Source Data Mapping Approach to CDMV6.0

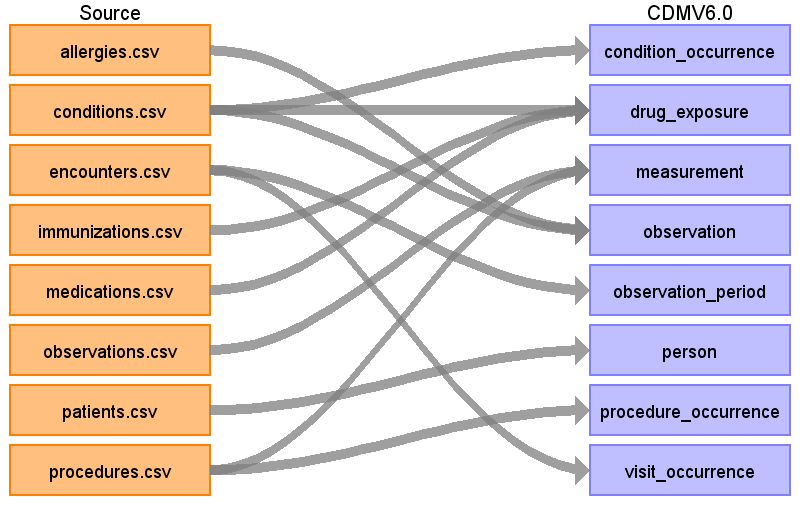
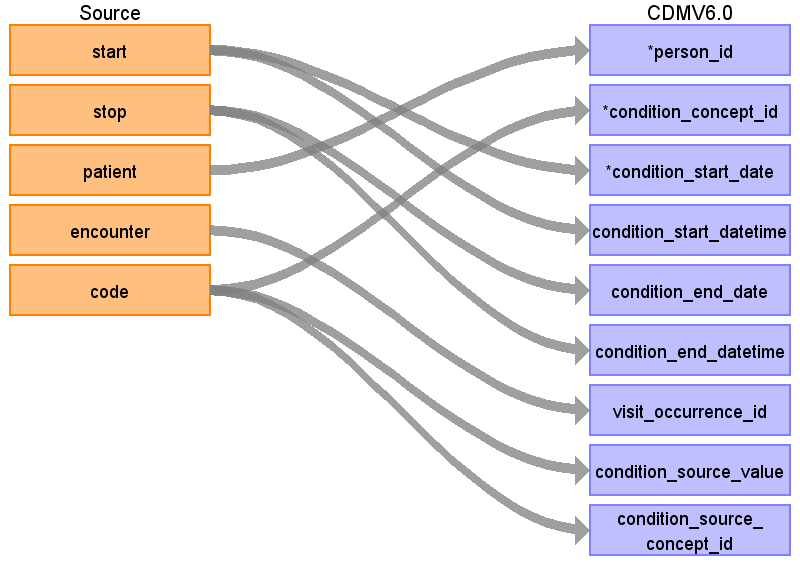


Table name: condition\_occurrence

Reading from conditions.csv

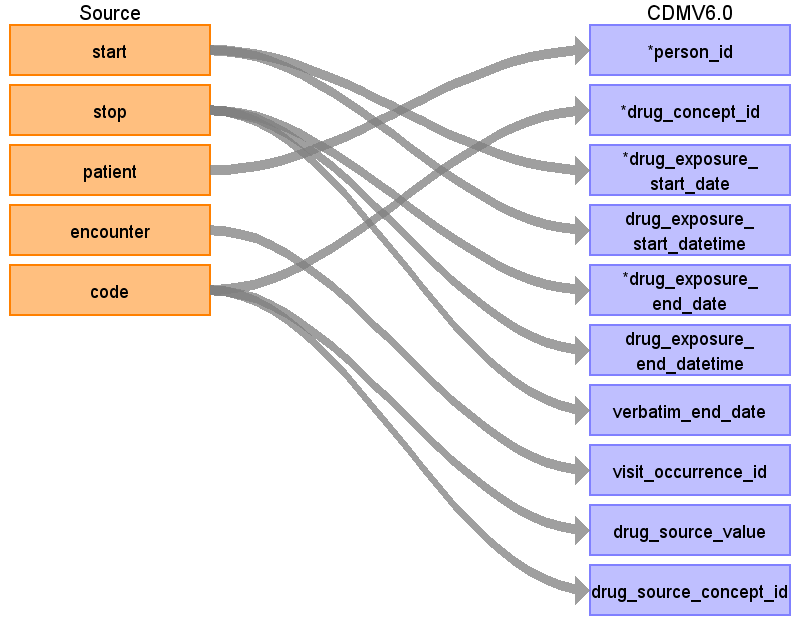


|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| condition\_occurrence\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| condition\_concept\_id | code | Use the following SQL code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from conditions c  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = c.code  and ctvm.target\_domain\_id = 'Condition'  and ctvm.target\_vocabulary\_id = 'SNOMED'  and ctvm.target\_invalid\_reason is NULL  and ctvm.target\_standard\_concept = 'S' |  |
| condition\_start\_date | start |  |  |
| condition\_start\_datetime | start |  |  |
| condition\_end\_date | stop |  |  |
| condition\_end\_datetime | stop |  |  |
| condition\_type\_concept\_id |  |  | Set to 32020 (EHR Encounter Diagnosis) for all records |
| stop\_reason |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| condition\_source\_value | code |  |  |
| condition\_source\_concept\_id | code | Use the following SQL code to lookup the source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join conditions c  on csvm.source\_code = c.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| condition\_status\_source\_value |  |  |  |
| condition\_status\_concept\_id |  |  |  |

Table name: device\_exposure

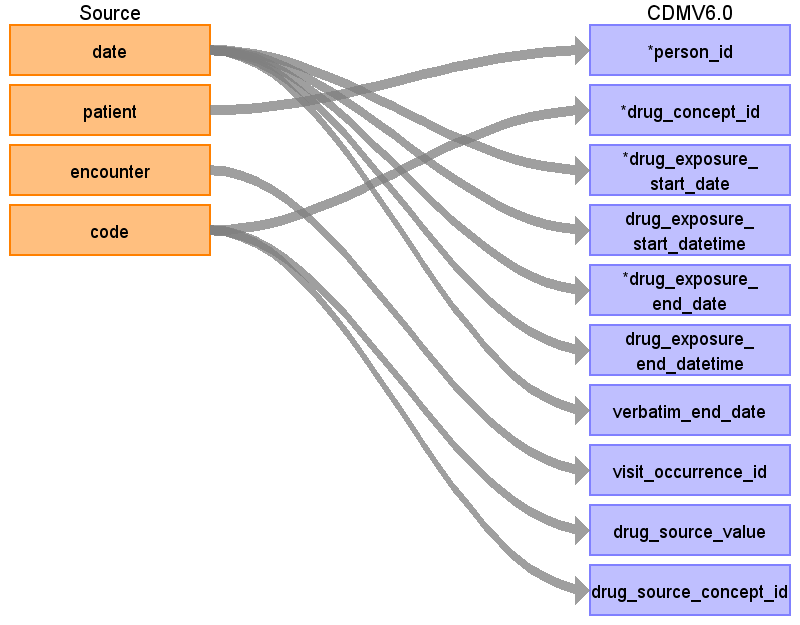
Table name: drug\_exposure

Reading from conditions.csv



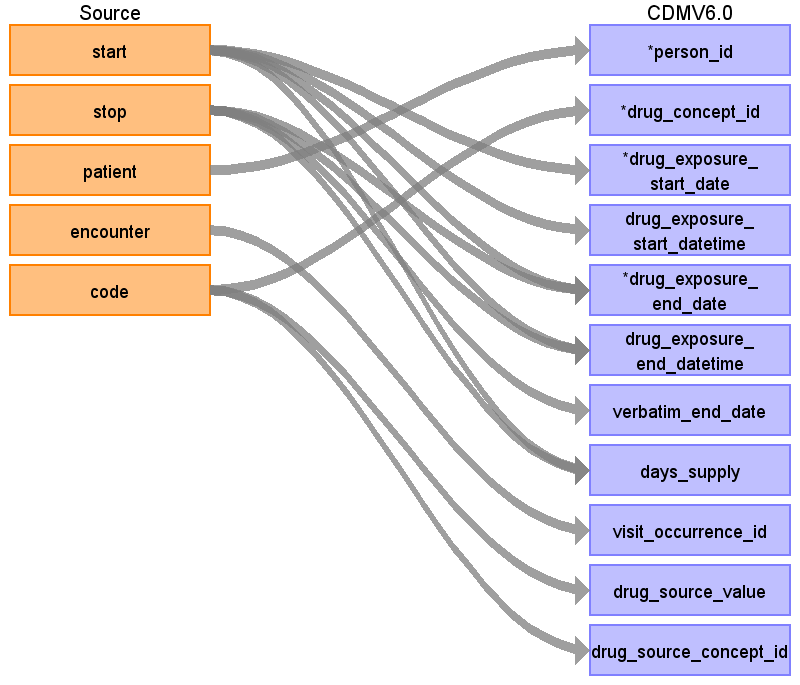
|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| drug\_exposure\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| drug\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from conditions c  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = c.code  and ctvm.target\_domain\_id = 'Drug'  and ctvm.target\_vocabulary\_id = 'RxNorm'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| drug\_exposure\_start\_date | start |  |  |
| drug\_exposure\_start\_datetime | start |  |  |
| drug\_exposure\_end\_date | stop |  | If there is no stop date, use the start date instead. |
| drug\_exposure\_end\_datetime | stop |  | If there is no stop date, use the start date instead. |
| verbatim\_end\_date | stop |  |  |
| drug\_type\_concept\_id |  |  | Use the concept\_id 581452 for all records from the immunizations and conditions tables and concept\_id 38000177 for all records from the medications table. |
| stop\_reason |  |  |  |
| refills |  |  |  |
| quantity |  |  |  |
| days\_supply |  |  | Subtract the start date from the stop date to find the days\_supply. If there is no stop date this should be 0. |
| sig |  |  |  |
| route\_concept\_id |  |  |  |
| lot\_number |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| drug\_source\_value | code |  |  |
| drug\_source\_concept\_id | code | Use code to lookup source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join conditions c  on csvm.source\_code = c.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| route\_source\_value |  |  |  |
| dose\_unit\_source\_value |  |  |  |

Reading from immunizations.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| drug\_exposure\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping immunizations.patient to person.person\_source\_value. |  |
| drug\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from immunizations i  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = i.code  and ctvm.target\_domain\_id = 'Drug'  and ctvm.target\_vocabulary\_id = 'CVX'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| drug\_exposure\_start\_date | date |  |  |
| drug\_exposure\_start\_datetime | date | Use 00:00:00 as the time. |  |
| drug\_exposure\_end\_date | date |  | If there is no stop date, use the start date instead. |
| drug\_exposure\_end\_datetime | date | Use 00:00:00 as the time. | If there is no stop date, use the start date instead. |
| verbatim\_end\_date | date |  |  |
| drug\_type\_concept\_id |  |  | Use the concept\_id 581452 for all records from the immunizations and conditions tables and concept\_id 38000177 for all records from the medications table. |
| stop\_reason |  |  |  |
| refills |  |  |  |
| quantity |  |  |  |
| days\_supply |  |  | Subtract the start date from the stop date to find the days\_supply. If there is no stop date this should be 0. |
| sig |  |  |  |
| route\_concept\_id |  |  |  |
| lot\_number |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| drug\_source\_value | code |  |  |
| drug\_source\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from immunizations i  join cte\_source\_vocab\_map csvm  on csvm.source\_code = i.code  and csvm.source\_vocabulary\_id = 'CVX' |  |
| route\_source\_value |  |  |  |
| dose\_unit\_source\_value |  |  |  |

Reading from medications.csv

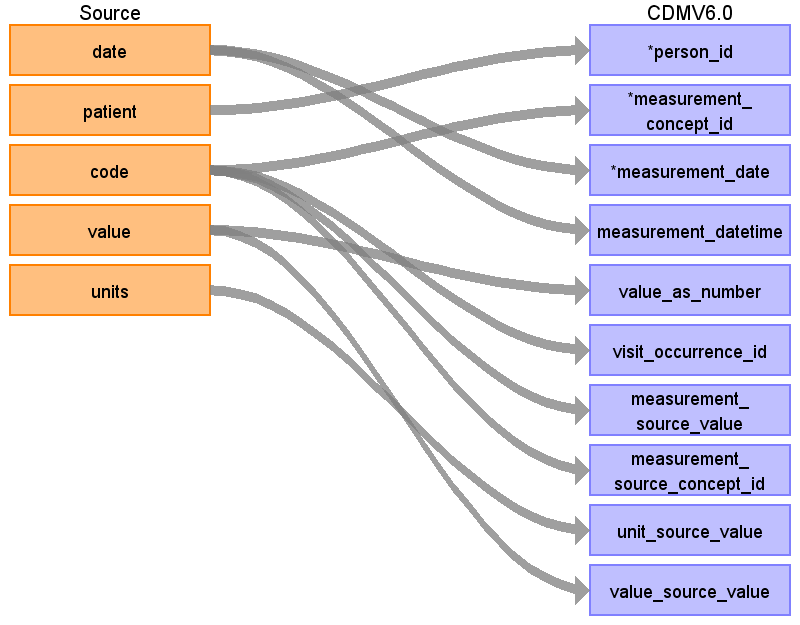


|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| drug\_exposure\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping medications.patient to person.person\_source\_value. |  |
| drug\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from medications m  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = m.code  and ctvm.target\_domain\_id = 'Drug'  and ctvm.target\_vocabulary\_id = 'RxNorm'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| drug\_exposure\_start\_date | start |  |  |
| drug\_exposure\_start\_datetime | start | Use 00:00:00 as the time. |  |
| drug\_exposure\_end\_date | stop  start |  | If there is no stop date, use the start date instead. |
| drug\_exposure\_end\_datetime | stop  start | Use 00:00:00 as the time. | If there is no stop date, use the start date instead. |
| verbatim\_end\_date | stop |  |  |
| drug\_type\_concept\_id |  |  | Use the concept\_id 581452 for all records from the immunizations and conditions tables and concept\_id 38000177 for all records from the medications table. |
| stop\_reason |  |  |  |
| refills |  |  |  |
| quantity |  |  |  |
| days\_supply | start  stop |  | Subtract the start date from the stop date to find the days\_supply. If there is no stop date this should be 0. |
| sig |  |  |  |
| route\_concept\_id |  |  |  |
| lot\_number |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| drug\_source\_value | code |  |  |
| drug\_source\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from medications m  join cte\_source\_vocab\_map csvm  on cvm.source\_code = m.code  and cvm.source\_vocabulary\_id = 'RxNorm' |  |
| route\_source\_value |  |  |  |
| dose\_unit\_source\_value |  |  |  |

Table name: fact\_relationship

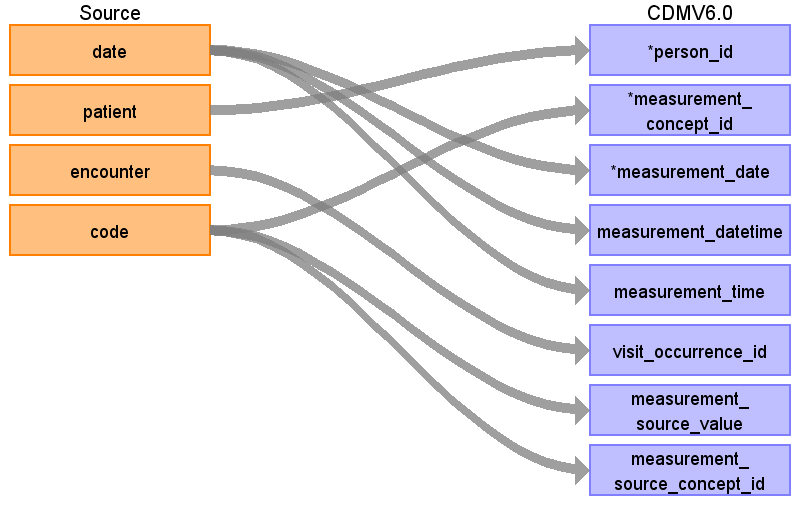
Table name: measurement

Reading from observations.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| measurement\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from observations o  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = o.code  and ctvm.target\_domain\_id = 'Measurement'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| measurement\_date | date |  |  |
| measurement\_datetime | date | Use 00:00:00 as time. |  |
| measurement\_time |  |  |  |
| measurement\_type\_concept\_id |  |  | Use concept\_id 5001 for all records |
| operator\_concept\_id |  |  |  |
| value\_as\_number | value |  |  |
| value\_as\_concept\_id |  |  |  |
| unit\_concept\_id |  |  |  |
| range\_low |  |  |  |
| range\_high |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | code | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| measurement\_source\_value | code |  |  |
| measurement\_source\_concept\_id | code | Use code to lookup source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join observations o  on csvm.source\_code = o.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| unit\_source\_value | units |  |  |
| value\_source\_value | value |  |  |

Reading from procedures.csv



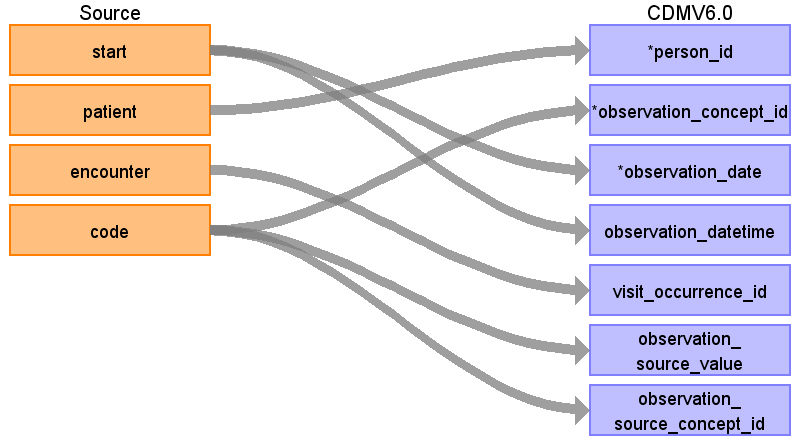
|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Autogenerate |
| person\_id | patient | Lookup in the person table: map by mapping person.person\_source\_value to patient. |  |
| measurement\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from procedures pr  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = pr.code  and ctvm.target\_domain\_id = 'Measurement'  and ctvm.target\_vocabulary\_id = 'SNOMED'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| measurement\_date | date |  |  |
| measurement\_datetime | date | Use 00:00:00 as the time. |  |
| measurement\_time | date |  |  |
| measurement\_type\_concept\_id |  |  | Use concept\_id 5001 for all records |
| operator\_concept\_id |  |  |  |
| value\_as\_number |  |  |  |
| value\_as\_concept\_id |  |  |  |
| unit\_concept\_id |  |  |  |
| range\_low |  |  |  |
| range\_high |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| measurement\_source\_value | code |  |  |
| measurement\_source\_concept\_id | code | Use code to lookup source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join procedures pr  on csvm.source\_code = pr.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| unit\_source\_value |  |  |  |
| value\_source\_value |  |  |  |

Table name: note

Table name: note\_nlp

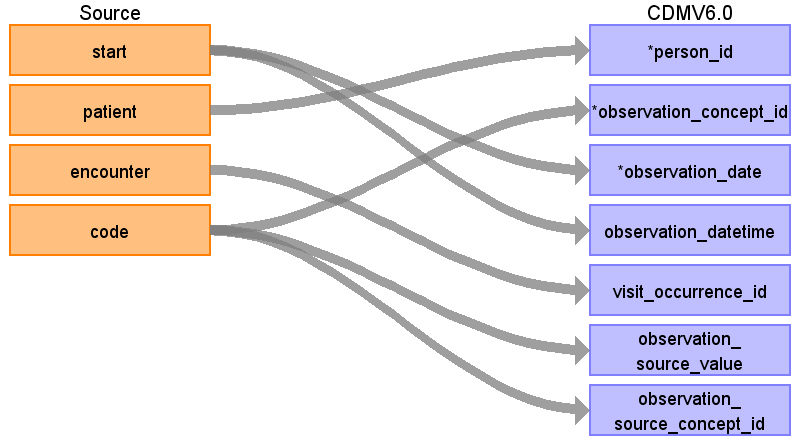
Table name: observation

Reading from allergies.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| observation\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping allergies.patient to person.person\_source\_value. |  |
| observation\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from allergies a  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = a.code  and ctvm.target\_domain\_id = 'Observation'  and ctvm.target\_vocabulary\_id = 'SNOMED'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| observation\_date | start |  |  |
| observation\_datetime | start | Use 00:00:00 as time. |  |
| observation\_type\_concept\_id |  |  | Set as 38000280 for all records. |
| value\_as\_number |  |  |  |
| value\_as\_string |  |  |  |
| value\_as\_concept\_id |  |  |  |
| qualifier\_concept\_id |  |  |  |
| unit\_concept\_id |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| observation\_source\_value | code |  |  |
| observation\_source\_concept\_id | code | Use code to lookup source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join allergies a  on csvm.source\_code = a.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| unit\_source\_value |  |  |  |
| qualifier\_source\_value |  |  |  |
| observation\_event\_id |  |  |  |
| obs\_event\_field\_concept\_id |  |  |  |
| value\_as\_datetime |  |  |  |

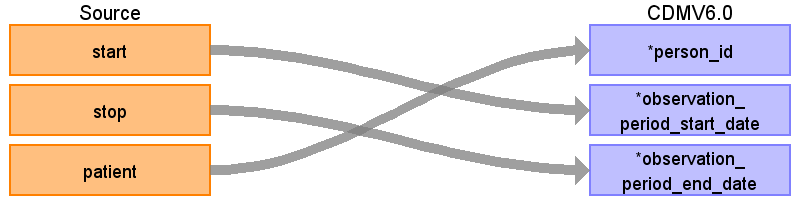
Reading from conditions.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| observation\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| observation\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from conditions c  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = c.code  and ctvm.target\_domain\_id = 'Observation'  and ctvm.target\_vocabulary\_id = 'SNOMED'  and ctvm.target\_standard\_concept = 'S'  and ctvm.target\_invalid\_reason is NULL |  |
| observation\_date | start |  |  |
| observation\_datetime | start |  |  |
| observation\_type\_concept\_id |  |  | Set as 38000280 for all records. |
| value\_as\_number |  |  |  |
| value\_as\_string |  |  |  |
| value\_as\_concept\_id |  |  |  |
| qualifier\_concept\_id |  |  |  |
| unit\_concept\_id |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| observation\_source\_value | code |  |  |
| observation\_source\_concept\_id | code | Use code to lookup source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join conditions c  on csvm.source\_code = c.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| unit\_source\_value |  |  |  |
| qualifier\_source\_value |  |  |  |
| observation\_event\_id |  |  |  |
| obs\_event\_field\_concept\_id |  |  |  |
| value\_as\_datetime |  |  |  |

Table name: observation\_period

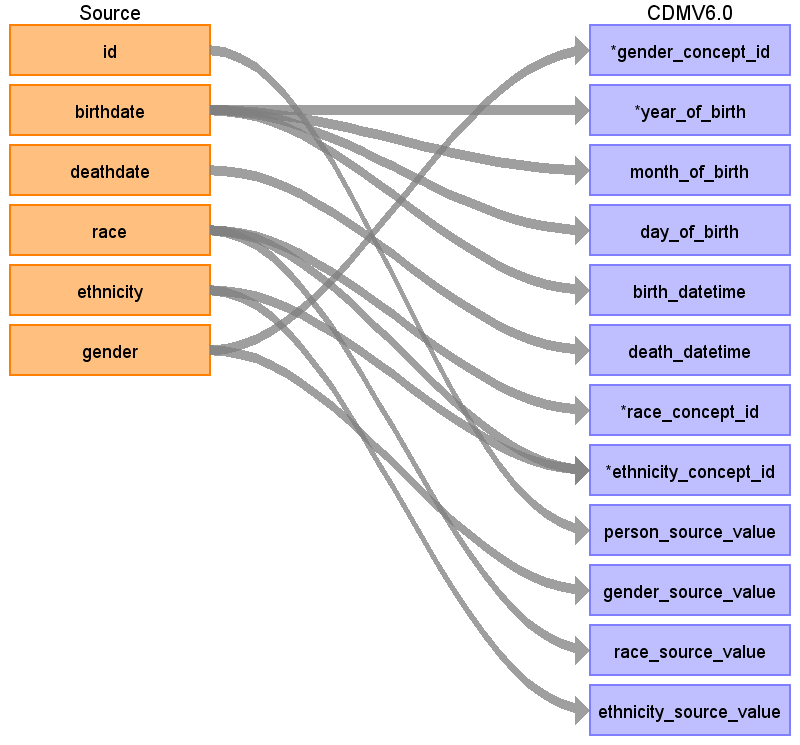
Reading from encounters.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| observation\_period\_id |  |  |  |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| observation\_period\_start\_date | start | min(start) group by patient  Take the earliest START per patient |  |
| observation\_period\_end\_date | stop | max(stop) group by patient  Take the latest STOP per patient |  |
| period\_type\_concept\_id |  |  | Set as concept 44814724 for all records |

Table name: person

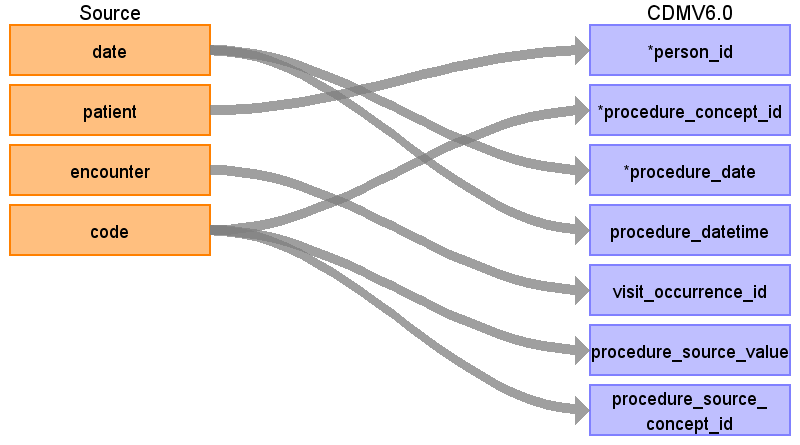
Reading from patients.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id |  |  | Autogenerate |
| gender\_concept\_id | gender | When gender = ‘M’ then set gender\_concept\_id to 8507, when gender = ‘F’ then set to 8532 | Drop any rows with missing/unknown gender. |
| year\_of\_birth | birthdate | Take year from birthdate |  |
| month\_of\_birth | birthdate | Take month from birthdate |  |
| day\_of\_birth | birthdate | Take day from birthdate |  |
| birth\_datetime | birthdate | With midnight as time 00:00:00 |  |
| death\_datetime | deathdate | With midnight as time 00:00:00 |  |
| race\_concept\_id | race | When race = ‘WHITE’ then set as 8527, when race = ‘BLACK’ then set as 8516, when race = ‘ASIAN’ then set as 8515, otherwise set as 0 |  |
| ethnicity\_concept\_id | race  ethnicity | When race = ‘HISPANIC’, or when ethnicity in (‘CENTRAL\_AMERICAN’, ‘DOMINICAN’, ‘MEXICAN’, ‘PUERTO\_RICAN’, ‘SOUTH\_AMERICAN’ ) then set as 38003563, otherwise set as 0  When race = ‘HISPANIC’, or when ethnicity in (‘CENTRAL\_AMERICAN’, ‘DOMINICAN’, ‘MEXICAN’, ‘PUERTO\_RICAN’, ‘SOUTH\_AMERICAN’ ) then set as 38003563, otherwise set as 0 |  |
| location\_id |  |  |  |
| provider\_id |  |  |  |
| care\_site\_id |  |  |  |
| person\_source\_value | id |  |  |
| gender\_source\_value | gender |  |  |
| gender\_source\_concept\_id |  |  |  |
| race\_source\_value | race |  |  |
| race\_source\_concept\_id |  |  |  |
| ethnicity\_source\_value | ethnicity |  |  |
| ethnicity\_source\_concept\_id |  |  |  |

Table name: procedure\_occurrence

Reading from procedures.csv



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| procedure\_occurrence\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| procedure\_concept\_id | code | Use code to lookup target\_concept\_id in CTE\_TARGET\_VOCAB\_MAP:  select ctvm.target\_concept\_id  from procedures pr  join cte\_target\_vocab\_map ctvm  on ctvm.source\_code = pr.code  and ctvm.target\_domain\_id = 'Procedure'  and ctvm.target\_vocabulary\_id = 'SNOMED'  and ctvm.target\_invalid\_reason is NULL  and ctvm.target\_standard\_concept = 'S' |  |
| procedure\_date | date |  |  |
| procedure\_datetime | date | Use 00:00:00 as the time. |  |
| procedure\_type\_concept\_id |  |  | Use concept\_id 38000275 for all records. |
| modifier\_concept\_id |  |  |  |
| quantity |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id | encounter | Lookup visit\_occurrence\_id using encounter, joining to temp table defined in AllVisitTable.sql. |  |
| visit\_detail\_id |  |  |  |
| procedure\_source\_value | code |  |  |
| procedure\_source\_concept\_id | code | Use code to lookup source\_concept\_id in CTE\_SOURCE\_VOCAB\_MAP:  select csvm.source\_concept\_id  from cte\_source\_vocab\_map csvm  join procedures pr  on csvm.source\_code = pr.code  and csvm.source\_vocabulary\_id = 'SNOMED' |  |
| modifier\_source\_value |  |  |  |

Table name: specimen

Table name: visit\_detail

Table name: visit\_occurrence

Reading from encounters.csv

STEP 1

For encounterclass = ‘Inpatient’ (IP):

Sort data in ascending order by PATIENT, START, END.

Then by PERSON\_ID, collapse lines of claim as long as the time between the END of one line and the START of the next is <=1 day.

Each consolidated inpatient claim is then considered as one inpatient visit, set

MIN(START) as VISIT\_START\_DATE

MAX(END) as VISIT\_END\_DATE

‘IP’ as PLACE\_OF\_SERVICE\_SOURCE\_VALUE

See if any records with encounterclass ‘outpatient’ (OP), ‘ambulatory’ (OP), ‘wellness’ (OP), ‘emergency’ (ER) or ‘urgentcare’ (ER) occur during an identified ‘inpatient’ visit. These should be consolidated into the ‘inpatient’ visit, unless it is an ‘emergency’ or ‘urgentcare’ visit that starts and ends on the first day of the ‘inpatient’ visit. Types of outpatient (OP) visits not collapsed:

If an OP starts before an IP but ends during an IP

If an OP starts before and ends after an IP visit. If an OP is collapsed into an IP and its VISIT\_END\_DATE is greater than the IP’s VISIT\_END\_DATE it does not change the IP VISIT\_END\_DATE.

For claim type in (‘emergency’,’urgentcare’) (ER)

Sort data in ascending order by PATIENT, START, END.

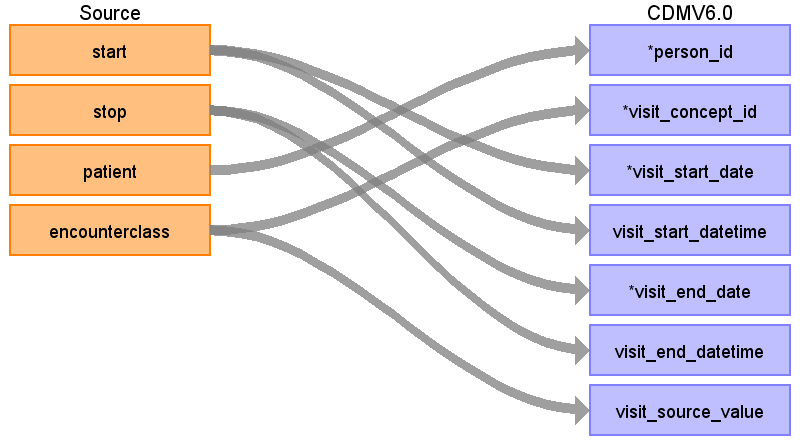
Then by PERSON\_ID, collapse all (ER) claims that start on the same day as one ER visit, then take START as VISIT\_START\_DATE, MAX (END) as VISIT\_END\_DATE, and ‘ER’ as PLACE\_OF\_SERVICE\_SOURCE\_VALUE.

For claim type in (‘ambulatory’, ‘wellness’, ‘outpatient’) (OP)

Sort data in ascending order by PATIENT, START, END.

Then by PERSON\_ID take START as VISIT\_START\_DATE, MAX(END) as VISIT\_END\_DATE, and ‘OP’ as PLACE\_OF\_SERVICE\_SOURCE\_VALUE.

Use logic found here: https://github.com/OHDSI/ETL-Synthea/blob/master/ETL/SQL/AllVisitTable.sql



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| visit\_occurrence\_id |  |  | Autogenerate |
| person\_id | patient | Map by mapping person.person\_source\_value to patient.  Find person.person\_id by mapping encouters.patient to person.person\_source\_value. |  |
| visit\_concept\_id | encounterclass | When encounterclass is 'emergency' or 'urgentcare' then set to 9203. When encounterclass is 'ambulatory', 'wellness', or 'outpatient' then set to 9202. When encounterclass is 'inpatient' then set to 9201. Otherwise set to 0. |  |
| visit\_start\_date | start |  |  |
| visit\_start\_datetime | start |  |  |
| visit\_end\_date | stop |  |  |
| visit\_end\_datetime | stop |  |  |
| visit\_type\_concept\_id |  |  | Set all records as concept\_id 44818517. |
| provider\_id |  |  |  |
| care\_site\_id |  |  |  |
| visit\_source\_value | encounterclass |  |  |
| visit\_source\_concept\_id |  |  |  |
| admitted\_from\_concept\_id |  |  |  |
| admitted\_from\_source\_value |  |  |  |
| discharge\_to\_concept\_id |  |  |  |
| discharge\_to\_source\_value |  |  |  |
| preceding\_visit\_occurrence\_id |  |  |  |

Table name: cohort

Table name: condition\_era

Table name: dose\_era

Table name: drug\_era

Table name: cost

Table name: payer\_plan\_period

Table name: care\_site

Table name: location

Table name: provider

Table name: cdm\_source

Table name: metadata

Table name: cohort\_definition

Table name: survey\_conduct

Table name: location\_history

Appendix: source tables

Table: allergies.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| start | date | 2008-03-24 |  |
| stop | date |  |  |
| patient | varchar | 99a94a01-bb91-4ad7-9a69-b9bd0f114d47 |  |
| encounter | varchar | c430e49d-9612-4ba0-92ec-70ab4afdcb5f |  |
| code | int | 419474003 |  |
| description | varchar | Allergy to mould |  |

Table: careplans.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| id | varchar | 906fceed-443f-4f59-acab-562f80c5cb44 |  |
| start | date | 2016-11-09 |  |
| stop | date |  |  |
| patient | varchar | b6750e05-91df-471b-97bb-a19ab9fe79c5 |  |
| encounter | varchar | 17781139-da60-46dc-ba6e-13637cb13c61 |  |
| code | int | 53950000 |  |
| description | varchar | Respiratory therapy |  |
| reasoncode | int | 10509002 |  |
| reasondescription | varchar | Acute bronchitis (disorder) |  |

Table: conditions.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| start | date | 2017-12-26 |  |
| stop | date |  |  |
| patient | varchar | ffab9299-0b3d-474f-96a4-bb279c244c7c |  |
| encounter | varchar | 50bbccc5-bb6b-4073-8244-045f10a28469 |  |
| code | int | 444814009 |  |
| description | varchar | Viral sinusitis (disorder) |  |

Table: encounters.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| id | varchar | 8ead2604-2f95-4ecb-a898-8c29f74709fb |  |
| start | varchar | 1983-04-24T05:36:43Z |  |
| stop | varchar | 2001-10-05T03:47:01Z |  |
| patient | varchar | 24121ee3-3413-4e42-8e17-0410204adf10 |  |
| encounterclass | varchar | ambulatory | case  when lower(encouterclass) = 'ambulatory' then 9202  when lower(encouterclass) = 'emergency' then 9203  when lower(encouterclass) = 'inpatient' then 9201  when lower(encouterclass) = 'wellness' then 9202  when lower(encouterclass) = 'urgentcare' then 9203  when lower(encouterclass) = 'outpatient' then 9202  else 0  end |
| code | int | 185349003 |  |
| description | varchar | Encounter for check up (procedure) |  |
| cost | real | 129.16 |  |
| reasoncode | int |  |  |
| reasondescription | varchar |  |  |

Table: imaging\_studies.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| id | varchar | 26cd0a95-10a6-4a30-8af1-ebc602d39b72 |  |
| date | date | 2015-10-26 |  |
| patient | varchar | 5dacca68-eb32-45ea-bf9d-f224f96e4299 |  |
| encounter | varchar | fcd7e8af-0242-4efb-a392-642ed4cb01d1 |  |
| bodysite\_code | int | 51185008 |  |
| bodysite\_description | varchar | Thoracic structure (body structure) |  |
| modality\_code | varchar | CT |  |
| modality\_description | varchar | Computed Tomography |  |
| sop\_code | varchar | 1.2.840.10008.5.1.4.1.1.2 |  |
| sop\_description | varchar | CT Image Storage |  |

Table: immunizations.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| date | date | 2016-06-16 |  |
| patient | varchar | 8734f7b8-4525-4db3-b8b7-60da22f729b5 |  |
| encounter | varchar | c9557a0c-7f03-48ba-9975-667d28f2dc39 |  |
| code | int | 140 |  |
| description | varchar | Influenza seasonal injectable preservative free |  |
| cost | real | 140.52 |  |

Table: medications.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| start | date | 2016-04-03 |  |
| stop | date |  |  |
| patient | varchar | 5dacca68-eb32-45ea-bf9d-f224f96e4299 |  |
| encounter | varchar | 93917d59-daf9-4206-b59c-c73262c4ca9f |  |
| code | int | 316672 |  |
| description | varchar | Simvistatin 10 MG |  |
| cost | real | 263.49 |  |
| dispenses | int | 1 |  |
| totalcost | real | 3161.88 |  |
| reasoncode | int |  |  |
| reasondescription | varchar |  |  |

Table: observations.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| date | date | 1986-06-20 |  |
| patient | varchar | cfe0b091-a7b4-4fbb-a0a2-d0fd86a596e4 |  |
| encounter | varchar | e84bdba6-e316-4c1f-a778-99a05f0dde1f |  |
| code | varchar | 72514-3 |  |
| description | varchar | Pain severity - 0-10 verbal numeric rating [Score] - Reported |  |
| value | varchar | Never smoker |  |
| units | varchar | mg/dL |  |
| type | varchar | numeric |  |

Table: patients.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| id | varchar | e31a194d-e8b1-4e81-a511-5da4ad99f8c8 |  |
| birthdate | date | 1926-02-23 |  |
| deathdate | date |  |  |
| ssn | varchar | 999-41-5589 |  |
| drivers | varchar |  |  |
| passport | varchar |  |  |
| prefix | varchar | Mr. |  |
| first | varchar | Benito209 |  |
| last | varchar | Marks830 |  |
| suffix | varchar |  |  |
| maiden | varchar |  |  |
| marital | varchar | M |  |
| race | varchar | white |  |
| ethnicity | varchar | irish |  |
| gender | varchar | M |  |
| birthplace | varchar | Boston |  |
| address | varchar | 192 MacGyver Dam |  |
| city | varchar | Boston |  |
| state | varchar | Massachusetts |  |
| zip | int | 02108 |  |

Table: procedures.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Type | Most freq. value | Comment |
| date | date | 2018-02-27 |  |
| patient | varchar | 24121ee3-3413-4e42-8e17-0410204adf10 |  |
| encounter | varchar | 7a9a4f7d-060c-4de9-b03d-0eee972ff7f1 |  |
| code | varchar | 428191000124101 |  |
| description | varchar | Documentation of current medications |  |
| cost | real | 516.65 |  |
| reasoncode | int | 72892002 |  |
| reasondescription | varchar | Normal pregnancy |  |