	Glucose Test, CSF
1027	Glucose GTT, Serum 2 hr	Tolerance Test
1028	Glucose GTT, Serum 3 hr	Tolerance Test
1029	Glucose GTT, Serum 4 hr	Tolerance Test
1030	Glucose GTT, Serum 5 hr	Tolerance Test
1031	Glucose GTT, Serum 6 hr	Tolerance Test
1034	Glucose Stick, Urine	Glucose Test, Urine
1035	Glucose Stick/Meter Whole Blood POC	Glucose Test
1036	Glucose, Synovial	Glucose Test, Body Fluid
1037	Glucose, Body Fluid	Glucose Test, Unk Spec
1038	Glucose, Serum 2 hr pp	Glucose Test
1039	Glucose, Serum Fasting	Glucose Test
1040	Glucose, Serum Random	Glucose Test
52	Glucose, Serum	Glucose Test
2287	Glucose, Capillary	Glucose Test
2288	Glucose, Dialysate	Glucose Test
…etc…	…etc…	…etc…

Health Facts® – Labs

Then join lab_procedure_id from
hf_d_lab_procedure
to detail_lab_procedure_id in hf_f_lab_procedure



Health Facts® – Labs

- *Always* examine the distribution of units – example, Hematocrit:

UNIT_DESC	COUNT(*)	MIN(NUMERIC_RESULT)	AVG(NUMERIC_RESULT)	MAX(NUMERIC_RESULT)	STDDEV(NUMERIC_RESULT)
NULL	59188	0	0.7	55	3.7
Not Mapped	32090	0	34.2	71.5	5.8
Percent	1723362	0	32.9	79	6.1

- Investigate units that have a different distribution than the typical units of measure for that test (investigate by hospital_id, test name, or both)

select hospital_id, unit_desc, count(*), min(numeric_result), avg(numeric_result), max(numeric_result), stddev(numeric_result)

from HCT1

where lower(UNIT_DE

group by hospital_key

HOSPITAL_ID	UNIT_DESC	COUNT(*)	MIN(NUMERIC_RESULT)	AVG(NUMERIC_RESULT)	MAX(NUMERIC_RESULT)	STDDEV(NUMERIC_RESULT)
36	NULL	58559	0	0.32	8.32	0.07
98	NULL	2	24.1	24.10	24.1	0.00
131	NULL	2	39	39.00	39	0.00
137	NULL	3	26	28.33	33	4.04
140	NULL	620	12	35.33	55	7.43
148	NULL	1	32.9	32.90	32.9	0.00
152	NULL	1	32.3	32.30	32.3	0.00

Health Facts® – Labs

- Consult with an internal expert to decide on “impossible” thresholds (values considered incompatible with life – example: glucose 5000 mg/dL, creatinine of 62 mg/dL)
- Handling values outside the “impossible” thresholds
  - If the lab test is not an outcome measure, and the percentage of impossible values is very low, consider deleting the values
  - Otherwise – here are some options
    - Examine the values by hospital, by test name, or both, to see if there is some pattern
    - Examine tests that are part of the same orderable (same order_lab_procedure_id and lab_drawn_dt_tm) to see if there are other

Health Facts – Glucose and Hemoglobin A1C

- Glucose
  - Lab_procedure_ids for blood/serum glucose (including glucose tolerance tests): 1027, 1028, 1029, 1030, 1031, 1035, 1021, 1022, 1024, 1050, 1038, 1039, 1040, 1045, 1046, 1047, 1048, 1049, 1652, 1692, 1748, 1749, 1750,1751, 52, 2292, 2293, 2287, 2290, 2291
  - Health Facts contains a relatively low number of glucoses designated as ‘fasting’.
  - If fasting blood/serum glucose is of interest, also consider including glucose values drawn early in the AM, e.g., 5AM – 7AM.
- Hemoglobin A1C
  - Lab_procedure_ids: 1093, 1665, 2404, 2405, 2406
  - The majority of results are reported as %
  - Some results are reported as g/dL

Addressing “missing” data: labs

- Not all patients have complete “admission labs”
  - Clinicians don’t consider all of them relevant in “real world” practice
  - For elective surgeries, pre-admission labs might be drawn outside the system, but if abnormal, they are highly likely to be repeated
  - In a recent infectious disease study ~10% of study patients did not have an admission hemoglobin and ~10% did not have a glucose
    - For otherwise healthy patients, neither one would be required to guide treatment decisions – indeed those without the labs had better outcomes
- A number of tests are only ordered when specifically indicated, so even an order is significant
  - Serum lactate
    - The presence of data rows implies physician concern over organ dysfunction
  - Consult with the study team regarding how to address missing lab values

Addressing “missing” data: labs (cont’d.)

- Most, but not all contributing hospitals capture fingerstick glucoses (POC) in the lab system
  - If assessing glycemic control, include mean/median number of glucose readings during encounter by facility
  - If there’s a need to include all sites, address statistically
- Some tests such as molecular diagnostics may not be captured if not mapped or if performed outside facility
  - Address up front with a query

Using lab data to define groups/subgroups

- Ideal for abnormalities defined by lab values
  - Anemia
  - Hepatic or renal dysfunction or toxicity
  - Hyperglycemia
  - Toxic drug levels
- Ideal for conditions known to be underdiagnosed
  - Renal insufficiency (eGFR)
  - Hypoglycemic events
- Can define based on
  - Pre-defined value or change in value
- Can leverage lab tests to identify drug exposure
  - Prothrombin/INR (with/without vitamin K) for warfarin-related bleeds
  - Drug levels for toxicity or overdose

Timing of lab specimen collection

- Presentation/Admission
  - Toxicity, complications
  - Baseline or admission labs (Cerner uses first value within 48 hours from admission/arrival)
- Nearest to index event (e.g. procedure/treatment, drug, first culture, mechanical ventilation)
- Any time during encounter
  - Examples: toxicity, organ dysfunction
  - Pre-defined thresholds
    - Worst value (e.g. peak or nadir)
    - First abnormal result (as defined)

Using micro data to define groups/subgroups

- Timing of specimen collection can help distinguish HAI
  - First positive C. difficile toxin >48 hours after admission
  - Respiratory cultures collected >48 hours after mechanical ventilation with either primary or non-primary dx of VAP
  - Respiratory culture collected within 24-48 hours of admission for patients admitted with pneumonia and no evidence of HAI (thus CAP)

- By specific isolates or number of isolates
  - At least one gram negative isolate
  - One or more specific pathogens of interest (MRSA, VRE, pseudomonas)

Health Facts®– hf_f_medication

- *Health Facts*® has medication orders that go through the hospital pharmacy
  - Meds given from floor stock or from a crash cart might or might not appear in the pharmacy order data.
  - Specialty areas such as surgery and radiology usually manage their own medications. Therefore, orders for these agents do not typically go through the pharmacy.

Health Facts®– hf_f_medication

- *Health Facts*® has the medication order and overall dispense information (not med administration)
  - Ceftriaxone 2 gm IV q24hrs started 1/12/2005 12:00 stopped 1/19/2005 11:59, 7 doses charged, 2 doses credited
  - Dextrose 50% 25 mL IV prn started 1/12/2005 12:00 stopped 1/19/2005 11:59, 0 doses charged
  - Dextrose 50% 25 mL IV prn started 3/23/2005 12:00 stopped 3/29/2005 23:59, 1 dose charged
  - Lantus 40 U SQ every evening started 4/5/2008 19:00 stopped 4/15/2008, 1x10mL vial charged, 0 credited.
  - Insulin lispro, as directed, started 8/23/2009 18:00 stopped 8/29/2009 17:45, 3x3mL pens charged, 0 credited

Health Facts® – Medications

- Obtain medication_ids of interest

Select *

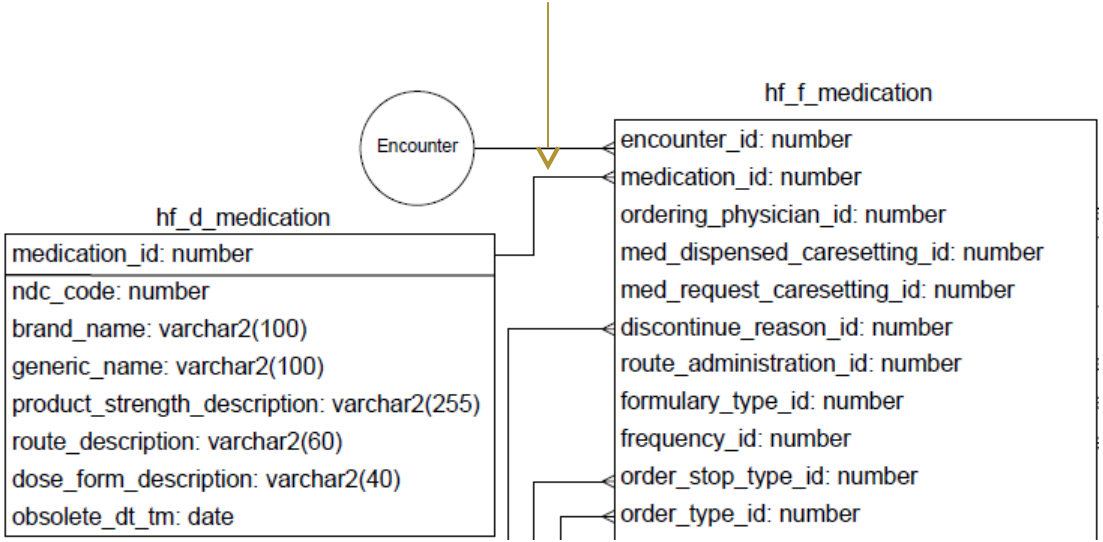
From hf_d_medication

Where lower(brand_name) like ('%lantus%')

MEDICATION_ID	NDC_CODE	BRAND_NAME	GENERIC_NAME
115940	88222033	Lantus	insulin glargine
1933262	2245689	Lantus	insulin glargine
2485283	88222052	Lantus	insulin glargine
2472463	68115000000	Lantus	insulin glargine
2729678	54868000000	Lantus	insulin glargine
3594872	54570000000	Lantus	insulin glargine
6382730	49999000000	Lantus	insulin glargine
7195962	55045000000	Lantus	insulin glargine
7194776	54869000000	Lantus OptiClik Cartridge	insulin glargine
5236308	88222060	Lantus Solostar Pen	insulin glargine
8064453	88221905	Lantus Solostar Pen	insulin glargine

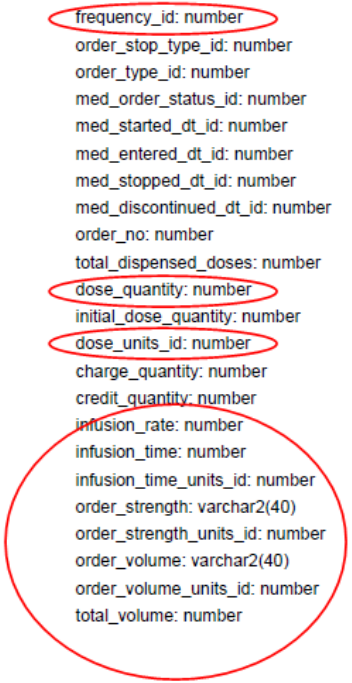
Health Facts® – Medications

Then join medication_id from hf_d_medication
to medication_id in hf_f_medication



Medications – Dosing

- Dosing is represented more than one way
- Determining daily dose requires examining the data and writing code to handle the various scenarios. Fields that might be used



Medications – Orders versus Dispensed Orders

- To differentiate a dispensed order from any order (e.g., standing orders where the patient may never have received the drug):
  - Make sure the order wasn't cancelled; AND
  - Make sure that either the charge quantity minus the credit quantity is greater than zero OR the value for total dispensed doses is greater than zero.
  - SQL coding logic:

```
where med_order_status_id != 2

    and (charge_quantity – credit_quantity > 0 or total_dispensed_doses > 0)
```

- Prior to 2009, extract architecture generally limited the orders in *Health Facts*® to those with dispensing events.
- In 2009, facilities began migrating to a new architecture that allows for the collection of all medication orders, though not all facilities have adopted it.

Health Facts®– hf_f_clinical_event

- Vital Signs, Height, Weight
  - Height, Weight, Blood Pressure, Heart Rate, Pulse, Temperature, BSA, BMI
  - The dates and times for the event start and end, if applicable
  - The result value captured in several possible forms such as number, date, etc… depending on the type of event,
  - Unit of measurement and low/high/critical value thresholds when applicable
  - The medical specialty for the performed personnel per event if available
- Patient Symptoms
  - Pain assessment
- Health History information
  - Previous hospitalizations, Smoking cessation

Discussion and Coding Examples

In this section we will…

- …discuss the typical steps for an inpatient study
- …discuss study considerations and additional topics
- …engage in Q&A
- …describe some SQL coding examples that can optionally be explored

Typical Steps for an Inpatient Study

- Apply inclusion/exclusion criteria to obtain a list of encounter_ids of interest
  - Typically, limit to encounters having diagnosis, lab, and medication data (join to these tables and only keep encounter_ids that are in all three tables)
- Obtain all encounters for these patients within a timeframe of interest (e.g., 1 year prior and 30 days post)
- Examine date-times of orders relative to admission and discharge, cleansing the admit/discharge dates if needed.
- Blend adjacent encounters together, if appropriate given the study objectives.

Assessing Feasibility – Key Questions/Issues

- What are your primary study objectives/hypotheses?
  - Exploratory hypotheses?
- What are the minimal data elements required to meet objectives?
  - Define study population and key endpoints of interest
  - Minimal data elements are required to match/adjust comparators
- What other data elements would provide additional insight?
  - Is it acceptable if only available for a portion of the study population?
- Minimum sample size needed
  - Consider power calculations
- Perform a query to test feasibility
  - Use specifications that are appropriate for an initial query

Selecting your minimal data elements

- Think real world data, not clinical trials
- Diagnosis codes: know strengths and limitations
  - ACS codes changed in 2005 to distinguish NSTEMI and STEMI
  - Cancer codes organized by anatomic location, not cell type
  - Procedure codes are from a facility perspective
- Pharmacy: ordered, or ordered and dispensed?
- Labs: which specific labs should be included and is more than one result required?
- Micro: any culture result, or any culture ordered?
  - Use all cultures to define and refine cultures of interest
- Clinical assessments (events):
  - Which specific measures are applicable?
  - Can some be combined via a cross-walk or mapping process?

Feasibility testing may require additional analysis

- Additional data exploration may be needed to test feasibility
  - Insight into size of specific patient groups or cohorts
    - Timing of drug exposure, specific qualifying cultures, etc.
  - Insight into inclusion and exclusion criteria
- Can do power calculations after initial data exploration
- Definitions of patient groups may be negotiable and can be tested
- When in doubt, **define more broadly up front**

Additional Topics – Patient Location

- Patient location can be inferred from
  - Discharge caresetting in the hf_f_encounter table
  - “caresetting” ids and date-time stamp on orders
    - Caresetting is the location where there order was placed/requested/dispensed, which

Column	Table	Date/Time
Lab_order_caresetting_id	Hf_f_lab_procedure	Lab_drawn_dt_tm
Med_request_caresetting_id	Hf_f_medication	Med_started_dt_tm
Med_dispensed_caresetting_id	Hf_f_medication	Med_started_dt_tm

Additional Topics – Patient Location/ICU

- If the goal is to determine whether or not the patient spent time in the ICU, look for the following caresettings:

CARESETTING_ID	CARESETTING_DESC
18	Coronary Care Unit
55	Intensive Care Unit
56	Intensive Care Unit - Cardiac
57	Intensive Care Unit - Medical
58	Intensive Care Unit - Neonatal
59	Intensive Care Unit - Neurology
60	Intensive Care Unit - Surgical

Additional Topics – Patient Location/ICU

- Find the first and last order from an ICU setting, occurring X hours apart (we have used 12 hours, but one could select shorter or longer timeframes, depending on the patient population)
- For that timeframe get the
  - % of orders with a valued caresetting and
  - % of orders with a valued caresetting coming from an ICU. The higher both of these percentages, the more likely it is that the patient was in the ICU during that timeframe.
- Also examine the discharge caresetting on the encounter table to see if it is an ICU location.
- Limitations:
  - Caresetting is not valued on every order (some hospitals are better than others)
  - Caresetting might not correspond to the patient's location if the provider placed the order from some other location.

One potential classification of caresetting

```
case when caresetting_id in (16, 84, 175, 176, 173, 174, 177) then -1  /* invalid/unknown locations  */
     when caresetting_id in (11, 12, 15, 22, 50, 63, 64, 65, 66, 67, 68, 69, 100, 116, 117, 119, 121, 124, 127,
                             128, 130, 131, 136, 137, 166, 25) then 0 /* procedure locations or ER */
     when caresetting_id in (18, 55, 56, 57, 58, 59, 60) then 1 /* ICU locations  */
     when caresetting_id in (140, 141, 142, 143) then 2 /* SDU locations */
     when caresetting_id in (2, 4, 5, 6, 7, 9, 17, 21, 23, 24, 26, 29, 30,31, 32, 34, 37, 41, 43, 45, 47, 48, 49, 62,
                             74, 78, 82, 90, 91, 92, 102, 104, 105, 108, 110, 112, 115, 120, 123, 129, 134, 138,
                             154, 156, 158, 161, 168, 170, 171, 172, 178) then 3 /* outpatient locations */
     when caresetting_id in (1, 3, 8, 10, 13, 14, 19, 20, 27, 28, 33, 35, 36, 38, 39, 40, 42, 44, 46, 51, 52, 53,
54,
                              61, 70, 71, 72, 73, 75, 76, 77, 79, 80, 81, 83, 85, 87, 88, 89, 93, 94, 95, 96, 97,
98,
```

Additional Topics - Total Charges

- *Health Facts*® has total charges, not costs
- For recent *Health Facts*® projects:
  - Cost-to-charge ratios obtained from HCUP data
  - Matched hospital demographics from HCUP data to those from *Health Facts*® data, assigning cost-to-charge ratio based on census region, bed size, urban vs rural, and teaching status for Medicare and non-Medicare patients

Discussion / Q&A

- Discussion of topics of interest / Q&A

Code examples

- A number of SQL code examples are available for (optional) independent study; please ask if interested