1. List of issues - Every time you experienced an issue where you had to make a decision affecting data in the OMOP CDM, and there was no clear guidance.

|  |  |  |
| --- | --- | --- |
| No | Table or column | issue |
| 1 | End\_date | If there is no exact end\_date, the date must be calculated. It is likely that the same calculation rules are needed between institutions. |
| 2 | Visit\_occurrence | When a previous visit, if the patient has scheduled a test or has been prescribed a medication, the patient may be provided with testing or medication services without a visit registration in the hospital. |
| 3 | Measurement | There are several dates associated with the test result information.  We needed an appropriate date to decide between Orddate(Date of prescription), Verifytm(Check result date and time), and Regtime(Date of registration). |
| 4 | Specimen | There are several dates related to the information of the test result of the sample. We needed an appropriate date to decide such as WORKDATE(Date of work), ORDDATE(Prescription date), SPCDATE( Date of specimen), COLLTIME(Date of blood collection), RECVERITIME(Recent sample results confirm date and time), VERIFYTM (Sample results confirm date and time), and RSLTTM (date of registration of the results). |
| 5 | Measurement | Sometimes the test result values are entered together with number and String. Ex) > 100, 100 less, +, P (52.93) and so on |
| 6 | Death | In the death certificate, there are three types of indirect, direct, and origin sign. We could not enter all three signs, so we had to decide the proper sign. |
| 7 | Person | In the case of a foreigner, the address information of the patient is entered as a temporary address (ex) hotel name).  Therefore, we need a way to enter the patient's location information. |
| 8 | Provider\_id | Several doctors are assigned to patients when they are prescribed or hospitalized. For example, when a patient is admitted to a hospital, the patient is assigned to Doctor of hospitalization, doctor in charge (CHADR), general doctor (GENDR), and specialist doctor (SPCDR). It was necessary to decide which of these doctors was right. |
| 9 | Note\_text | It is necessary to divide the contents of the clinical record into columns and input the entire data into the Note\_text column. |

1. Solution list. - All the conventions you created and might be worth standardizing.
2. End\_date calculation rules

|  |  |  |
| --- | --- | --- |
| Visit\_occurrence | | |
|  | Inpatient | Date of discharge |
|  | outpaient | Date of visit (Same as startdate) |
| Drug\_exposure | | |
|  | all | (Date of prescription+ Prescription days) -1day |
| Procedure\_occurrence | | |
|  | all | Date of treatment (Same as startdate) |
| Condition\_occurrence | | |
|  | Inpatient | \* If there is a discharge drug : (discharge prescription day + Prescription days) -1day  \* If there is no discharge drug : discharge prescription day |
|  | outpaient | \* If there is a prescription : prescription day + Prescription days) -1day  \* If there is no prescription : STARTDATE |
| Device\_exposure | | |
|  | all | Date of equipment use (Same as startdate) |

1. Visit\_occurrence\_id = null

* First, we extract the case where visit\_occurrence\_id is null in each table. Second, the person\_id and the date are distinct, and then a new row is created in the visit\_occurrence table (visit\_concept\_id =9202). Finally, we update the null value of visit\_occurrence\_id of each table with the newly generated id.

1. Measurement\_date

* We decided to enter in the measurement\_date the date that entered from the measuring instrument or the date that was collected sample.

1. Specimen\_date

* We decided to enter in the specimen\_date the date that the specimen was collected.

1. Measurement value preprocessing

* In the original data, extract the non-numeric data type of the result value(where isnumeric(result value) = 0). Since the types of values are variable, only the inequality, positive, negative, and titer results are classified. We created a new column in which a string and a number are to be input. Then numbers were converted to VALUE\_AS\_NUMBER columns and characters were mapped to OPERATOR\_CONCEPT\_ID or VALUE\_AS\_CONCEPT\_ID.

|  |  |
| --- | --- |
| **operator\_concept\_id** | |
| **CONCEPT\_ID** | **CONCEPT\_NAME** |
| 4172703 | = |
| 4172704 | > |
| 4171756 | < |
| 4171755 | >= |
| 4171754 | <= |
| **value\_as\_concept\_id** | |
| **CONCEPT\_ID** | **CONCEPT\_NAME** |
| 9191 | Positive |
| 9189 | Negative |

1. death\_type\_concept\_id

* When the patient died, the direct sign(immediate cause) must be filled out. it is an important sign. So, we decided to select the direct sign(immediate cause) from among the three signs.

1. location\_id in PERSON table

* Decide to enter location\_id as a temporary address. That is, we decided to map the location\_id to the hotel's zipcode.

1. Provider\_id in several table

* In the hospitalization history, we chose doctor in charge (CHADR). When prescription data is used, it is recommended to select the prescription doctor as the first choice, and if not, choose the doctor in charge.

1. Note\_text

* Columns are separated in XML format and combined into one sentence. After that, the name was anonymized. As a result of the anonymization of the name, only the last name was left. Ex) “Kim\*\*”.

1. Concept list - All the concepts you have picked to represent certain data where you had ambiguous or unclear choices.
2. Domain\_type\_concept\_id
   * Procedure\_type\_concept\_id = 38000275 (EHR order list entry)
   * Device\_type\_concept\_id = 44818707(EHR Detail)

* select \* from [dbo].[CONCEPT] where domain\_id = 'Type Concept' and vocabulary\_id = ‘Domain Type'
  + We use [computerized physician order entry](http://blog.daum.net/ontologicallearning/6942055) (CPOE) and Electronic Medical Record (EMR). But there is no corresponding concept list, so the most similar EHR is selected and mapped.

1. admitting\_source\_concept\_id, discharge\_to\_concept\_id
   * In case of outpatients, we will enter null because we can not know the route of the patient in detail. In case of inpatient, we can only check information on ambulance and emergency helicopter, and the discharge route includes information about home, other hospital transfer, and federally qualified health center to know. So we extracted the concept\_id list as shown in the table below.

