****

**OMOP Common Data Model (CDM V5.0)**

**ETL Mapping Specification Premier**

**Version 1.0**

**June 2015**

Rupa Makadia

Table of Contents

[1.0 Introduction 4](#_Toc427766907)

[2.0 Source Data Mapping Approach 5](#_Toc427766908)

[3.0 Source Data Mapping 8](#_Toc427766909)

[3.1.1 Table Name: CDM\_SOURCE 9](#_Toc427766910)

[3.1.2 Table Name: PERSON 10](#_Toc427766911)

[3.1.3 Table Name: VISIT\_OCCURRENCE 13](#_Toc427766912)

[3.1.4 Table Name: VISIT\_COST 19](#_Toc427766913)

[3.1.5 Table Name: OBSERVATION\_PERIOD 20](#_Toc427766914)

[3.1.6 Table Name: PAYER\_PLAN\_PERIOD 22](#_Toc427766915)

[3.1.7 Table Name: DEATH 23](#_Toc427766916)

[3.1.8 Table Name: OBSERVATION 25](#_Toc427766917)

[3.1.9 Table Name: SPECIMEN 27](#_Toc427766918)

[3.1.10 Table Name: MEASUREMENT 28](#_Toc427766919)

[3.1.11 Table Name: DRUG\_EXPOSURE 30](#_Toc427766920)

[3.1.12 Table Name: DRUG\_COST 33](#_Toc427766921)

[3.1.13 Table Name: CONDITION\_OCCURRENCE 35](#_Toc427766922)

[3.1.14 Table Name: PROCEDURE\_OCCURRENCE 38](#_Toc427766923)

[3.1.15 Table Name: PROCEDURE\_COST 42](#_Toc427766924)

[3.1.16 Table Name: DEVICE\_EXPOSURE 43](#_Toc427766925)

[3.1.17 Table Name: DEVICE\_COST 45](#_Toc427766926)

[3.1.18 Table Name: NOTE 47](#_Toc427766927)

[3.1.19 Table Name: PROVIDER 47](#_Toc427766928)

[3.1.20 Table Name: CARE\_SITE 49](#_Toc427766929)

[3.1.21 Table Name: FACT\_RELATIONSHIP 50](#_Toc427766930)

[3.1.22 Table Name: LOCATION 50](#_Toc427766931)

[3.1.23 Table Name: COHORT 51](#_Toc427766933)

[3.2 Source Independent Data Mapping 52](#_Toc427766934)

[3.2.1 Table Name: DRUG\_ERA 52](#_Toc427766935)

[3.2.2 Table Name: CONDITION\_ERA 54](#_Toc427766936)

[3.2.3 table name: DOSE\_ERA 55](#_Toc427766937)

[4.0 Appendix 56](#_Toc427766938)

[4.1 Premier mapping to CDM fields 57](#_Toc427766939)

[4.2 STD\_CHG\_CODE to SOURCE\_TO\_CONCEPT\_MAP table 57](#_Toc427766940)

[4.3 Provider Specialty mapping table 57](#_Toc427766941)

[4.4 STD\_CHG\_CODE to HOSP\_CHG\_ID Mapping table 57](#_Toc427766942)

# Introduction

This document reflects the requirements, assumptions, business rules, and transformations for the implementation of the Common Data Model Version 5.0 (CDM) as implemented by Rupa Makadia, Epidemiology Analytics, Janssen Research and Development.

The purpose of this document is to describe the ETL mapping of the proprietary or licensed data from ***Premier*** into the OMOP Common Data Model. All table references made to the source database, Premier, will be referenced in bold and italic.

Premier is a hospital based system that houses inpatient and outpatient visits from 619 hospitals with over 62 million discharges. Premier is a hospital based system that captures visits and is not a claims database. Each visit is linked to billing records. The data captures 1 in every 5 inpatient stays in the US.

The document is composed of two main sections:

* Source Data Mapping. Describes major tables of the CDM schema and special data handling required for each table.
* Source Independent Data Mapping. Describes mapping process of the drug and condition eras.

In each section, the tables and their mapping are individually reviewed along with any source-specific rules and exceptions.

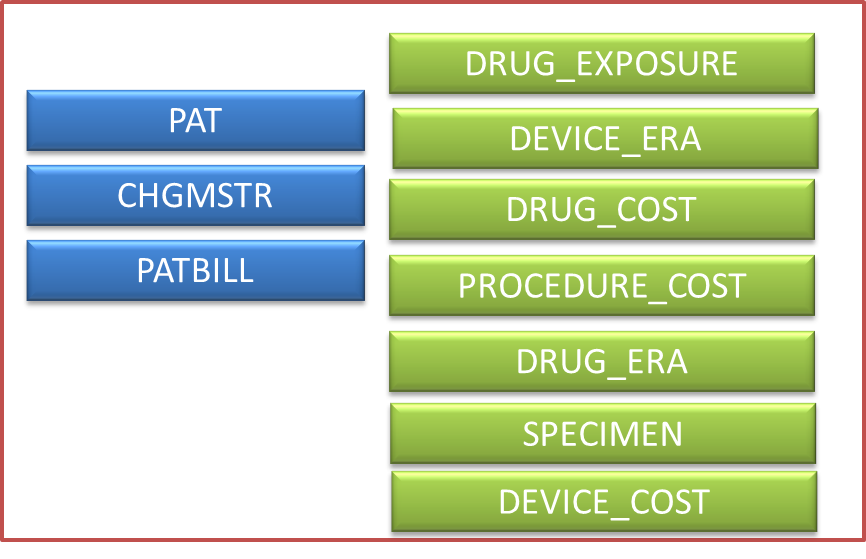
The intended audience for this document is researchers wishing to use the experience and learning in their own CDM construction.

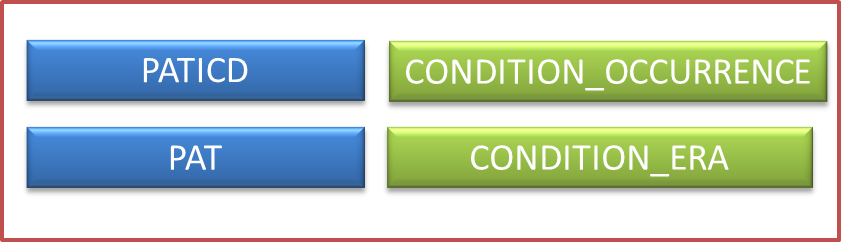
# Source Data Mapping Approach

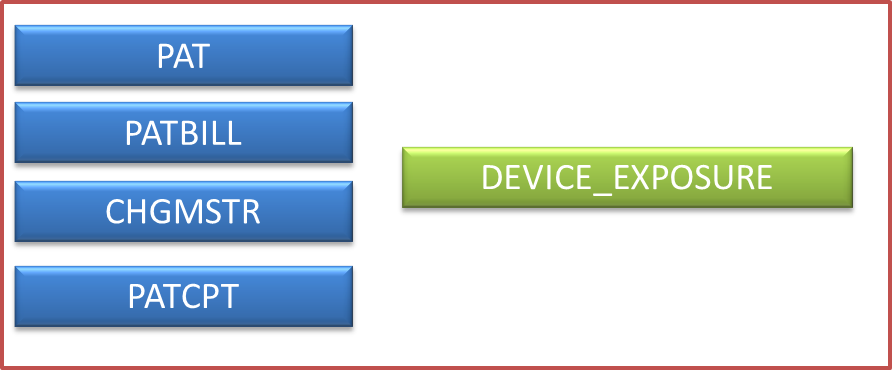
The figure below represents the general approach to mapping the source data tables that comprise Premier to the CDM data schema. Additional ancillary tables and look up tables are included in the chart. The following tables are null: ORGANIZATION and COHORT. The blue boxes represent the input tables in Premier and the green boxes represent the resulting CDM tables. For example, information in the ***PAT*** source table maps to the CDM PERSON table, the VISIT\_OCCURRENCE table, , the OBSERVATION\_PERIOD table, the PAYER\_PLAN\_PERIOD\_TABLE, and the OBSERVATION\_TABLE.

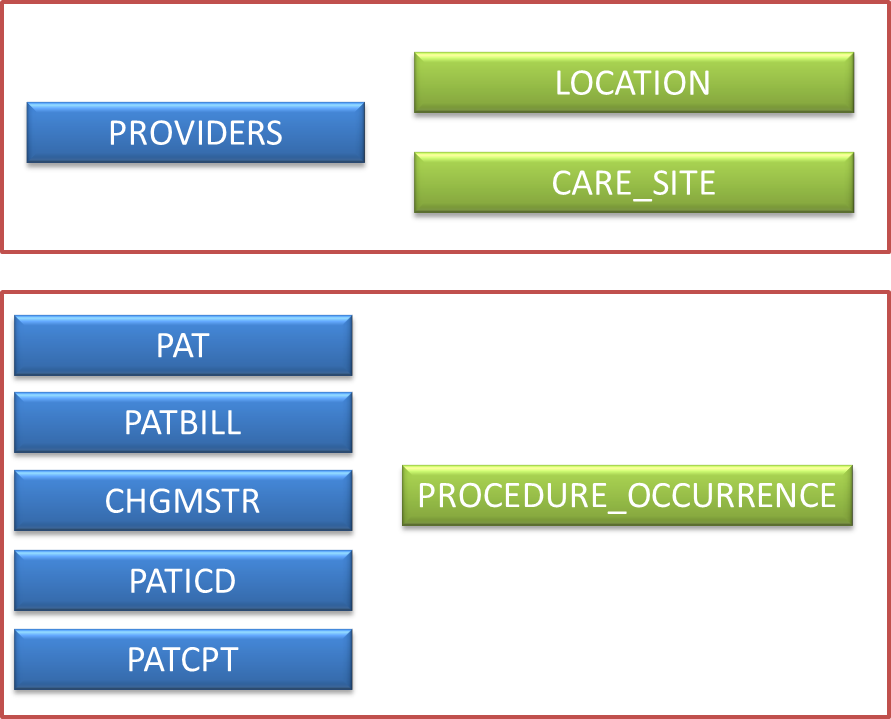
In addition to the high level table mapping that is represented below, a review was conducted of each source data field that is available in all of the following tables: ***PAT, PATBILL, PATICD,*** and ***PATCPT***. The mapping at the field level and justification of why selected fields are not included is located in the appendix.





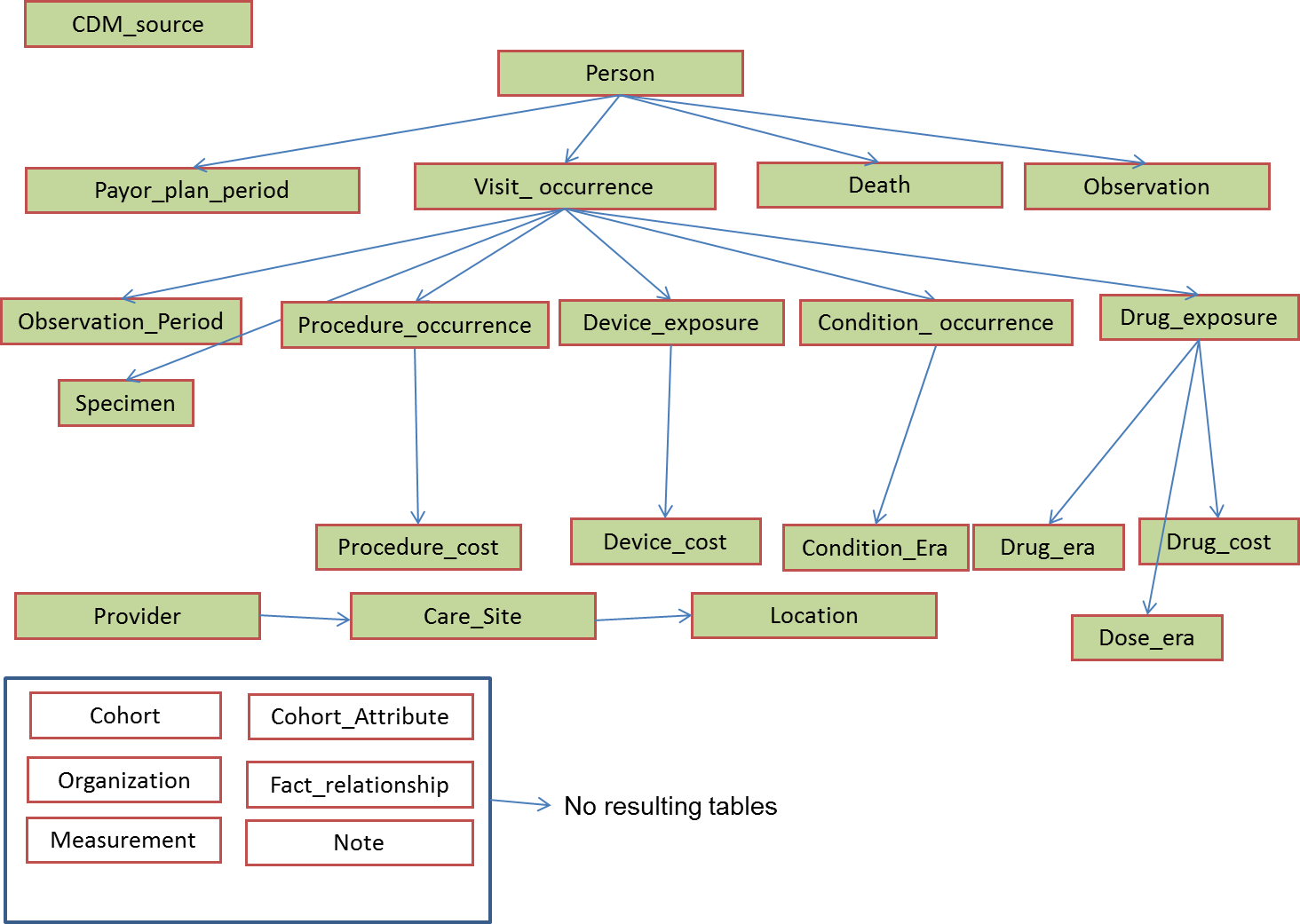






# Source Data Mapping

The figure below illustrates the order in which the CDM tables are generated. The **VISIT\_OCCURRENCE table must generated first** because procedure occurrence, device exposure, condition occurrence, and drug exposure dates are subsequently generated using visit start date. The start and end date of each visit are derived from the maximum number of service days recorded during a visit. The service days for each visit are located in ***PATBILL*** table where, for each visit, the maximum value in this field is obtained. The logic transformation for these dates is explained in the sections for each respective table.



### Table Name: CDM\_SOURCE

*Describe how the CDM\_SOURCE mapping and transformations are designed.*

The CDM\_SOURCE table houses meta data about the version of the CDM that is populated, including key elements such as the vocabulary version used in generating the CDM.

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| CDM\_SOURCE\_NAME |  | Premier |  |
| CDM\_SOURCE\_ABBRIVIATION |  | Premier |  |
| CDM\_HOLDER |  | Janssen R&D |  |
| SOURCE\_DESCRIPTION |  | Anonymized hospital transactional database from over 500 hospitals from 2000-present day includes inpatient, outpatient, and emergency room visits.  The database is a visit-oriented database, with each visit having its own unique id.  Conditions are coded as ICD9 codes and procedures are coded both in ICD9, CPT, and HCPCS procedure codes. Drugs, labs, and other procedures are coded as a standard charge code and occur as a transaction.  Cost information is associated to each transaction including charges and quantity of each transaction is recorded in the billing table. |  |
| SOURCE\_DOCUMENTATION\_REFERENCE |  | http://hicoe.jnj.com/DataSources/Premier |  |
| CDM\_ETL\_REFERENCE |  | http://www.ohdsi.org/web/wiki/  doku.php?id=documentation:example\_etls |  |
| SOURCE\_RELEASE\_DATE |  | SELECT VERSION\_DATE  FROM [\_Version] | Get from the raw source tables. |
| CDM\_RELEASE\_DATE |  | SELECT CONVERT(VARCHAR(10), GETDATE(),102) | Get the date the run completes. |
| CDM\_VERSION |  | V5.0 |  |
| VOCABULARY\_VERSION |  | SELECT VOCABULARY\_VERSION  FROM VOCABULARY  WHERE VOCABULARY\_ID = 'None' | Taken from the Vocabulary loaded into the CDM. |

### Table Name: PERSON

*Describe how the Person mapping and transformations are designed.*

In Premier, the ***PAT*** table contains all demographic, admission, and total cost data for each visit. There are multiple entries per person, thus a single record needs to be created to populate the PERSON table. MONTH\_OF\_BIRTH and DAY\_OF\_BIRTH are not available in Premier, because age is the only available field. YEAR\_OF\_BIRTH is calculated from the first transformed admission date. The admission year minus the age results in the YEAR\_OF\_BIRTH. Since no address information is available in Premier for each person, LOCATION\_ID is null. Primary care providers for each person are not known, thus PROVIDER\_ID and CARE\_SITE are NULL. Race can vary among records for the same person in the ***PAT*** table, so the most common race value is used for these people. Ethnicity is available in the race field so logic is applied to parse out the ethnicity from each record. Hispanic is the only ethnicity available in Premier so for those with ethnicity recorded as Hispanic, their race is considered ‘Other’. For populating the ETHNICITY field, if the race is Hispanic then ETHNICITY is assigned Hispanic otherwise the ethnicity is coded as NULL.

Delete any patients that have invalid birth years < 1900 or > the current year. After birth year has been determined delete any individual that has an OBSERVATION\_PERIOD that is >= 2 years prior to the YEAR\_OF\_BIRTH. Due to data discrepancies in Premier additional logic has been applied to generating gender and age. If a person has YEAR\_OF\_BIRTH that varies over two years then that person is dropped. Also, if a person has multiple genders recorded then those records are dropped. The exclusion criterion for the PERSON table removes about 1% of the population in the database.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** | Field is a randomly generated identifier that is available in Premier |  |
| GENDER\_CONCEPT\_ID | ***PAT.GENDER*** | When ***PAT.GENDER***=M then GENDER\_CONCEPT\_ID=8507  When ***PAT.GENDER***=F then GENDER\_CONCEPT\_ID=8532  Unknown gender mapped to zero | CONCEPT\_ID’s are VOCABULARY\_ID=Gender |
| YEAR\_OF\_BIRTH | ***VISIT\_OCCURRENCE.VISIT\_START\_DATEPAT.AGE*** | Year(***VISIT\_OCCURRENCE.VISIT\_START\_DATE***)-age  Year\_of\_birth needs to be > 1900 and <=current year |  |
| MONTH\_OF\_BIRTH | ***-*** | NULL | Premier only provides age |
| DAY\_OF\_BIRTH | ***-*** | NULL | Premier only provides age |
| TIME\_OF\_BIRTH | ***-*** | NULL | Premier only provides age |
| RACE\_CONCEPT\_ID | ***PAT.RACE*** | When ***PAT.RACE***=1 then RACE\_CONCEPT\_ID=8527  When ***PAT.RACE***=2 then RACE\_CONCEPT\_ID =8516  Race value of 9 and U gets mapped to 0 | Premier combines both race and ethnicity into one field. Ethnicity is removed from race. If multiple race records per person, see logic to obtain the max value of race that occurs in all records. |
| ETHNICITY\_CONCEPT\_ID | ***PAT.RACE*** | When ***PAT.RACE***=3 then  ETHNICITY\_CONCEPT\_ID=38003563 else NULL | If race is not Hispanic set ethnicity to 0 |
| LOCATION\_ID | - | NULL |  |
| PROVIDER\_ID | - | NULL |  |
| CARE\_SITE\_ID | - | NULL |  |
| PERSON\_SOURCE\_VALUE | ***PAT.MEDREC\_KEY*** |  |  |
| GENDER\_SOURCE\_VALUE | ***PAT.GENDER*** |  |  |
| GENDER\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |
| RACE\_SOURCE\_VALUE | ***PAT.RACE*** |  |  |
| RACE\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |
| ETHNICITY\_SOURCE\_VALUE | ***PAT.RACE*** |  |  |
| ETHNICITY\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |

### Table Name: VISIT\_OCCURRENCE

*Describe how the Visit\_Occurrence mapping and transformation are designed.*

Premier data are visit oriented, thus each visit has its own unique visit identifier. The ***PAT*** table includes admission date and discharge date for each visit. Each visit is stored as a date but the day of the stay is always coded as the first of the month. Because a person can have more than one visit in the same year-month combination, an additional field is included to preserve the order of visits because the day of month is lost. The field ***PAT.DISC\_MON\_SEQ*** is included in Premier to preserve the order of visits based on unique discharge dates. The ***LOS*** field on the ***PAT*** table is populated for inpatient stays and is recorded as the number of 24 hour increments that a patient spends in the hospital. The ***LOS*** field is usually off by one calendar day for inpatient visits. Outpatient visits have a ***LOS*** of 0 and are typically only one day in length which is represented in the billing tables. For about 6% of outpatient visits, a single visit is comprised of multiple procedures that occur on different days for procedures such as chemotherapy or dialysis. The Premier billing system does not separate these into individual visits and for purposes of the ETL these will be considered one continuous visit. The ***PATBILL*** table houses billing records that occur each calendar day during the visit. Some visits may include a service day record of zero, which are considered to be pre-visit tests or procedures. For the purposes of the ETL, the information is considered as the first day of the visit, which can occur for inpatient and outpatient stays. The length of the stay is determined by the ***PAT\_BILL*** table using the field ***SERV\_DAY*.** The max service day is determined for each person. Logic is defined to take each person, visit, and sequence order and add the maximum number of service days to their stay starting with the first of the month. Each visit that occurs after the first will use the previous end date +1 as the visit start date, and then the service days are added to the start date to determine to end date. The same logic is repeated until all observations have a valid start date and end date that do not overlap. For those visits that have an admission date and discharge date in different months, the logic works backward to obtain the visit start date from the max service days and visit end date. If there are multiple visits that occur in the same admission month that span over months, the records may overlap, which affects roughly 0.10% of patients. For those records that have service days greater than the number of days in the month, the remaining days are truncated to the last day of the month. Logic guarantees that length of visits is preserved as well as the sequential order of visits. The specific day of the month in start date or end date is not necessary accurate, and the time between to two visits is not necessary accurate. The algorithm will always set the new visit start date plus 1 day or use the backward logic and assume that the visit end date is the first day of the new month.

An example for both sets of logic is displayed below:



Logic for ER stays is varied due to data changes from Premier in identifying ER visits. ER Visits in Premier are identified through point of origin or admission source and those visits are identified as an outpatient stay if a person enters the ER and then leaves. If an inpatient stay results from an ER visit, that is identified as an Inpatient stay. Additional logic has been added to constrain dates and fields which reflect changes in Premier’s classification of emergency room visits. In Premier, the value of admission source that designates emergency room was discontinued 7/1/2010 because it no longer was a required variable for CMS. Point of origin represents the last physical location of a patient before entering the hospital. This field is populated in Premier after 7/1/2010 with emergency room visits. Thus, a combination of point of origin, admission source, and admission type is used to determine if a patient had a valid ER stay for Premier. ***LOS***, the length of stay field, is not used because each billing record corresponds to a service day in the ***PATBILL*** table.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** |  |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| VISIT\_CONCEPT\_ID |  | When ***PAT.ADM\_DATE*** <= ‘6/1/2010’ and ***ADM\_SOURCE***=7 and ***I\_O\_IND*** =’O’ then concept\_id=9203  When ***PAT.ADM\_DATE*** >= ‘7/1/2010’ and (***POINT\_OF\_ORIGIN***=7 or ***ADM\_SOURCE*** =1) and ***I\_O\_IND*** =’O’ then concept\_id=9203  When ***I\_O\_IND*** =’I’ then concept\_id=9201  When ***I\_O\_IND*** =’O’ then concept\_id=9202 | Logic is in the following order:  Patient enters through the ER but doesn’t result in an inpatient stay will be identified as an ER visit.  Patients who may have entered through the ER but get admitted will be identified as an Inpatient visit. |
| VISIT\_START\_DATE | ***PAT.ADM\_DATE***  ***PATBILL.SERV\_DAY***  ***PAT.DISC\_MON\_SEQ*** | See logic in code and explanation above |  |
| VISIT\_START\_TIME | ***-*** | NULL |  |
| VISIT\_END\_DATE | ***PAT.DISC\_DATE***  ***PATBILL.SERV\_DAY***  ***PAT.DISC\_MON\_SEQ*** | See logic in code and explanation above |  |
| VISIT\_END\_TIME | ***-*** | NULL |  |
| VISIT\_TYPE\_CONCEPT\_ID |  | 44818517- Visit derived from encounter on claim |  |
| PROVIDER\_ID | ***PAT.ADM\_PHY*** |  |  |
| CARE\_SITE\_ID | ***PAT.PROV\_ID*** |  |  |
| VISIT\_SOURCE\_VALUE | ***PAT.I\_O\_IND*** |  |  |
| VISIT\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |

**SAMPLE CODE: T-SQL logic**

--sort table by adm\_date and dis\_mon\_seq, make sure to pick up max\_serv\_day

DECLARE visit\_cursor CURSOR FOR

select pat.pat\_key as id,

pat.medrec\_key as person\_id,

pat.adm\_date as StartDate,

pat.disc\_date as EndDate,

pat.disc\_mon\_seq,

case when max(patbill.serv\_day)=0 then 1 else max(patbill.serv\_day)

end as max\_serv\_day

from pat

inner join patbill

on pat.pat\_key = patbill.pat\_key

where pat.medrec\_key in {0}

group by pat.patkey, pat.medrec\_key, pat.adm\_date, pat.disc\_date, pat.dis\_mon\_seq

order by pat.disc\_date asc, pat.dis\_mon\_seq asc

;

--create temp variables for cursoring

DECLARE @lastenddate date

DECLARE @laststartdate date

DECLARE @lastendmonth date

DECLARE @currentid int

DECLARE @oldpersonid int

DECLARE @currentpersonid int

DECLARE @currentstartdate date

DECLARE @currentenddate date

DECLARE @currentdismonseq int

DECLARE @currentmaxservday int

SET @laststartdate = '1/1/1900'

SET @lastenddate = '1/1/1900'

SET @lastendmonth = '1/1/1900'

SET @oldpersonid=0

OPEN visit\_cursor

FETCH NEXT FROM visit\_cursor INTO @currentid, @currentpersonid, @currentstartdate, @currentenddate, @currentdismonseq, @currentmaxservday

WHILE @@FETCH\_STATUS = 0

BEGIN

IF @oldpersonid <> @currentpersonid

BEGIN

SET @laststartdate = '1/1/1900'

SET @lastenddate = '1/1/1900'

SET @lastendmonth = '1/1/1900'

set @oldpersonid=@currentpersonid

END

IF @currentstartdate = @currentenddate

BEGIN

IF @lastendmonth < @currentstartdate --should only fire on first record or when dis\_mon\_seq = 1

BEGIN

SET @currentenddate = case when dateadd(dd,@currentmaxservday,@currentstartdate) < dateadd(mm,1,@currentstartdate)

then dateadd(dd,@currentmaxservday-1,@currentstartdate) --if endate falls in month, use it

else dateadd(dd,-1, dateadd(mm,1,@currentstartdate)) --otherwise, set to last day of the month

end

SET @laststartdate = @currentstartdate

SET @lastendmonth = @currentstartdate

SET @lastenddate = @currentenddate

END

ELSE --IF @lastendmonth = @currentstartdate

BEGIN

SET @currentstartdate = dateadd(dd, case when @lastenddate < dateadd(dd,-1,dateadd(mm,1,@lastendmonth)) then 1 else 0 end , @lastenddate) --increment prior visit by 1 unless you're already at the end of the month

SET @currentenddate = case when dateadd(dd,@currentmaxservday,@lastenddate) < dateadd(dd,-1,dateadd(mm,1,@lastendmonth))

then dateadd(dd,@currentmaxservday,@lastenddate)

else dateadd(dd,-1,dateadd(mm,1,@lastendmonth)) end

--use prior visit + 1 and add the maxsrvdate, unless either date exceeds the end of the month, or else just use the end of te month

SET @laststartdate = @currentstartdate

SET @lastenddate = @currentenddate

END

END

ELSE --startdate <> enddate, which means the visit spans across months

BEGIN

IF @lastendmonth = @currentstartdate

BEGIN

SET @currentstartdate = case when datediff(dd, dateadd(dd, case when @lastenddate < dateadd(dd,-1,dateadd(mm,1,@lastendmonth)) then 1 else 0 end , @lastenddate) ,@currentenddate) > @currentmaxservday

then dateadd(dd, -1\*(@currentmaxservday-1), @currentenddate)

else dateadd(dd, 1, @lastenddate)

end

SET @lastendmonth = @currentenddate

SET @currentenddate = case when dateadd(dd,@currentmaxservday-1,@currentstartdate) < DATEADD(mm,1,@currentenddate)

then dateadd(dd,@currentmaxservday-1,@currentstartdate)

else dateadd(dd,-1,DATEADD(mm,1,@currentenddate))

end

SET @lastenddate = @currentenddate

END

ELSE --@lastendmonth < @currentstartdate --if the visit spanning months is the first visit in the month, start it on first day of adm month and let it go through into the disc month

BEGIN

SET @lastendmonth = @currentenddate

SET @currentstartdate =DATEADD(dd,-1\*(@currentmaxservday-1), @currentenddate)

SET @lastenddate = @currentenddate

END

END

INSERT INTO scratch.dbo.VISIT\_OCCURRENCE (VISIT\_OCCURRENCE\_ID, PERSON\_ID, VISIT\_START\_DATE, VISIT\_END\_DATE)

VALUES (@currentid, @currentpersonid, @currentstartdate, @currentenddate)

FETCH NEXT FROM visit\_cursor INTO @currentid, @currentpersonid, @currentstartdate, @currentenddate, @currentdismonseq, @currentmaxservday

END

CLOSE visit\_cursor

DEALLOCATE visit\_cursor

Adding mapping logic for PLACE\_OF\_SERVICE and PROVIDER

select

visit\_occurance\_id,

person\_id,

visit\_start\_date,

visit\_end\_Date,

case when adm\_date <= '06/01/2010' and pat.adm\_source='7' and pat.i\_o\_ind='O' then **9203**

when adm\_date >= '07/01/2010' and (pat.point\_of\_origin='7' OR pat.adm\_source='1') and pat.i\_o\_ind='O' then **9203**

when pat.i\_o\_ind='I' then **9201**

when pat.i\_o\_ind='O' then **9202**

end as place\_of\_service\_concept\_id,

PAT.PROV\_ID AS care\_Site\_id,

PAT.i\_o\_ind as place\_of\_service\_source\_value

from visit\_occurrence

join pat on visit\_occurrence.visit\_occurrence\_id=pat.pat\_key

### Table Name: VISIT\_COST

*Describe how the VISIT\_COST mapping and transformation are designed.*

Premier provides the total cost that is associated with each visit. The VISIT\_COST table will house the total charge amount of the visit which includes costs from drugs, procedures, and devices including unmapped charges that are incurred during the stay. A patient can only have a single cost record for each visit.

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| VISIT\_COST\_ID | ***-*** | System generated |  |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** | - |  |
| CURRENCY\_CONCEPT\_ID | ***-*** | 44818668- American dollar ***-*** | USD |
| PAID\_COPAY | ***-*** | - |  |
| PAID\_COINSURANCE | ***-*** | - |  |
| PAID\_TOWARD\_DEDUCTIABLE | ***-*** | - |  |
| PAID\_BY\_PAYER | ***-*** | - |  |
| PAID\_BY\_COORDINATION\_OF\_BENEFITS | ***-*** | - |  |
| TOTAL\_OUT\_OF\_POCKET | ***-*** | - |  |
| TOTAL\_PAID | ***PAT.PAT\_CHARGES*** | - | The total charge amount of the visit |
| PAYER\_PLAN\_PERIOD\_ID | ***PAYOR\_PLAN\_PERIOD.PAYER\_PLAN\_PERIOD\_ID*** | Use PAYER\_PLAN\_PERIOD table and join by person. Include filtering to assure that the visit start date falls within the payer start date and payer end date |  |

### Table Name: OBSERVATION\_PERIOD

*Describe how the Observation\_Period mapping and transformation are designed.*

Because of the lack of enrollment data in Premier, the observation period for each patient will be defined by unique visits from the VISIT\_OCCURRENCE table for each unique patient. Derived admission and discharge dates are created using the number of service days and the sequence of visits as defined by the field ***PAT.DISC\_MON\_SEQ*** if the visits occurred twice in the same discharge month/year. (See VISIT\_OCCURRENCE specification).

**All overlapping visits will be collapsed into one observation period. An overlapping visit is defined by a visit that has an admit date that is within 31 days of the previous discharge date.** For example a patient has an admission date of ‘2011-02-01’ and a discharge date of ‘2011-02-05’ and the next visit occurs in ‘2011-02-19’ and a discharge date of ‘2011-03-01’. These records would be collapsed into a single observation period.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| OBSERVATION\_PERIOD\_ID | ***-*** | System generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| OBSERVATION\_PERIOD\_START\_DATE | ***VISIT\_OCCURRENCE.VISIT\_START\_DATE*** |  |  |
| OBSERVATION\_PERIOD\_END\_DATE | ***VISIT\_OCCURRENCE.VISIT\_END\_DATE*** |  |  |
| PERIOD\_TYPE\_CONCEPT\_ID | ***-*** | 44814725= Period inferred by algorithm |  |

### Table Name: PAYER\_PLAN\_PERIOD

*Describe how the PAYER\_PLAN\_PERIOD mapping and transformation are designed.*

Payer information presented in Premier exists in the ***PAT*** table, in the field ***PAT.STD\_PAYOR,*** which represents standard payer categories. Since information about how long the patient remains with a payer or is enrolled is unavailable, the PAYER\_PLAN\_PERIOD\_START\_DATE and PAYER\_PLAN\_PERIOD\_END\_DATE are the same as the calculated OBSERVATION\_PERIOD for each patient. If a patient changes payer within an observation period, then the payer plan period will be segmented to reflect the change in payers. If multiple payers are attributed to one observation period, use visit start and visit end to determine the payer.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| PAYER\_PLAN\_PERIOD\_ID | - | System- generated value |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| PAYER\_PLAN\_PERIOD\_START\_DATE | ***OBSERVATION\_PERIOD.OBSERVATION\_PERIOD\_START\_DATE*** |  |  |
| PAYER\_PLAN\_PERIOD\_END\_DATE | ***OBSERVATION\_PERIOD.OBSERVATION\_PERIOD\_START\_DATE*** |  |  |
| PAYER\_SOURCE\_VALUE | ***PAYOR.STD\_PAYER\_DESC*** |  |  |
| PLAN\_SOURCE\_VALUE | - | NULL |  |
| FAMILY\_SOURCE\_VALUE | - | NULL |  |

An example is represented below:



### Table Name: DEATH

*Describe how the DEATH mapping and transformation are designed.*

Death is mapped from discharge status and ICD9 codes. The cause of death is not available in Premier. Discharge status from the ***PAT*** table should be used first, and if no codes are found, then ICD9 codes are used. The ICD9 codes are obtained from the vocabulary and identified in the code below. Keep only one record for each patient, if both discharge status and ICD9 codes indicate death, use discharge status first. No records should be populated for that person 32 days after the death date. The ICD9 diagnosis codes that indicate that a patient has died are:

|  |  |
| --- | --- |
| Maternal death affecting fetus or newborn | 761.6 |
| Accidental mechanical suffocation by plastic bag | E913.1 |
| Sudden infant death syndrome | 798.0 |
| Death occurring in less than 24 hours from onset of symptoms, not otherwise explained | 798.2 |
| Unattended death | 798.9 |
| Instantaneous death | 798.1 |
| Sudden death, cause unknown | 798 |
| Legal execution | E978 |

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| DEATH\_DATE | ***VISIT\_OCCURRENCE.VISIT\_END\_DATE*** |  | The exact date of death cannot be determined thus the VISIT\_END date is used. |
| DEATH\_TYPE\_CONCEPT\_ID | ***PAT.DISC\_STATUS***  OR  ***PATICD.ICD\_CODE*** | Logic based on discharge status or ICD9 diagnosis code. If discharge code is present then assign death, otherwise search records for ICD9 codes Discharge status of 20, 40, 41, 42 indicates death. |  |
| CAUSE\_CONCEPT\_ID | - | NULL |  |
| CAUSE\_SOURCE\_VALUE | - | NULL |  |
| CAUSE\_SOURCE\_CONCEPT\_ID | - | NULL |  |

### Table Name: OBSERVATION

*Describe how the Observation mapping and transformation are designed.*

The observation table houses additional demographic and visit data that is housed in Premier. DRG codes and certain ICD9, CPT, and standard charge codes map to standardized observation table concept. Marital status, admission information, discharge status, and patient type records are specific to Premier and map to non-standard observation table concepts.

***PAT.DRG, PATICD.ICD\_CODE, PATCPT.CPT\_CODE***, and ***PATBILL.STD\_CHG\_CODE*** map to OBSERVATION.OBSERVATION\_CONCEPT\_ID using the source to standard cte\_vocab\_map. These records also map to OBSERVATION.OBSERVATION\_SOURCE\_CONCEPT\_ID using the source to source cte\_vocab\_map.

Note: DRG codes contain 3 digits, including those codes with leading zeros (i.e. 00# and 0##). These leading zeros are missing in Premier data and must be added in the ETL process for accurate mapping using the source to standard cte\_vocab\_map.

***PAT.MART\_STATUS, PAT.POINT\_OF\_ORIGIN, PAT.DISC\_STATUS***, and ***PAT.PATTYPE*** map to set OBSERVATION\_CONCEPT\_ID codes described in the table below and OBSERVATION\_SOURCE\_CONCEPT\_ID=0.

The observation start date is assigned the VISIT\_START\_DATE. The ASSOCIATED\_PROVIDER\_ID that is provided is the randomly generated key provided by Premier for the provider that admitted the patient. There are two providers that exist in Premier, the admitting and attending. This ETL makes the decision to use admitting because it is unknown whether or not the admitting provider, attending provider or another person diagnosed the person.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| OBSERVATION\_ID | - | System-generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| OBSERVATION\_CONCEPT\_ID |  | For PAT.DRG, PATICD.ICD\_CODE, PATCPT.CPT\_CODE, and PATBILL.STD\_CHG\_CODE:  QUERY: SOURCE TO STANDARD  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('HCPCS', 'CPT4', 'ICD9CM', 'DRG')  AND TARGET\_DOMAIN\_ID ='Observation'  For PAT.MART\_STATUS, OBSERVATION\_CONCEPT\_ID=4053609  For PAT.POINT\_OF\_ORIGIN, OBSERVATION\_CONCEPT\_ID=40757183  For PAT.DISC\_STATUS, OBSERVATION\_CONCEPT\_ID= 40757177  For PAT.PATTYPE, OBSERVATION\_CONCEPT\_ID= 40769091 |  |
| OBSERVATION\_DATE | ***VISIT\_OCCURRENCE.VISIT\_START\_DATE*** | NULL |  |
| OBSERVATION\_TIME | ***-*** | NULL |  |
| OBSERVATION\_TYPE\_CONCEPT\_ID | ***-*** | 38000281 Observation recorded from EHR with text result |  |
| VALUE\_AS\_NUMBER | ***-*** | NULL |  |
| VALUE\_AS\_STRING | ***PAT.MART\_STATUS***  ***PAT.POINT\_OF\_ORIGIN***  ***PAT.DISC\_STATUS***  ***PAT.PATTYPE*** | Value\_as\_string only populated for Premier-specific fields mart\_status, point\_of\_origin, disc\_status, and pat\_type  select mart\_status\_desc from martstat m  join pat p on p.mart\_status=m.martstat  select point\_of\_origin\_desc from poorgin po  join pat p on p.mart\_status=po.point\_of\_origin  select disc\_status from poorgin po  join pat p on p.mart\_status=po.point\_of\_origin  select pat\_type\_desc from pattype p  join pat p1 on p1.pat\_type=p.pat\_type | Use look up values in the text fields. |
| VALUE\_AS\_CONCEPT\_ID | ***-*** | NULL |  |
| QUALIFER\_CONCEPT\_ID | ***-*** | NULL |  |
| UNIT\_CONCEPT\_ID | ***-*** | NULL |  |
| PROVIDER\_CONCEPT\_ID | ***PAT.ADMPHY\_SPEC*** |  |  |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** |  |  |
| OBSERVATION\_SOURCE\_VALUE | ***PAT.DRG***  ***PATICD.ICD\_CODE***  ***PATCPT.CPT\_CODE***  ***CHARGE CODE*** | Standard charge code value:  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id |  |
| OBSERVATION\_SOURCE\_CONCEPT\_ID | ***-*** | QUERY: SOURCE TO SOURCE  SELECT SOURCE\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('ICD9CM', 'CPT4', 'HCPCS', ‘DRG’)  AND TARGET\_VOCABULARY\_ID IN ('ICD9CM', 'CPT4', 'HCPCS', ‘DRG’) and TARGET\_DOMAIN\_ID='Observation' |  |
| UNITS\_SOURCE\_VALUE | ***-*** | NULL |  |
| QUALIFIER\_SOURCE\_VALUE | ***-*** | NULL |  |

### Table Name: SPECIMEN

*Describe how the Specimen mapping and transformation are designed.*

Premier does not provide information for specimen’s collected during a person’s stay.

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| SPECIMEN\_ID | ***-*** | NULL |  |
| PERSON\_ID | ***-*** | NULL |  |
| SPECIMEN\_CONCEPT\_ID | ***-*** | NULL |  |
| SPECIMEN\_TYPE\_CONCEPT\_ID | ***-*** | NULL |  |
| SPECIMEN\_DATE | ***-*** | NULL |  |
| SPECIMEN\_TIME | ***-*** | NULL |  |
| QUANTITY | ***-*** | NULL |  |
| UNIT\_CONCEPT\_ID | ***-*** | NULL |  |
| ANATOMIC\_SITE\_CONCEPT\_ID | ***-*** | NULL |  |
| DISEASE\_STATUS\_CONCEPT\_ID | ***-*** | NULL |  |
| SPECIMEN\_SOURCE\_ID | ***-*** | NULL |  |
| SPECIMEN\_SOURCE\_VALUE | ***-*** | NULL |  |
| UNIT\_SOURCE\_VALUE | ***-*** | NULL |  |
| ANATOMIC\_SITE\_SOURCE\_VALUE | ***-*** | NULL |  |
| DISEASE\_STATUS\_SOURCE\_VALUE | ***-*** | NULL |  |

### Table Name: MEASUREMENT

*Describe how the Measurement mapping and transformation are designed.*

The MEASUREMENT table will house records from ***PATBILL*** and ***PATCPT*** that have been mapped to the measurement domain.

Measurements are recorded in the ***PATBILL*** table as standard charges. Premier captures the day the measurement is made in the ***SERV\_DAY*** field thus, the MEASUREMENT\_DATE is determined from the VISIT\_START\_DATE from VISIT\_OCCURRENCE and ***PATBILL.SERV\_DAY*** unless the start date is greater than the end of the month, then it’s truncated to the end of month.For measurements recorded in the ***PATCPT*** table, the day the measurement was made is unknown so MEASUREMENT\_DATE is recorded as VISIT\_END\_DATE.There are two providers that exist in Premier, the admitting and attending. This ETL makes the decision to use admitting because it is unknown whether or not the admitting provider, attending provider or another person diagnosed the person. Only records that fall within an OBSERVATION\_PERIOD are available for each person. The VISIT\_OCCURRENCE table must be created before the MEASUREMENT table is created.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| MEASUREMENT\_ID | - | System generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| MEASUREMENT\_CONCEPT\_ID |  | QUERY: SOURCE TO STANDARD  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('JNJ\_PMR\_PROC\_CHRG\_CD', 'HCPCS', 'CPT4')  AND TARGET\_DOMAIN\_ID ='Measurement' | Only capture those records that have a domain map to Measurement. |
| MEASUREMENT\_DATE | **VISIT\_OCCURRENCE.VISIT\_START\_DATE**  ***PATBILL.SERV\_DAY***  ***Or***  **VISIT\_OCCURRENCE.VISIT\_END\_DATE** |  | If measurement is from ***PATBILL*** use a combination of service day and visit start date unless the service day is greater than the end of the month  If measurement comes from ***PATCPT*** then use visit end date |
| MEASUREMENT\_TIME | ***-*** | NULL |  |
| MEASUREMENT\_TYPE\_CONCEPT\_ID | ***-*** | 44818701-From physical examination |  |
| OPERATOR\_CONCEPT\_ID | ***-*** | NULL |  |
| VALUE\_AS\_NUMBER | ***-*** | NULL |  |
| VALUE\_AS\_CONCEPT\_ID | ***-*** | NULL |  |
| UNIT\_CONCEPT\_ID | ***-*** | NULL |  |
| RANGE\_LOW | ***-*** | NULL |  |
| RANGE\_HIGH | ***-*** | NULL |  |
| PROVIDER\_ID | ***PAT.ADMPHY\_SPEC*** |  |  |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** |  |  |
| MEASUREMENT\_SOURCE\_VALUE |  | SELECT SOURCE\_VALUE FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A  UNION  (  SELECT CPT\_CODE AS SOURCE\_VALUE FROM PATCPT  ) |  |
| MEASUREMENT\_SOURCE\_CONCEPT\_ID | ***-*** | QUERY: SOURCE TO SOURCE  SELECT SOURCE\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('CPT4', 'HCPCS')  AND TARGET\_VOCABULARY\_ID IN ('CPT4', 'HCPCS') AND DOMAIN\_ID=’Measurement’ | Only populated for standard coding CPT4, and HCPCS codes |
| UNIT\_SOURCE\_VALUE | ***-*** | NULL |  |
| VALUE\_SOURCE\_VALUE | ***-*** | NULL |  |

### Table Name: DRUG\_EXPOSURE

*Describe how the Drug\_Exposure mapping and transformation are designed.*

The DRUG\_EXPOSURE table will house records from ***PATBILL***, ***PATCPT,*** and ***PATICD*** that have been mapped to the drug or metadata domain.

Administrations of drugs are recorded in the ***PATBILL*** table as standard charges. Premier captures the day of administration in the ***SERV\_DAY*** field. DRUG\_EXPOSURE\_START\_DATE is determined by adding the number of service days to the visit start day using VISIT\_OCCURRENCE .VISIT\_START\_DATE and ***PATBILL.SERV\_DAY.*** If the start date is greater than the end of the month, then it’s truncated to the end of month. Procedure drugs reside in the ***PATCPT*** table, and ***PATICD*** table. DRUG\_EXPOSURE\_START\_DATE for procedures is the last day of visit or VISIT\_END\_DATE, since dates for the administration of procedure drugs is not recorded, the assumption is made that the procedure occurred sometime before the end of the visit. DRUG\_EXPOSURE\_END\_DATE cannot be determined because the patient is not followed after each stay and days’ supply information is not available.

Premier does not provide NDC codes for drugs that are administered during a visit. The PRESCRIBING\_PROVIDER\_ID is determined from the visit using the admitting provider id of the visit. In Premier the admitting and attending providers are provided and due to the similarity of both the fields, the admitting provider id is used. The determination cannot be made if the admitting provider was the provider that prescribed the medication but that is the only information that is available. Drug type is considered inpatient administration for all drugs, except those drugs that are procedures and come from ***PATCPT*** or ***PATICD***. Both HCPCS codes and CPT codes are available in ***PATCPT***. The quantity of drugs administered as captured from the ***QTY*** field in ***PATBILL***.

Each standard charge from PATBILL maps to the ***CHGMSTR*** table that houses additional information regarded the department, and descriptions about the item. In order to map each drug to an appropriate concept, USAGI was used to map ***STD\_CHG\_DESC*** to the value of an RxNorm concept. The ***CHGMSTR*** table is segmented by the department ***STD\_DEPT\_DESC***. The drug records represented in this table are captured from the Pharmacy department. Any mapping that cannot be correctly identified is mapped to a CONCEPT\_ID of zero. All drugs will be mapped and included in this table even if they don’t have a valid concept. All charges are loaded into the SOURCE\_TO\_CONCEPT\_MAP table, and the table is attached in the appendix. The ***STD\_CHG\_CODE*** is mapped to a ***HOSP\_CHG*** using the table ***HOSPCHG***, and each ***HOSP\_CHG*** has a description that is displayed in the CDM as the source value.

Only records that fall within an OBSERVATION\_PERIOD are available for each person. The VISIT\_OCCURRENCE table must be created before the DRUG\_EXPOSURE table is created.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| DRUG\_EXPOSURE\_ID | - | System generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| DRUG\_CONCEPT\_ID |  | QUERY: SOURCE TO STANDARD  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('JNJ\_PMR\_DRUG\_CHRG\_CD', 'CPT4', 'HCPCS', 'ICD9CM')  AND TARGET\_DOMAIN\_ID in ('Drug', 'Metadata') | Include all concepts that map to a concept id of zero. |
| DRUG\_EXPOSURE\_START\_DATE | ***PATBILL.SERV\_DAY***  **VISIT\_OCCURRENCE.VISIT\_START\_DATE**  ***Or***  **VISIT\_OCCURRENCE.VISIT\_END\_DATE** | If drug is from ***PATBILL*** use a combination of service day and visit start date unless the service day is greater than the end of the month  If drug comes from ***PATCPT*** or ***PATICD*** then use visit end date |  |
| DRUG\_EXPOSURE\_END\_DATE | ***-*** | NULL |  |
| DRUG\_TYPE\_CONCEPT\_ID |  | 38000180- Inpatient administration |  |
| STOP\_REASON | ***-*** | NULL |  |
| REFILLS | ***-*** | NULL |  |
| QUANTITY | ***PATBILL.STD\_QTY*** |  | Value is applied only to records that come from PATBILL, else records from PATCPT or PATICD are NULL |
| DAYS\_SUPPLY | ***-*** | NULL |  |
| SIG | ***-*** | NULL |  |
| ROUTE\_CONCEPT\_ID | ***-*** | NULL |  |
| EFFECTIVE\_DRUG\_DOSE | ***-*** | NULL |  |
| DOSE\_UNIT\_CONCEPT\_ID | ***-*** | NULL |  |
| LOT\_NUMBER | ***-*** | NULL |  |
| PROVIDER\_ID | ***PAT.ADMPHY\_SPEC*** | NULL |  |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** |  |  |
| DRUG\_SOURCE\_VALUE |  | SELECT SOURCE\_VALUE FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A |  |
| DRUG\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |
| ROUTE\_SOURCE\_VALUE | ***-*** | NULL |  |
| DOSE\_UNIT\_SOURCE\_VALUE | ***-*** | NULL |  |

### Table Name: DRUG\_COST

*Describe how the DRUG\_COST mapping and transformation are designed.*

Cost in Premier is captured in the ***PATBILL*** table. Each ***STD\_CHG\_CODE*** has an associated cost, which is represented as the total paid in ***BILL\_CHARGES***. Additional cost variables in Premier that represent fixed and variable costs are available but not used to populate the CDM. Copays, and amounts paid by the payer are not included in Premier. Drugs that come from the ***PATBILL*** table will have costs attributed to them, drugs that come from ***PATICD*** or ***PATCPT*** will not have a cost represented. Using the mapping from standard charge codes to drugs will identify total drug costs. In Premier each visit is associated with a standard payer, and the PAYER\_PLAN\_PERIOD\_ID is populated for each cost record. The DRUG\_EXPOSURE\_START\_DATE must be between the PAYER\_PLAN\_PERIOD\_START\_DATE and PAYER\_PLAN\_PERIOD\_END\_DATE to ensure only one standard payer is populated per record.

Only drug costs that are not procedure drugs will be included as in the DRUG\_COST table, all procedure drugs are included in the PROCEDURE\_COST table.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| DRUG\_COST\_ID | - | Same as drug\_exposure\_id |  |
| DRUG\_EXPOSURE\_ID | - | Same as drug\_exposure\_id |  |
| CURRENCY\_CONCEPT\_ID | - | 44818668- American dollar |  |
| PAID\_COPAY | - | NULL |  |
| PAID\_COINSURANCE | - | NULL |  |
| PAID\_TOWARD\_DEDUCTIBLE | - | NULL |  |
| PAID\_BY\_PAYER | - | NULL |  |
| PAID\_BY\_COORDINATION\_BENEFITS | - | NULL |  |
| TOTAL\_OUT\_OF\_POCKET | - | NULL |  |
| TOTAL\_PAID | ***PATBILL.BILL\_CHARGES*** | SELECT BILL\_CHARGES FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE, BILL\_CHARGES FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A | Use mapping table to map standard charges to drugs |
| INGREDIENT\_COST | - | NULL |  |
| DISPENSING\_FEE | - | NULL |  |
| AVERAGE\_WHOLESALE\_PRICE | - | NULL |  |
| PAYER\_PLAN\_PERIOD\_ID | ***PAYOR\_PLAN\_PERIOD.PAYER\_PLAN\_PERIOD\_ID*** | Use PAYER\_PLAN\_PERIOD table and join by person. Include filtering to assure that the visit start date falls within the payer start date and payer end date |  |

### Table Name: CONDITION\_OCCURRENCE

*Describe how the Condition\_Occurrence mapping and transformation are designed.*

The CONDITION\_OCCURRENCE table will house records from ***PATICD*** that have been mapped to the condition domain and SNOMED vocabulary.

The conditions in Premier are stored in ***PATICD***. The table houses admitting, primary and secondary diagnosis by visit. Procedure codes and diagnosis codes are stored in the same table. By using the mapping table, ***ICDCODE***, the diagnosis codes can be separated. The CDM transformation captures all 3 types of diagnoses. In many cases patients will have the same admitting and primary diagnosis. The condition start date is determined as the visit start date from the VISIT\_OCCURRENCE table. The exact day of diagnosis is not recorded in Premier, thus the assumption is made that the diagnosis is made on the VISIT\_START\_DATE. The CONDITION\_END\_DATE is null because in Premier we are unaware of when the condition is no longer relevant to the patient. The ASSOCIATED\_PROVIDER\_ID that is provided is the randomly generated key provided by Premier for the provider that admitted the patient. There are two providers that exist in Premier, the admitting and attending. This ETL makes the decision to use admitting because it is unknown whether or not the admitting provider, attending provider or another person diagnosed the person.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| CONDITION\_OCCURRENCE\_ID | ***-*** | System-generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| CONDITION\_CONCEPT\_ID | ***PATICD.ICD\_CODE*** | QUERY: SOURCE TO STANDARD  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('ICD9CM')  AND TARGET\_VOCABUALRY\_ID IN ('SNOMED')  and TARGET\_DOMAIN\_ID = 'Condition' | ICD9 diagnosis codes are mapped to SNOMED concepts |
| CONDITION\_START\_DATE | ***VISIT\_OCCURRENCE.VISIT\_START\_DATE*** |  |  |
| CONDITION\_END\_DATE | ***-*** | NULL |  |
| CONDITION\_TYPE\_CONCEPT\_ID | ***-*** | when ***PAT.I\_O\_IND*** =’I’ and ***ICD.PRI\_SEC*** =’P’ then 38000183  when ***PAT.I\_O\_IND***=’I’ and ***ICD.PRI\_SEC*** =’S’ then 38000185  when ***PAT.I\_O\_IND***=’O’ and ***ICD.PRI\_SEC*** =’P’ then 38000215  when ***PAT.I\_O\_IND***=’O’ and ***ICD.PRI\_SEC*** =’S’ then 38000216  when ***PAT.I\_O\_IND*** in (’I’ , “O”) and ***ICD.PRI\_SEC*** =’A’ then 4203942 |  |
| STOP\_REASON | ***-*** | NULL |  |
| PROVIDER\_ID | ***PAT.ADMPHY\_SPEC*** | NULL |  |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** |  |  |
| CONDITION\_SOURCE\_VALUE | ***PATICD.ICD\_CODE*** |  |  |
| CONDITION\_SOURCE\_CONCEPT\_ID |  | QUERY: SOURCE TO SOURCE  SELECT SOURCE\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('ICD9CM')  AND TARGET\_VOCABULARY\_ID IN ('ICD9CM') AND DOMAIN\_ID=’CONDITION’ |  |

### Table Name: PROCEDURE\_OCCURRENCE

*Describe how the Procedure\_Occurrence mapping and transformation are designed.*

The PROCEDURE\_OCCURRENCE table will house records from ***PATBILL***, ***PATCPT,*** and ***PATICD.*** Procedure records from ***PATBILL*** are mapped to the procedure domain; procedure records from ***PATCPT*** and procedure records from ***PATICD*** (i.e. Vocabulary\_id=’ICD9Proc’) are mapped to the SNOMED vocabulary.

The ***PATBILL*** table holds all charges that were consumed within a visit. For the purpose of our CDM, the drugs are separated and inserted into the DRUG\_EXPOSURE table, and all other billing records are entered into the PROCEDURE\_OCCURRENCE table. For records that are obtained through ***PATBILL***, the start date is determined from the service day in ***PATBILL*** and VISIT\_START\_DATE. If the combination of start date and service day records result in a date greater than the end of the month, the VISIT\_END\_DATE is assigned.

***PATCPT*** houses the HCPCS and CPT codes by visit, and it is unknown when the procedure was performed. ***PATICD*** holds both diagnosis codes and procedure codes. The mapping table ***ICDCODE*** identifies the two different code types. Procedure drugs are recorded as procedure drugs and move to the DRUG\_EXPOSURE table. The procedure start date is identified as the VISIT\_END\_DATE from VISIT\_OCCURRENCE. Procedure type is determined by the indicator of whether or not it was an inpatient stay or outpatient stay. The ASSOCIATED\_PROVIDER\_ID that is provided is the randomly generated key provided by Premier for the provider that admitted the patient. There are two providers that exist in Premier, the admitting and attending. This ETL makes the decision to use admitting because it is unknown whether or not the admitting provider, attending provider or another person diagnosed the person.

In order to map each drug to an appropriate concept, USAGI was used on the ***STD\_CHG\_DESC*** to map the value to a concept; all concepts that map into the procedure domain are included in this table. The ***STD\_CHG\_CODE*** is mapped to a ***HOSP\_CHG*** using ***HOSPCHG***, and each ***HOSP\_CHG*** has a description that is displayed in the CDM along with the standard change code descriptions. Billing records that do not map to a target concept are moved to PROCEDURE\_OCCURRENCE with CONCEPT\_ID=0.

Records that have a valid OBSERVATION\_PERIOD for each patient are included.

The field mapping is performed as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| PROCEDURE\_OCCURRENCE\_ID | - | System-generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| PROCEDURE\_CONCEPT\_ID |  | QUERY SOURCE TO STANDARD:  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('ICD9Proc', 'CPT', 'HCPCS')  AND TARGET\_VOCABUALRY\_ID IN ('SNOMED')  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('JNJ\_PMR\_PROC\_CHRG\_CD' AND TARGET\_DOMAIN\_ID ='Procedure'  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('JNJ\_PMR\_PROC\_CHRG\_CD' AND TARGET\_CONCEPT\_ID=0 | . Also mapped are any ICD9 codes, CPT codes that are mapped to the Procedure Domain. Any concepts that are mapped to zero get placed into the procedure table. |
| PROCEDURE\_DATE | ***VISIT\_OCCURRENCE.VISIT\_END\_DATE*** or  ***VISIT\_OCCURRENCE.VISIT\_START\_DATE***  ***PATBILL.SERV\_DAY*** |  | If the procedure is a CPT code or ICD9 procedure code then discharge date is used as procedure date because the exact date is unknown. If the row is coming from PATBILL then a combination or admit date and service date is used. |
| PROCEDURE\_TYPE\_CONCEPT\_ID | - | When ***PAT.I\_O\_IND*** =’I’ then 38000249  When ***PAT.I\_O\_IND*** =’O’ then 38000267 | All CHGMSTR procedures will be assigned a PROCEDURE\_TYPE\_CONCEPT\_ID indicating 1st position. |
| MODIFIER\_CONCEPT\_ID | - | NULL |  |
| QUANTITY | PATBILL.STD\_QTY | Quantities are populated for all records obtained from the billing record. |  |
| PROVIDER\_ID | ***PAT.ADMPHY\_SPEC*** |  |  |
| VISIT\_OCCURRENCE\_ID | ***PAT.PAT\_KEY*** |  |  |
| PROCEDURE\_SOURCE\_VALUE | ***PATICD.ICD CODE***  ***Or***  ***PATCPT.CPT\_CODE***  ***For all other procedures:***  ***CHGMSTR.STD\_CHG\_CODE\_DESC/ HOSP\_CHG.HOSP\_CHG\_DESC*** | SELECT SOURCE\_VALUE FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A  SELECT SOURCE\_VALUE FROM  (  SELECT ICD\_CODE FROM PATICD A  JOIN CONCEPT C ON  C.CONCEPT\_CODE=A.PATICD  WHERE VOCABULARY\_ID=’ICDProc’  ) A  UNION  (  SELECT CPT\_CODE AS SOURCE\_VALUE FROM PATCPT  ) | To preserve the most detailed description of procedures, if hospital charge descriptions are available, they are to be used, otherwise standard charge code description is displayed |
| PROCEDURE\_SOURCE\_CONCEPT\_ID | ***-*** | SELECT  SOURCE\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('ICD9Proc', 'CPT4', 'HCPCS')  AND TARGET\_VOCABULARY\_ID IN ('ICD9Proc', 'CPT4', 'HCPCS') AND DOMAIN\_ID=’Procedure’  SELECT SOURCE\_VALUE FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A |  |
| QUALIFER\_SOURCE\_VALUE | ***-*** | NULL |  |

### 

### Table Name: PROCEDURE\_COST

*Describe how the PROCEDURE\_COST mapping and transformation are designed.*

Cost in Premier is captured in the PATBILL table. Each ***STD\_CHG\_CODE*** has an associated cost, which is represented as the total paid or ***BILL\_CHARGES***. Additional cost variables in Premier that represent fixed and variable costs are available but not used to populate the CDM. Copays, and amounts paid by the payer are not included in Premier. Procedures that come from the ***PATBILL*** table will have costs attributed to them; procedures that come from ***PATCPT*** or ***PATICD*** will not have a cost represented. The PROCEDURE\_OCCURRENCE\_START\_DATE must be between the PAYER\_PLAN\_PERIOD\_START\_DATE and PAYER\_PLAN\_PERIOD\_END\_DATE to ensure only one standard payer is populated per record.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| PROCEDURE\_COST\_ID | - | Same as procedure\_occurrence\_id |  |
| PROCEDURE\_OCCURRENCE\_ID | - | Same as procedure\_occurrence\_id |  |
| CURRENCY\_CONCEPT\_ID | - | 44818668- American dollar |  |
| PAID\_COPAY | - | NULL |  |
| PAID\_COINSURANCE | - | NULL |  |
| PAID\_TOWARD\_DEDUCTIBLE | - | NULL |  |
| PAID\_BY\_PAYER | - | NULL |  |
| PAID\_BY\_COORDINATION\_BENEFITS | - | NULL |  |
| TOTAL\_OUT\_OF\_POCKET | - | NULL |  |
| TOTAL\_PAID | ***PATBILL.BILL\_CHARGES*** | SELECT BILL\_CHARGES FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE, BILL\_CHARGES FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A |  |
| REVENUE\_CODE\_CONCEPT\_ID | - | NULL |  |
| PAYER\_PLAN\_PERIOD\_ID | ***PAYOR\_PLAN\_PERIOD.PAYER\_PLAN\_PERIOD\_ID*** | Use PAYER\_PLAN\_PERIOD table and join by person. Include filtering to assure that the visit start date falls within the payer start date and payer end date |  |
| REVENUE\_CODE\_SOURCE\_VALUE | - | NULL |  |

### Table Name: DEVICE\_EXPOSURE

*Describe how the DEVICE\_EXPOSURE mapping and transformation are designed.*

The DEVICE\_EXPOSURE table will house records from ***PATBILL, PATCPT,*** and ***PATICD*** that have been mapped to the device domain.

The Premier database has information captured about the various devices in the billing table PATBILL and PATCPT includes codes that are mapped into the device domain. USAGI is used to make PATBILL records to standard concepts, and concepts that are in the Device domain are included. The ASSOCIATED\_PROVIDER\_ID that is provided is the randomly generated key provided by Premier for the provider that admitted the patient. There are two providers that exist in Premier, the admitting and attending. This ETL makes the decision to use admitting because it is unknown whether or not the admitting provider, attending provider or another person diagnosed the person. The ***STD\_CHG\_CODE*** is mapped to a ***HOSP\_CHG*** using ***HOSPCHG***, and each ***HOSP\_CHG*** has a description that is displayed in the CDM along with the standard change code descriptions.

Records that have a valid OBSERVATION\_PERIOD for each patient are included.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| DEVICE\_EXPOSURE\_ID | ***-*** | System-generated |  |
| PERSON\_ID | ***PAT.MEDREC\_KEY*** |  |  |
| DEVICE\_CONCEPT\_ID | - | QUERY:SOURCE To STANDARD:  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID IN ('ICD9Proc', 'CPT', 'HCPCS', 'JNJ\_PMR\_PROC\_CHRG\_CD')  AND TARGET\_DOMAIN\_ID IN ('Device') |  |
| DEVICE\_EXPOSURE\_START\_DATE | ***VISIT\_OCCURRENCE.VISIT\_END\_DATE***  or  ***VISIT\_OCCURRENCE.VISIT\_START\_DATE***  ***PATBILL.SERV\_DAY*** |  | If the device is a CPT code or HCPCS code then discharge date is used as device date because the exact date is unknown. If the row is coming from PATBILL then a combination or admit date and service date is used. |
| DEVICE\_EXPOSURE\_END\_DATE |  |  |  |
| DEVICE\_TYPE\_CONCEPT\_ID | - | 44818705 Inferred from procedure claim |  |
| UNIQUE\_DEVICE\_ID | ***-*** | NULL |  |
| PROVIDER\_ID | ***PAT.ADMPHY\_SPEC*** |  |  |
| VISIT\_OCCURRENCE\_ID | PAT.PAT\_KEY |  |  |
| DEVICE\_SOURCE\_VALUE | ***PATCPT.CPT\_CODE***  ***For all other procedures:***  ***CHGMSTR.STD\_CHG\_CODE\_DESC/ HOSP\_CHG.HOSP\_CHG\_DESC*** | SELECT SOURCE\_VALUE FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A  UNION  (  SELECT CPT\_CODE AS SOURCE\_VALUE FROM PATCPT  ) | To preserve the most detailed description of procedures, if hospital charge descriptions are available, they are to be used, otherwise standard charge code description is displayed |
| DEVICE\_SOURCE\_CONCEPT\_ID | - | NULL |  |

### Table Name: DEVICE\_COST

*Describe how the DEVICE\_COST mapping and transformation are designed.*

Cost in Premier is captured in the PATBILL table. Each ***STD\_CHG\_CODE*** has an associated cost, which is represented as the total paid or ***BILL\_CHARGES***. Additional cost variables in Premier that represent fixed and variable costs are available but not used to populate the CDM. Copays, and amounts paid by the payer are not included in Premier. Devices that come from the ***PATBILL*** table will have costs attributed to them; procedures that come from ***PATCPT*** or PATICD will not have a cost represented. The DEVICE\_OCCURRENCE\_START\_DATE must be between the PAYER\_PLAN\_PERIOD\_START\_DATE and PAYER\_PLAN\_PERIOD\_END\_DATE to ensure only one standard payer is populated per record.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| DEVICE\_COST\_ID |  | Same as device\_concept\_id |  |
| DEVICE\_EXPOSURE\_ID |  | Same as device concept id |  |
| CURRENCY\_CONCEPT\_ID | - | 44818668- American dollar |  |
| PAID\_COPAY | - | NULL |  |
| PAID\_COINSURANCE | - | NULL |  |
| PAID\_TOWARD\_DECDUCTABLE | - | NULL |  |
| PAID\_BY\_PAYER | - | NULL |  |
| PAID\_BY\_COORDINATION\_BENEFITS | - | NULL |  |
| TOTAL\_OUT\_OF\_POCKET | - | NULL |  |
| TOTAL\_PAID | ***PATBILL.BILL\_CHARGES*** | SELECT BILL\_CHARGES FROM  (  SELECT CONCAT(STD\_CHG\_DESC, ' / ', HOSP\_CHG\_DESC) AS SOURCE\_VALUE, BILL\_CHARGES FROM PATBILL A  JOIN CHGMSTR B ON A.STD\_CHG\_CODE=B.STD\_CHG\_CODE  JOIN hospchg C ON A.hosp\_chg\_id=C.hosp\_chg\_id  ) A |  |
| PAYER\_PLAN\_ID | ***PAYOR\_PLAN\_PERIOD.PAYER\_PLAN\_PERIOD\_ID*** | Use PAYER\_PLAN\_PERIOD table and join by person. Include filtering to assure that the visit start date falls within the payer start date and payer end date |  |

### Table Name: NOTE

*Describe how the NOTE mapping and transformation are designed.*

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| NOTE\_ID |  |  |  |
| PERSON\_ID |  |  |  |
| NOTE\_DATE |  |  |  |
| NOTE\_TIME |  |  |  |
| NOTE\_TYPE\_CONCEPT\_ID |  |  |  |
| NOTE\_TEXT |  |  |  |
| PROVIDER\_ID |  |  |  |
| NOTE\_SOURCE\_VALUE |  |  |  |
| VISIT\_OCCURRENCE\_ID |  |  |  |

### Table Name: PROVIDER

*Describe how the PROVIDER mapping and transformation are designed.*

Premier does provide individual providers for each visit, and it houses the admitting and attending provider specialty for each visit. For Premier, each hospital will have a set of providers which will be identified with a unique key generated by Premier. Two providers will be indicated for each visit, an admitting provider, and an attending provider. In most cases these fields are the same, thus the assumption is made to use admitting provider for each stay. Each provider is linked to a specialty. In addition, each provider will be linked to a Premier hospital through the ***PROV\_ID***. The mapping for the specialties from the ***ATTSPEC*** table to valid concepts is located in the mapping table in the appendix. Any providers that are not listed in the look up tables are added manually and associated to an unknown provider. The mapping table uses the most closely matching concept value. Due to orphan records in the PAT table, the following record must be added manually to the provider table:

INSERT INTO provider

VALUES (900, null, null, 38004514,57824101,900, ‘Unknown Physician’)

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| PROVIDER\_ID | ***PAT.ADM\_PHY*** | System generated unique code |  |
| PROVIDER\_NAME | ***-*** | NULL |  |
| NPI | - | NULL |  |
| DEA | - | NULL |  |
| SPECIALTY\_CONCEPT\_ID |  | QUERY: SOURCE TO STANDARD  SELECT TARGET\_CONCEPT\_ID  FROM CTE\_VOCAB\_MAP  WHERE SOURCE\_VOCABULARY\_ID = 'JNJ\_PMR\_P\_SPCLTY' |  |
| CARE\_SITE\_ID | ***PAT.PROV\_ID*** |  |  |
| YEAR\_OF\_BIRTH | ***-*** | NULL |  |
| GENDER\_CONCEPT\_ID | ***-*** | NULL |  |
| PROVIDER\_SOURCE\_VALUE | ***PAT.ADMPHY\_SPEC*** |  |  |
| SPECIALTY\_SOURCE\_VALUE | ***ATTSPEC***.***ADMPHY\_SPEC\_DESC*** |  |  |
| SPECIALITY\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |
| GENDER\_SOURCE\_VALUE | ***-*** | NULL |  |
| GENDER\_SOURCE\_CONCEPT\_ID | ***-*** | NULL |  |

### Table Name: CARE\_SITE

*Describe how the CARE\_SITE mapping and transformation are designed.*

In Premier each CARE\_SITE is considered a hospital. Premier has additional information about the whether or not the hospital is in an urban area versus a rural area, if it’s a teaching hospital, and the number of beds. The information is inserted as a concatenated field in the PLACE\_OF\_SERVICE\_SOURCE\_VALUE field. Each ***PAT\_KEY*** or visit, a ***PAT.PROV\_ID*** is assigned as the hospital. The ***PROVIDER*** table provides the additional information about the hospital.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| CARE\_SITE\_ID | ***PROVIDERS.PROV\_ID*** |  |  |
| CARE\_SITE\_NAME | ***-*** | NULL |  |
| PLACE\_OF\_SERVICE\_CONCEPT\_ID | ***-*** | NULL |  |
| LOCATION\_ID |  | Link to LOCATION\_ID using LOCATION table |  |
|  |  |  |  |
| CARE\_SITE\_SOURCE\_VALUE | ***PROVIDERS.PROV\_ID*** | NULL |  |
| PLACE\_OF\_SERVICE\_SOURCE\_VALUE | ***PROVIDERS.TEACHING and***  ***PROVIDERS.URBAN\_RURAL***  ***and***  ***PROVIDERS.BEDS*** | Provides the values for the additional place of service variables, in the specified format below:  Urban\_rural: <value> Teaching: < value> Beds: <value> |  |

**SAMPLE CODE:**

SELECT distinct a.prov\_id as care\_site\_id,

b.location\_id,

**0** as organization\_id,

**0** as place\_of\_service\_concept\_id,

prov\_id as care\_site\_source\_value,

('Urban\_Rural:' + urban\_rural + ' ' + 'Teaching:' + teaching + ' ' + 'Beds:' + cast(beds as varchar(**5**))) as place\_of\_service\_source\_value

FROM (SELECT A.\*, PROV\_AREA + ' ' + PROV\_DIVISION AS LOCATION\_SOURCE\_VALUE FROM rndusrdhit02.Premier.dbo.providers A) a

Join rndusrdhit02.cdm\_Premier.dbo.LOCATION b on a.LOCATION\_SOURCE\_VALUE=b.location\_source\_value

### Table Name: FACT\_RELATIONSHIP

*Describe how the FACT\_RELATIONSHIP mapping and transformation are designed.*

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| DOMAIN\_CONCEPT\_ID\_1 |  | System generated |  |
| FACT\_ID\_1 | - | NULL |  |
| DOMAIN\_CONCEPT\_ID\_2 | - | NULL |  |
| FACT\_ID\_2 | - | NULL |  |
| RELATIONSHIP\_CONCEPT\_ID |  |  |  |

### Table Name: LOCATION

*Describe how the LOCATION mapping and transformation are designed.*

Premier has the census values for the location of each Premier hospital. The region that the hospital or CARE\_SITE is in, one of 4 valid values is: Northeast, Midwest, South, and West.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| LOCATION\_ID |  | System generated |  |
| ADDRESS\_1 | - | NULL |  |
| ADDRESS\_2 | - | NULL |  |
| CITY | - | NULL |  |
| STATE | - | NULL |  |
| ZIP | - | NULL |  |
| COUNTY | - | NULL |  |
| LOCATION\_SOURCE\_VALUE | ***PROVIDERS.PROV\_AREA*** |  | Only 4 values are valid in Premier |

**SAMPLE CODE:**

select distinct prov\_area from providers

### Table Name: COHORT

*Describe how the COHORT mapping and transformation are designed.*

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| COHORT\_ID |  |  |  |
| COHORT\_CONCEPT\_ID |  |  |  |
| COHORT\_START\_DATE |  |  |  |
| COHORT\_END\_DATE |  |  |  |
| SUBJECT\_ID |  |  |  |
| STOP\_REASON |  |  |  |

## Source Independent Data Mapping

### Table Name: DRUG\_ERA

A Drug Era is defined as a span of time when the Person is assumed to be exposed to a particular drug. Successive periods of Drug Exposures are combined under certain rules to produce continuous Drug Eras. The DRUG\_ERA table is populated by pulling from the DRUG\_EXPOSURE table within the CDM. Drug eras are consolidated to their respective ingredient off the DRUG\_EXPOSURE table. A drug era is therefore understood as exposure to a certain compound over a certain period of time. There will only be one type of persistence window (duration that is allowed to elapse between drug exposures) applied to this CDM, which is 30 days.

Drugs that are mapped to a DRUG\_CONCEPT\_ID=0 should not be mapped. The logic below is used to map DRUG\_CONCEPT\_ID’s to ingredients.

SELECT DISTINCT ca.ANCESTOR\_CONCEPT\_ID /\*ingredient level\*/,

ca.DESCENDANT\_CONCEPT\_ID /\*this is where you set the DRUG\_EXPOSURE.DRUG\_CONCEPT\_ID to\*/

FROM CONCEPT c1

JOIN CONCEPT\_ANCESTOR ca

ON ca.DESCENDANT\_CONCEPT\_ID = c1.CONCEPT\_ID

JOIN CONCEPT c2

ON c2.CONCEPT\_ID = ca.ANCESTOR\_CONCEPT\_ID

AND c2.CONCEPT\_VOCABULARY\_ID = 8

AND c2.CONCEPT\_LEVEL = 2

WHERE c1.CONCEPT\_VOCABULARY\_ID = 8

Do not include records that cannot be mapped to the ingredient level. The DRUG\_EXPOSURE\_END\_DATE is the DRUG\_EXPOSURE\_START\_DATE.

The field mapping is as follows:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| DRUG\_ERA\_ID |  | System generated |  |
| PERSON\_ID | PERSON\_ID |  |  |
| DRUG\_CONCEPT\_ID | DRUG\_CONCEPT\_ID | Do no create DRUG\_ERAs where the DRUG\_EXPOSURE.DRUG\_CONCEPT\_ID is 0.  Use the map above to map DRUG\_EXPOSURE.DRUG\_CONCEPT\_ID to the ingredient level DRUG\_CONCEPT\_ID used in the DRUG\_ERA. |  |
| DRUG\_ERA\_START\_DATE | DRUG\_EXPOSURE\_START\_DATE |  | The start date for the drug era constructed from the individual instances of drug exposures. It is the start date of the very first chronologically recorded instance of utilization of a drug. |
| DRUG\_ERA\_END\_DATE | DRUG\_EXPOSURE.START\_DATE |  |  |
| DRUG\_TYPE\_CONCEPT\_ID | - | Apply a 30 day persistence window and label as CONCEPT\_ID 38000182 (Drug era - 30 days persistence window). | Falls under CONCEPT\_VOCABULARY\_ID = 36 as a Drug Exposure Type. |
| DRUG\_EXPOSURE\_COUNT | - | Sum up the number of DRUG\_EXPOSURES for this PERSON\_ID and this CONCEPT\_ID during the exposure window being built. |  |

### 

### Table Name: CONDITION\_ERA

CONDITION\_ERAs are chronological periods of condition occurrence. There will only be one type of persistence window (duration that is allowed to elapse between condition occurrences) applied to this CDM, which is 30 days. CONDITION\_END\_DATE will be the CONDITION\_START\_DATE.

Exclude records with a CONDITION\_CONCEPT\_ID=0.

All Condition Eras are recorded in the CONDITION\_ERA table based on the following field mapping:

| **Destination Field** | **Source Field** | **Applied Rule** | **Comment** |
| --- | --- | --- | --- |
| CONDITION\_ERA\_ID |  | System-generated |  |
| PERSON\_ID | PERSON\_ID |  |  |
| CONDITION\_CONCEPT\_ID | CONDITION\_CONCEPT\_ID | Do not build condition\_era where the condition\_occurance\_condition\_concept\_id=0 |  |
| CONDITION\_ERA\_START\_DATE | CONDITION\_START\_DATE |  | The start date for the condition era constructed from the individual instances of condition occurrences. It is the start date of the very first chronologically recorded instance of the condition. |
| CONDITION\_ERA\_END\_DATE | CONDITION\_START\_DATE |  | The end date for the condition era constructed from the individual instances of condition occurrences. It is the end date of the final continuously recorded instance of the condition. |
| CONDITION\_TYPE\_CONCEPT\_ID |  | Apply a 30 day persistence window and label as CONCEPT\_ID 38000247 (Condition era - 30 days persistence window). | Falls under CONCEPT\_VOCABULARY\_ID = 37 - OMOP Condition Occurrence Type. |
| CONDITION\_OCCURRENCE\_COUNT |  | Sum up the number of CONDITION\_OCCURRENCEs for this PERSON\_ID and this CONCEPT\_ID during the exposure window being built. |  |

### 

### table name: DOSE\_ERA

A Dose Era is defined as a span of time when the Person is assumed to be exposed to a constant dose of a specific active ingredient.

| Destination Field | Source Field | Applied Rule | Comment |
| --- | --- | --- | --- |
| DOSE\_ERA\_ID |  | System-generated |  |
| PERSON\_ID | PERSON\_ID |  |  |
| DRUG\_CONCEPT\_ID | DRUG\_CONCEPT\_ID | Do not build dose\_era where the drug\_concept\_id=0 |  |
| UNIT\_CONCEPT\_ID |  |  |  |
| DOSE\_VALUE |  |  | Numeric value of dose |
| DOSE\_ERA\_START\_DATE | DRUG\_EXPOSURE\_START\_DATE |  | The start date for the dose era constructed from the individual instances of drug exposures. It is the start date of the very first chronologically recorded instance of utilization of a drug. |
| DOSE\_ERA\_END\_DATE |  |  |  |

# Appendix

Queries for VOCABULARY:

--SOURCE TO SOURCE

WITH CTE\_VOCAB\_MAP AS (

SELECT c.concept\_code AS SOURCE\_CODE, c.concept\_id AS SOURCE\_CONCEPT\_ID, c.vocabulary\_id AS SOURCE\_VOCABULARY\_ID, c.domain\_id AS SOURCE\_DOMAIN\_ID, c.concept\_class\_id AS SOURCE\_CONCEPT\_CLASS\_ID,

c.invalid\_reason AS SOURCE\_INVALID\_REASON,

c.concept\_ID as TARGET\_CONCEPT\_ID, c.vocabulary\_id AS TARGET\_VOCABULARY\_ID, c.domain\_id AS TARGET\_DOMAIN\_ID, c.concept\_class\_id AS TARGET\_CONCEPT\_CLASS\_ID, c.INVALID\_REASON AS TARGET\_INVALID\_REASON,

c.STANDARD\_CONCEPT AS TARGET\_STANDARD\_CONCEPT

FROM CONCEPT c

UNION

SELECT source\_code, SOURCE\_CONCEPT\_ID, source\_vocabulary\_id, c1.domain\_id AS SOURCE\_DOMAIN\_ID, c2.CONCEPT\_CLASS\_ID AS SOURCE\_CONCEPT\_CLASS\_ID,

stcm.INVALID\_REASON AS SOURCE\_INVALID\_REASON,target\_concept\_id, target\_vocabulary\_id, c2.domain\_id AS TARGET\_DOMAIN\_ID, c2.concept\_class\_id AS TARGET\_CONCEPT\_CLASS\_ID,

c2.INVALID\_REASON AS TARGET\_INVALID\_REASON, c2.standard\_concept AS TARGET\_STANDARD\_CONCEPT

FROM source\_to\_concept\_map stcm

LEFT OUTER JOIN CONCEPT c1

ON c1.concept\_id = stcm.source\_concept\_id

LEFT OUTER JOIN CONCEPT c2

ON c2.CONCEPT\_ID = stcm.target\_concept\_id

WHERE stcm.INVALID\_REASON IS NULL

)

--SOURCE TO STANDARD

WITH CTE\_VOCAB\_MAP AS (

SELECT c.concept\_code AS SOURCE\_CODE, c.concept\_id AS SOURCE\_CONCEPT\_ID, c.vocabulary\_id AS SOURCE\_VOCABULARY\_ID,

c.domain\_id AS SOURCE\_DOMAIN\_ID, c.CONCEPT\_CLASS\_ID AS SOURCE\_CONCEPT\_CLASS\_ID, c.INVALID\_REASON AS SOURCE\_INVALID\_REASON,

c1.concept\_id AS TARGET\_CONCEPT\_ID, c1.VOCABULARY\_ID AS TARGET\_VOCABUALRY\_ID, c1.domain\_id AS TARGET\_DOMAIN\_ID, c1.concept\_class\_id AS TARGET\_CONCEPT\_CLASS\_ID,

c1.INVALID\_REASON AS TARGET\_INVALID\_REASON, c1.standard\_concept AS TARGET\_STANDARD\_CONCEPT

FROM CONCEPT C

JOIN CONCEPT\_RELATIONSHIP CR

ON C.CONCEPT\_ID = CR.CONCEPT\_ID\_1

AND CR.invalid\_reason IS NULL

AND cr.relationship\_id = 'Maps To'

JOIN CONCEPT C1

ON CR.CONCEPT\_ID\_2 = C1.CONCEPT\_ID

AND C1.INVALID\_REASON IS NULL

UNION

SELECT source\_code, SOURCE\_CONCEPT\_ID, source\_vocabulary\_id, c1.domain\_id AS SOURCE\_DOMAIN\_ID, c2.CONCEPT\_CLASS\_ID AS SOURCE\_CONCEPT\_CLASS\_ID,

stcm.INVALID\_REASON AS SOURCE\_INVALID\_REASON,

target\_concept\_id, target\_vocabulary\_id, c2.domain\_id AS TARGET\_DOMAIN\_ID, c2.concept\_class\_id AS TARGET\_CONCEPT\_CLASS\_ID, c2.INVALID\_REASON AS TARGET\_INVALID\_REASON,

c2.standard\_concept AS TARGET\_STANDARD\_CONCEPT

FROM source\_to\_concept\_map stcm

LEFT OUTER JOIN CONCEPT c1

ON c1.concept\_id = stcm.source\_concept\_id

LEFT OUTER JOIN CONCEPT c2

ON c2.CONCEPT\_ID = stcm.target\_concept\_id

WHERE stcm.INVALID\_REASON IS NULL

)

## Premier mapping to CDM fields



## STD\_CHG\_CODE to SOURCE\_TO\_CONCEPT\_MAP table

## Provider Specialty mapping table



## STD\_CHG\_CODE to HOSP\_CHG\_ID Mapping table