

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

OMOP Common Data Model (CDM) ETL Mapping Specification — Humana HSRC

Version 2.0
11-Dec-2009

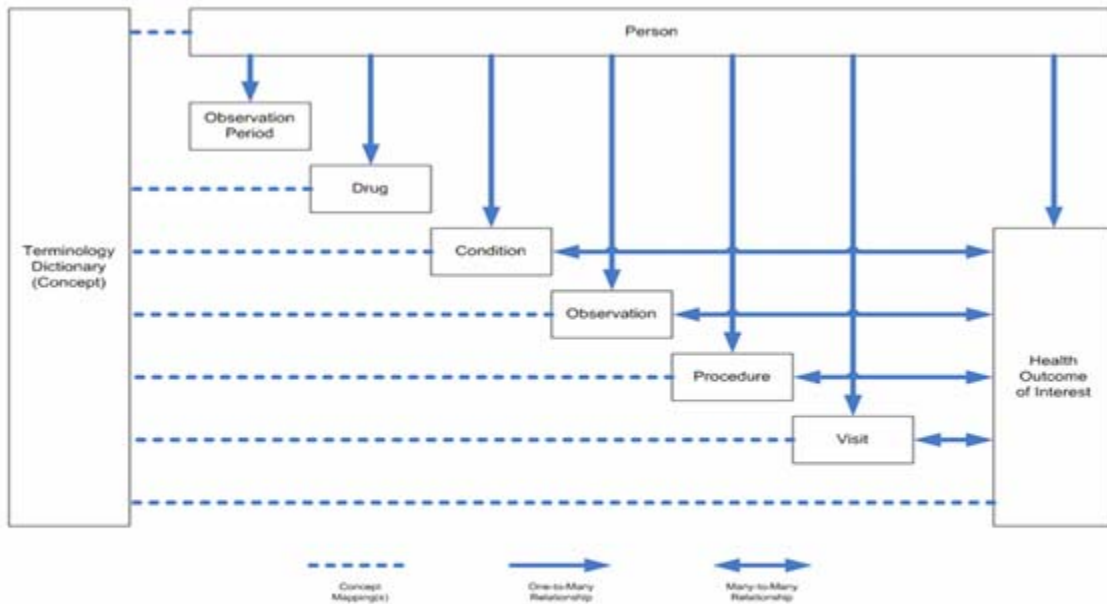


Table of Contents

1.0 Introduction	4
2.0 Source Data Mapping Approach	4
3.0 Source Data Mapping	5
3.1 Data Mapping	5
3.1.1 Table Name: PERSON	6
3.1.2 Table Name: DRUG_EXPOSURE	9
3.1.3 Table Name: CONDITION_OCCURRENCE	13
3.1.4 Table Name: VISIT_OCCURRENCE	18
3.1.5 Table Name: PROCEDURE_OCCURRENCE	20
3.1.6 Table Name: OBSERVATION	22
3.1.7 Table Name: OBSERVATION_PERIOD	24
3.2 Source Independent Data Mapping	26
3.2.1 Table Name: DRUG_ERA	26
3.2.2 Table Name: CONDITION_ERA	28
3.3 Reference Tables	30
3.3.1 Table Name: DRUG_EXPOSURE_REF	30
3.3.2 Table Name: CONDITION_OCCURRENCE_REF	30
3.3.3 Table Name: PROC_OCCURRENCE_REF	33
3.3.4 Table Name: OBSERVATION_TYPE_REF	35
3.3.5 Table Name: VOCABULARY_REF	35
3.3.6 Table Name: RELATIONSHIP_TYPE	36

Document Control

Authors and Contributors

Name	Organization	Title
Qianli Ma	Miami-Humana HSRC	Scientist

Reviewers

Name	Role	Title	Date Reviewed
Vinit Nair	Miami-Humana HSRC	Director	Dec 9 th 2009

Document References

Document Title	Type of Reference	Document Location
OMOP ETL Mapping Specification V2.0	Business Rules	OMOP Download Center
Standard Terminology Specifications V2.0	Vocabulary Terminology	OMOP Download Center

1.0 Introduction

This document reflects the requirements, assumptions, business rules and transformations for the implementation of the Common Data Model (CDM) as implemented by **Miami-Humana Health Services Research Center (HSRC)**

The purpose of this document is to describe the ETL mapping of the proprietary or licensed data from Humana Inc. into the OMOP Common Data Model.

It is based on the OMOP ETL Specifications. General information that is covered by the OMOP ETL Specification will not be covered in this document, but a detailed discussion of the Humana -specific aspects of mapping and converting data to the standard CDM is provided.

The document is composed of three 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 Era's.
- **Data Mapping Reference tables.**

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 will include both researchers that want to use the experience and learning in order to incorporate them into their own CDM construction.

2.0 Source Data Mapping Approach

The design follows the agreed upon general project assumptions:¹

- **Financial Information:** The CDM model makes no use of financial information such as Fees, Payments, Deductibles, Copayment, etc.
- **Plan Detail Information:** The model makes no use of any fields related to Plan or Coverage details such as Benefit Plan, Plan Indicator, etc. Of the administrative information in the claims data the model makes use of medical coverage period and eligibility for prescription drugs.

¹ Other, domain specific assumptions are documented in the respective sections

- Cleansing and Validation: The selected data fields will be handled (whether loaded directly or as part of a transformation) without any prior cleansing or validation.
- Data Privacy: Humana claims data is de-identified as a part of constructing the Common Data Model and is HIPAA compliant, thus allowing the design to use the unique person identifier as candidates for Primary keys without the risk of violating the person's private information.²
- Concept Identifiers: In the CDM, the main domains used for drug outcomes research are Drug, Condition, Observation, Procedure, Visit and Demographics. Data are represented through standard concept identifiers using a standardized terminology. During ETL, source data representations (raw data codes) will be translated to standard concept identifiers through a mapping process. If no standard concept identifier is available, the concept identifier field will contain '0' as a value.
- Indicators: Indicators are source tables containing additional information that were used for ETL rules in order to translate source definitions into concept identifiers. Some of the indicators will be retained in the CDM others are used only during ETL conversions.

3.0 Source Data Mapping

This section covers the high-level assumptions and approach to extraction, transformation and loading (ETL) of raw source data into the Common Data Model (CDM). The section covers each of the major tables in the CDM separately, elaborating the distinct handling required for each.

Unless otherwise specified, missing attributes will not disqualify data from being loaded into the Common Data Model. Missing attributes for Concept Identifiers will be populated with the value zero (0) in the CDM, while the rest of the missing attributes will be populated with NULL.

3.1 Data Mapping

A three year extract of administrative claims data (10/01/2006 -- 09/30/2009) residing on Humana SAS server was used to build the CDM. Members enrolled in Commercial HMO/PPO, Medicare Advantage HMO/PPO/PFFS and Medicare PDP were included in this study Claims data after 05/31/2009 was excluded due to claim lag.

² The model allows for 3rd party deployments outside of OMOP where the Primary key does refer to a Person as well as a field for encrypting such information in case the user needs to revert to source data for verification or further study.

Five tables were used to construct the CDM tables:

- MTH_MBR_COV_PLAN_PCP: Member coverage summary table
- CMS_MBR_PREM_FACT: Summary table for enrollment in medicare plans. This table contains race information for medicare enrollees.
- MED_CLAIMS: Medical claim data
- RX_CLAIMS: Pharmacy claim data
- LAB_CLAIMS: Laboratory claim data

3.1.1 TABLE NAME: PERSON

In the Humana database, person demographics data is recorded in MTH_MBR_COV_PLAN_PCP table. Race information is only available for Medicare enrollees and recorded in CMS_MBR_PREM_FACT table. Values for the individual source attributes are mapped to standard concept identifiers in the Dictionary and the corresponding concept identifiers which are stored as attribute values in the CDM PERSON table.

At least one record exists in the MTH_MBR_COV_PLAN_PCP table for each eligible person for each year for which data is loaded. Since the Humana data includes observations from many years, multiple records exist for each person. These were consolidated into one unique entry for each person including the demographic attributes. If the person demographics changes during the period of analysis, then the last known value was used for each demographic attribute.

Lessons learned, challenges and Humana modification:

- (1) Mapping for Location is incorrectly described in OMOP ETL mapping document:
OMOP: Extract TARGET_CONCEPT_ID for MAPPING_TYPE of 'LOCATION'
Humana: Extract TARGET_CONCEPT_ID for MAPPING_TYPE of 'ZIPCODE'
- (2) Additional cleaning for Humana MTH_MBR_COV_PLAN_PCP (**this step was also performed before constructing OBSERVATION_PERIOD table**):
MTH_MBR_COV_PLAN_PCP table also contains enrollment for dental, supplement and ASO plans, as well as records outside our study period (10/01/2006 -- 05/31/2009). These records were excluded before constructing PERSON table.
- (3) Derived and transformed fields: YEAR_OF_BIRTH, SOURCE_GENDER_CODE, SOURCE_LOCATION_CODE and SOURCE_RACE_CODE. See applied rules in the mapping table below.
- (4) Sorting and index: This table was sorted by PERSON_ID. We also created indexes on columns PERSON_ID, GENDER_CONCEPT_ID, RACE_CONCEPT_ID and LOCATION_CONCEPT_ID.

The field mapping is performed as follows:

Destination Field	Source Field	Applied Rule	Comment
PERSON_ID	[MTH_MBR_COV_PLAN_PCP].PERS_GEN_KEY		PERS_GEN_KEY is a unique system generated identifier with no real life significance as a person identifier.
YEAR_OF_BIRTH	[MTH_MBR_COV_PLAN_PCP].BIRTH_DATE	Keep only the birth year of [MTH_MBR_COV_PLAN_PCP].BIRTH_DATE	
GENDER_CONCEPT_ID	SOURCE_GENDER_CODE	<p>Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule:</p> <p>Extract TARGET_CONCEPT_ID for -MAPPING_TYPE of 'GENDER' SOURCE_VOCABULARY_CODE corresponding to 'GE' and SOURCE_CODE matching SOURCE_GENDER_CODE</p> <p>If no matching values found, store the value zero (0).</p>	
RACE_CONCEPT_ID	SOURCE_RACE_CODE	<p>Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule:</p> <p>Extract TARGET_CONCEPT_ID for -MAPPING_TYPE of 'RACE' SOURCE_VOCABULARY_CODE corresponding to 'GE' and SOURCE_CODE_DESCRIPTION matching SOURCE_RACE_CODE.</p> <p>Both fields need to be converted to UPPERCASE prior to matching.</p> <p>If no matching values found, store the value zero (0).</p>	

Destination Field	Source Field	Applied Rule	Comment
LOCATION_CONCEPT_ID	SOURCE_LOCATION_CODE	<p>Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule:</p> <p>Extract TARGET_CONCEPT_ID for -MAPPING_TYPE of 'ZIPCODE' SOURCE_VOCABULARY_CODE corresponding to 'GE' and SOURCE_CODE matching SOURCE_LOCATION_CODE</p> <p>If no matching values found, store the value zero (0).</p>	
SOURCE_PERSON_KEY		NULL	
SOURCE_GENDER_CODE	[MTH_MBR_COV_PLAN_PCP].SEX_CD	[MTH_MBR_COV_PLAN_PCP].SEX_CD values were recoded using the following rules: 'F' → 'FEMALE'; 'M' → 'MALE'; Missing or other → 'UNKNOWN/NOT STATED'.	
SOURCE_LOCATION_CODE	[MTH_MBR_COV_PLAN_PCP].ZIP_CD	Keep only the first three digits of [MTH_MBR_COV_PLAN_PCP].ZIP_CD	

Destination Field	Source Field	Applied Rule	Comment
SOURCE_RACE_CODE	[CMS_MBR_PREM_FACT].RACE_CD	[CMS_MBR_PREM_FACT].RACE_CD values were recoded using the following rule: '1'→ 'WHITE'; '2'→ 'BLACK'; '3'→ 'OTHER'; '4'→ 'ASIAN'; '5'→ 'HISPANIC'; '6'→ 'NATIVE AMERICAN'; '0' or missing → 'UNKNOWN'	

3.1.2 TABLE NAME: DRUG_EXPOSURE

Humana claims data carry two separate sources of Drug Exposure information. The primary source of the drug exposure data is filled prescription claims (RX_CLAIMS table). Additional Drug Exposure data are inferred based on the person procedures from medical claims data (MED_CLAIMS table). The mapped tables for data from these two sources were created separately and merged as the final Drug Exposure table. The final table was then sorted by PERSON_ID, DRUG_CONCEPT_ID and DRUG_EXPOSURE_START_DATE. We also created indexes on columns PERSON_ID and DRUG_CONCEPT_ID of the final table.

3.1.2.1 Transformation from Humana Prescriptions Data

Pharmacy claims carry an entry for each filled prescription and every subsequent refill.

Lessons learned, challenges and Humana modification:

- (1) Reversal claims: Reversal claims will introduce duplicate counts and negative days of supply. We use REVERSAL_IND (indicator of reversal claim: Y/N) and RX_COUNT (indicator of claim type: reversal /original claim) to remove reversal claims and only keep original claims.
- (2) Missing NDC: We deleted claim records with missing NDC which was recorded as blank or '0000000000'.

The field mapping is performed as follows:

Destination Field	Source Field	Applied Rule	Comment
-------------------	--------------	--------------	---------

Destination Field	Source Field	Applied Rule	Comment
DRUG_EXPOSURE_ID	[RX_CLAIMS].DOCUMENT_KEY	System generated unique identifier	
DRUG_EXPOSURE_START_DATE	[RX_CLAIMS].SERVICE_DATE		
DRUG_EXPOSURE_END_DATE		NULL	
PERSON_ID	[RX_CLAIMS].PERS_GEN_KEY		
DRUG_CONCEPT_ID	SOURCE_DRUG_CODE	[RX_CLAIMS].NDC_ID mapped to Drug Concept. Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule: Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'DRUG' SOURCE_VOCABULARY_CODE corresponding to 'NDC' and SOURCE_CODE matching SOURCE_DRUG_CODE If no matching values found, store the value zero (0).	
DRUG_EXPOSURE_TYPE		DRUG_EXPOSURE_TYPE is extracted from DRUG_EXPOSURE_REF For Records from Humana Pharmacy Claims data: DRUG_EXPOSURE_TYPE matching DRUG_EXPOSURE_TYPE_DESC of 'Prescription Dispensed' If no matching values found, store the value zero (0).	Defined type identifier based on the source from which the Drug Exposure was recorded.
STOP_REASON		NULL	
REFILLS		NULL	
DRUG_QUANTITY	[RX_CLAIMS]		

Destination Field	Source Field	Applied Rule	Comment
NTITY].PAYABLE_QTY		
DAYS_SUPPLY	[RX_CLAIMS].PAY_DAY_SUPPLY_CNT		
SOURCE_DRUG_CODE	[RX_CLAIMS].NDC_ID		

3.1.2.2 Transformation from Humana Procedure Data

Drug Procedure data were stored in MED_CLAIMS table and primarily captured in the form of procedure codes. They include:

- HCPCS Level II codes (i.e., J and Q series codes) that capture medication information.
- CPT-4 codes that record the administration and management of medications.
- ICD-9-CM procedure codes that relate to the administration and management of medications.

MED_CLAIMS table has 7 columns for procedures codes: 6 columns for ICD-9-CM procedure codes (ICD9_PROC_CD1 -- ICD9_PROC_CD6) and 1 column for CPT/HCPCS/RV code (SERV_TYPE_CD).

Lessons learned, challenges and Humana modification:

- (1) Additional cleaning for MED_CLAIMS table (this step was also performed before constructing **CONDITION_OCCURRENCE**, **VISIT_OCCURRENCE**, and **PROCEDURE_OCCURRENCE** tables): Claim records in Humana database are from billing process, paid, denied and reversal claims are all recorded. To reduce duplicate counts, we cleaned medical claims using PYMT_CAT_CD, and only claims with PYMT_CAT_CD in ('A', 'C', 'E') were kept. A = ADJUDICATED, C = PAID/COMPLETED, E = ENCOUNTER (PREPAID).
- (2) Duplicates within valid claim: Even after data cleaning aforementioned, duplicated records still exist. For example, a claim (identified by CLM_GEN_KEY) may contain 3 records for insulin injection, record 1 and 2 have the same information, while record 3 has negative amount of payment. By aggregating the payment, only one service for insulin injection is identified. However, using financial information is out of scope for OMOP project. To prevent the overestimate of procedure drug usage, we used the unique combination of PERS_GEN_KEY, CLM_GEN_KEY, SERV_FROM_DATE, POT_CD and procedure drug code to define a service of procedure drug.

- (3) Non-standard format of SOURCE_CODE in SOURCE_TO_CONCEPT_MAP: 00.13 was coded as 0.13. Our transformation of procedure code captured this format.
- (4) Derived and transformed fields: DRUG_EXPOSURE_TYPE and SOURCE_DRUG_CODE. See applied rules in the mapping table below.

The field mapping used to load Drug Exposure table is as follows. Only procedures which map to a drug concept are extracted:

Destination Field	Source Field	Applied Rule	Comment
DRUG_EXP OSURE_ID	[MED_CLAIMS].MEDCLM_KEY	Assign [MED_CLAIMS].MEDCLM_KEY*10+ relative position.	
DRUG_EXP OSURE_STA RT_DATE	[MED_CLAIMS].SERV_F ROM_DATE		
DRUG_EXP OSURE_EN D_DATE		NULL	
PERSON_ID	[MED_CLAIMS].PERS_ GEN_KEY		
DRUG_CON CEPT_ID	SOURCE_D RUG_CODE	Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule: Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'PROCEDURE DRUG' SOURCE_CODE matching SOURCE_DRUG_CODE	
DRUG_EXP OSURE_TYP E	[MED_CLAIMS].POT_C D	DRUG_EXPOSURE_TYPE is extracted from DRUG_EXPOSURE_RE: IF [MED_CLAIMS].POT_CD in ('21', '27','28','51','61'), DRUG_EXPOSURE_TYPE matching DRUG_EXPOSURE_TYPE_DESC of 'Inpatient Administration'; Else DRUG_EXPOSURE_TYPE matching DRUG_EXPOSURE_TYPE_DESC of 'Physician Administered Drug'.	

Destination Field	Source Field	Applied Rule	Comment
STOP_REASON		NULL	
REFILLS		NULL	
DRUG_QUANTITY		NULL	
DAYS_SUPPLY		NULL	
SOURCE_DRUG_CODE	[MED_CLAIMS].SERV_TYPE_CD [MED_CLAIMS].ICD9_PROC_CD1 - ICD9_PROC_CD6	[MED_CLAIMS]. ICD9_PROC_CD1 - ICD9_PROC_CD6 were recoded using the following rule: If Length of Code > 2 then SOURCE_DRUG_CODE =compress (substr(Code,1,2) ' ' substr(Code,3)). If substr (SOURCE_DRUG_CODE, 1,2) = '00' then substr (SOURCE_DRUG_CODE, 1,2) ='0'. [MED_CLAIMS]. SERV_TYPE_CD was used to exact CPT/HCPCS code using the following rule: if SERV_CD_TYPE_CD1 in ('CP','HC') or SERV_CD_TYPE_CD2 in ('CP','HC') then SOURCE_DRUG_CODE = trim(left(uppercase(SERV_TYPE_CD)));	

3.1.3 TABLE NAME: CONDITION_OCCURRENCE

Conditions data are recorded as ICD-9-CM diagnosis codes in MED_CLAIMS table. In addition, the person mortality is recorded based on the person discharge status in MED_CLAIMS table.

The step by step approach would be as follows:

Step 1: MED_CLAIMS table was scanned and all ICD-9-CM diagnosis codes are extracted. The diagnosis codes extracted are PRIMARY_DIAG_CD (Principal Diagnosis), DIAG_CD2, DIAG_CD3, DIAG_CD4, DIAG_CD5, DIAG_CD6, DIAG_CD7, DIAG_CD8 and DIAG_CD9. The ICD-9-CM code, the relative positioning of the code within the claim is noted (e.g.: Principal, Diagnosis1, Diagnosis 2, etc.)

Step 2: Dates associated with the ICD-9-CM diagnosis code are extracted as the Condition Start date in each case. The following attributes are not available for Humana claims data: Condition End Date, Stop Reason and Diagnosis (DX) Qualifier.

Step 3: A Condition Occurrence type is assigned to each of the extracted ICD-9-CM diagnosis codes based on place of service (POT_CD) and the relative positioning of the code within the claim. The Condition Occurrence type values and the associated descriptions are listed in the section for [CONDITION OCCURRENCE REF.](#)

Step 4: Each of the ICD-9-CM diagnosis codes is now mapped to a standard condition concept. The standard condition concept identifier related to the ICD-9-CM diagnosis code is determined by referencing the ICD-9-CM condition mapping table.

Step 5: Discharge status code attribute from MED_CLAIMS are now scanned to check for any record for Person Mortality. The presence of the following range of Discharge Status codes are taken as indicators of Person Mortality:

- Discharge status codes 20 to 29
- Discharge status codes 40 to 42

The date of the discharge (if missing, use SERV_END_DATE) is recorded as the Condition Start date and the event is assigned a Condition Occurrence type by scanning the CONDITION_OCCURRENCE_REF table for a description matching 'Death at Discharge'.

Lessons learned, challenges and Humana modification:

- (1) Additional cleaning for MED_CLAIMS table: Refer to section 3.1.2.2
- (2) Duplication of records: In Humana claim database, claims with multiple records carried same diagnosis for each record. Also a provider might submit multiple claims for a single visit, and these claims might have the same diagnosis codes. Same issue could happen to discharge status code. We used the unique combination of PERS_GEN_KEY, ICD9 diagnosis code, the relative positioning of the code, SERV_FROM_DATE and POT_CD to define a condition occurrence. For person mortality, only one record was kept for each patient.
- (3) Duplicated mapping for ICD9 diagnosis V code in SOURCE_TO_CONCEPT_MAP table: 33 V codes had ≥ 2 records. We kept only 1 record for each of them (see appended excel file for deleted records)
- (4) Mapping for SOURCE_CONDITION_CODE is incorrectly described in OMOP ETL mapping document:
 OMOP: Extract TARGET_CONCEPT_ID for MAPPING_TYPE of 'CONDITION' SOURCE_VOCABULARY_CODE corresponding to:
 - For [FACILITY HEADER].DSTATUS, 'Discharge Status'
 - For [INPATIENT ADMISSION].DSTATUS, 'Discharge Status'
 - For all others, 'ICD9 CM'
 Humana: (1) ICD-9-CM Diagnosis Codes:

Extract TARGET_CONCEPT_ID for MAPPING_TYPE of
 'CONDITION'
 SOURCE_VOCABULARY_CODE corresponding to 'ICD-9-CM'
 SOURCE_CODE matching SOURCE_CONDITION_CODE

(2) Discharge Status:

Extract TARGET_CONCEPT_ID for -MAPPING_TYPE of
 'DISCHARGE STATUS'
 SOURCE_CODE matching SOURCE_CONDITION_CODE

(5) Derived and transformed fields: CONDITION_OCCURRENCE_TYPE,
 SOURCE_CONDITION_CODE and CONDITION_START_DATE. See applied
 rules in the mapping table below.

(6) Sorting and Index: This table was sorted by PERSON_ID,
 CONDITION_CONCEPT_ID and CONDITION_START_DATE. We also created
 indexes on columns PERSON_ID and CONDITION_CONCEPT_ID.

The field mapping is performed as follows:

Destination Field	Source Field	Applied Rule	Comment
CONDITION_OCCURRENCE_ID	[MED_CLAIMS].MEDCLM_KEY	For the ICD-9-CM diagnosis codes, assign [MED_CLAIMS].MEDCLM_KEY*10+ relative position. For discharge status code, assign [MED_CLAIMS].MEDCLM_KEY	
CONDITION_START_DATE	[MED_CLAIMS].SERV_FROM_DATE [MED_CLAIMS].SRC_DISCHARGE_DATE [MED_CLAIMS].SERV_END_DATE	For the ICD-9-CM diagnosis codes, assign, [MED_CLAIMS].SERV_FROM_DATE For discharge status code, if [MED_CLAIMS].SRC_DISCHARGE_DATE is not missing, use it; Else use [MED_CLAIMS].SERV_END_DATE	
PERSON_ID	[MED_CLAIMS].PERS_GEN_KEY		
CONDITION_END_DATE		NULL	

Destination Field	Source Field	Applied Rule	Comment
CONDITION_OCCURRENCE_TYPE	CONDITION_OCCURRENCE_REF	<p>The type code is read from the CONDITION_OCCURRENCE_REF for a description matching:</p> <p>(1) Define Condition Occurrence Type Description:</p> <p>IF [MED_CLAIMS].POT_CD in ('21', '27', '28', '51', '61'), set as ' Inpatient Header ';</p> <p>Else set as 'Outpatient Header'</p> <p>(2) Define Condition Occurrence Position:</p> <p>'Primary' for PRIMARY_DIAG_CD</p> <p>1' for DIAG_CD2,</p> <p>'8' for DIAG_CD9 etc</p> <p>Death at Discharge</p>	

Destination Field	Source Field	Applied Rule	Comment
CONDITION_CONCEPT_ID	SOURCE_CONDITION_CODE	<p>Standardized concept, mapped from ICD-9-CM Diagnosis Codes or Discharge status codes. Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule:</p> <p>(1) ICD-9-CM Diagnosis Codes:</p> <p>Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'CONDITION'</p> <p>SOURCE_VOCABULARY_CODE corresponding to 'ICD-9-CM'</p> <p>SOURCE_CODE matching SOURCE_CONDITION_CODE</p> <p>(2) Discharge Status:</p> <p>Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'DISCHARGE STATUS'</p> <p>SOURCE_CODE matching SOURCE_CONDITION_CODE</p> <p>If no matching values found, store the value zero (0).</p>	
STOP_REASON		NULL	
DX_QUALIFIER		NULL	

Destination Field	Source Field	Applied Rule	Comment
SOURCE_CONDITION_CODE	[MED_CLAIMS]. PRIMARY_DIAG_CD, DIAG_CD2 - DAG_CD9 [MED_CLAIMS]. DISCHG_STAT_CD	[MED_CLAIMS]. PRIMARY_DIAG_CD, DIAG_CD2 - DAG_CD9 are recoded using the following rule: If substr(code,1,1)='E' and length(code) >4, SOURCE_CONDITION_CODE = compress(substr(code,1,4) ' ' substr(code,5)); Else If substr(code,1,1) ne 'E' and length(code) >3, SOURCE_CONDITION_CODE = compress(substr(code,1,3) ' ' substr(code,4))	The logic will traverse each diagnostic value and create a separate record for each non NULL diagnosis code. DSTATUS is added as a separate record if it is one of the following codes 20-29: Died 40-42: Other died status Each of these values is extracted as a separate record.

3.1.4 TABLE NAME: VISIT_OCCURRENCE

Visit occurrences are captured in MED_CLAIMS table. The place of treatment code (POT_CD) is used to determine the visit concept identifier from the source to concept mapping data in the Dictionary.

Lessons learned, challenges and Humana modification:

- (1) Additional cleaning for MED_CLAIMS table: Refer to section 3.1.2.2
- (2) In Humana claim database, inpatient admissions and outpatient services were aggregated into MED_CLAIMS table. We used the following rule to define a unique visit:
If the place of treatment code (POT_CD) was in ('21', '27','28', '51','61'), this was considered as inpatient admission and we used the unique combination of PERS_GEN_KEY, POT_CD, SERV_FROM_DATE and SERV_END_DATE to define a visit occurrence. Otherwise, we use the unique combination of PERS_GEN_KEY, POT_CD, SERV_FROM_DATE, SERV_END_DATE and PROV_TAX_ID to define a visit occurrence
- (3) Derived and transformed field: SOURCE_VISIT_CODE. See applied rules in the mapping table below.
- (4) Mapping for SOURCE_VISIT_CODE is incorrectly described in OMOP ETL mapping document:
OMOP: Extract TARGET_CONCEPT_ID for -MAPPING_TYPE of 'VISIT' SOURCE_VOCABULARY_CODE corresponding to 'PLACE OF SERVICE'
Humana: Extract TARGET_CONCEPT_ID for -MAPPING_TYPE of 'PLACE OF SERVICE' and source_vocabulary_code = '52'

(5) Sorting and index: This table was sorted by PERSON_ID and VISIT_START_DATE. We also created indexes on columns PERSON_ID and VISIT_CONCEPT_ID.

The field mapping is performed as follows:

Destination Field	Source Field	Applied Rule	Comment
VISIT_OCCURRENCE_ID	[MED_CLAIMS].MEDCLM_KEY	System generated unique identifier	
VISIT_START_DATE	[MED_CLAIMS].SERVICE_FROM_DATE		Start date for the Person Visit.
VISIT_END_DATE	[MED_CLAIMS].SERVICE_END_DATE		End date for the Person Visit.
PERSON_ID	[MED_CLAIMS].PERSON_KEY		
VISIT_CONCEPT_ID	SOURCE_VISIT_CODE	[MED_CLAIMS].POT_CD is mapped to the Visit concept Id. Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule: Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'PLACE OF SERVICE' SOURCE_CODE matching SOURCE_VISIT_CODE If no matching values found, store the value zero (0).	
SOURCE_VISIT_CODE	[MED_CLAIMS].POT_CD	The leading zero of POT_CD will be removed.	The place of Service field.

3.1.5 TABLE NAME: PROCEDURE_OCCURRENCE

Procedure codes are captured in MED_CLAIMS table. This table carries CPT-4, HCPCS and ICD-9-CM procedure codes, and each record in this table contains 6 columns for ICD-9-CM procedure codes (ICD9_PROC_CD1 -- ICD9_PROC_CD6) and 1 column for CPT/HCPCS/RV/HG code (SERV_TYPE_CD). Revenue code (RV) and Home growth code (HG) were excluded.

The following steps are outlined to ensure that all procedure occurrences are recorded:

- The place of treatment code (POT_CD) is used to determine the Procedure Occurrence Type Description:
 IF POT_CD in ('21', '27','28', '51','61'), set as 'Inpatient Header';
 Else set as 'Outpatient Header'

- Define Procedure Occurrence Position:
 '1' for ICD9_PROC_CD1,

 '6' for ICD9_PROC_CD6,
 '7' for SERV_TYPE_CD

Lessons learned, challenges and Humana modification:

- (1) Additional cleaning for MED_CLAIMS table: Refer to section 3.1.2.2
- (2) Duplicates within valid claim: Refer to section 3.1.2.2
- (3) Duplicated procedure codes in SOURCE_TO_CONCEPT_MAP table: For example, 52 can be both ICD9 procedure and CPT codes.

This was fixed by assigning Ptype for both SOURCE_CODE (dictionary table) and SOURCE_PROCEDURE_CODE (medical claim) using the following rule: 'ICD-9-Procedure' → Ptype = '1', 'CPT-4','HCPCS' → Ptype = '2'. Matching Concept ID for procedure code is extracted from the SOURCE_TO_CONCEPT_MAP table only when SOURCE_CODE matches SOURCE_PROCEDURE_CODE and their Ptype matches.

- (4) Derived and transformed fields: SOURCE_PROCEDURE_CODE and PROCEDURE_OCCURRENCE_TYPE. See applied rules in the mapping table below.
- (5) Sorting and index: This table was sorted by PERSON_ID, PROCEDURE_CONCEPT_ID and PROCEDURE_DATE. We also created indexes on columns PERSON_ID and PROCEDURE_CONCEPT_ID.

The field mapping is performed as follows:

Destination Field	Source Field	Applied Rule	Comment

Destination Field	Source Field	Applied Rule	Comment
PROCEDURE_OCCURRENCE_ID	[MED_CLAIMS].MEDCLM_KEY	Assign [MED_CLAIMS].MEDCLM_KEY*10+ relative position.	
PROCEDURE_DATE	[MED_CLAIMS].SERVICE_END_DATE		
PERSON_ID	[MED_CLAIMS].PERSON_KEY		
PROCEDURE_CONCEPT_ID	SOURCE_PROCEDURE_CODE	<p>Defined procedure type using the following rule for both SOURCE_CODE and SOURCE_PROCEDURE_CODE: 'ICD-9-Procedure' → Ptype = '1', 'CPT-4','HCPCS' → Ptype = '2'.</p> <p>Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule:</p> <p>Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'PROCEDURE'</p> <p>SOURCE_VOCABULARY_CODE corresponding to 'ICD-9-Procedure','CPT-4','HCPCS'</p> <p>SOURCE_CODE matching SOURCE_PROCEDURE_CODE</p> <p>Ptype defined above for SOURCE_CODE and SOURCE_PROCEDURE matches</p> <p>If no matching values found, store the value zero (0).</p>	
SOURCE_PROCEDURE_CODE	[MED_CLAIMS]. for ICD9_PROC_CD1- for ICD9_PROC_CD6 [MED_CLAIMS].	<p>Only Non NULL values will be included.</p> <p>[MED_CLAIMS]. ICD9_PROC_CD1 - ICD9_PROC_CD6 are recoded using the following rule:</p> <p>If Length of Code > 2 then SOURCE_DRUG_CODE =compress(substr(Code,1,2) ' ' substr(Co</p>	Each of the values is extracted as a separate record.

Destination Field	Source Field	Applied Rule	Comment
	SERV_TYPE_CD	de,3)). [MED_CLAIMS]. SERV_TYPE_CD is extracted using the following rule: If [MED_CLAIMS]. SERV_CD_TYPE_CD1 in ('CP','HC') or [MED_CLAIMS]. SERV_CD_TYPE_CD2 in ('CP','HC'), then keep [MED_CLAIMS]. SERV_TYPE_CD.	
PROCEDURE_OCCURRENCE_TYPE	PROCEDURE_OCCURRENCE_REF	The type code is extracted from the PROCEDURE_OCCURRENCE_REF for a description matching the Procedure Occurrence Type Description and Procedure Occurrence Position	Type code from which the Procedure Occurrence was drawn or inferred, and indicates whether a Procedure was primary or secondary and the relative positioning within a person Procedure record

3.1.6 TABLE NAME: OBSERVATION

Observation data from Humana claims data only includes Lab Observations data recorded in the LAB_CLAIMS table. The Lab Observations data include character values recorded in the LAB_RESULTS attribute and the units of measure stored in the RESULTS_UNITS attribute. Lab Observations are tracked using standard LOINC codes which are mapped to standard concept identifiers. Only those observations containing standard codes are included in the extraction. Units of measure are also mapped to standard concept identifiers in the dictionary for standardization

Lesson learned, challenges and Humana modification:

- (1) Mapping to Units of Measure:
We fixed incorrect specifications for UOM mapping in SOURCE_TO_CONCEPT_MAP table. In addition, we added 200 new records for mapping UOM based on observations from Humana lab claim data. (See appended excel file for Humana modification)
- (2) Derived and transformed fields: OBS_TYPE, OBS_VALUE_AS_NUMBER, OBS_VALUE_AS_STRING, OBS_RANGE_LOW and OBS_RANGE_HIGH. See applied rules the mapping table below.
- (3) Sorting and index: This table was sorted by PERSON_ID, OBS_CONCEPT_ID and OBS_DATE. We also created indexes on columns PERSON_ID, OBS_CONCEPT_ID and OBS_UNITS_CONCEPT_ID.

The field mapping is performed as follows:

Destination Field	Source Field	Applied Rule	Comment
-------------------	--------------	--------------	---------

Destination Field	Source Field	Applied Rule	Comment
OBS_OCCURRENCE_ID	[LAB_CLAIMS].LAB_PAT_CNTL_ID [LAB_CLAIMS].LAB_ORIG_SEQ_NBR	Combine the values of [LAB_CLAIMS].LAB_PAT_CNTL_ID and [LAB_CLAIMS].LAB_ORIG_SEQ_NBR with a _ in between.	
PERSON_ID	[LAB_CLAIMS].PERS_GEN_KEY		
SOURCE_OBS_CODE	[LAB_CLAIMS].LOINC_CD		
OBS_CONCEPT_ID	SOURCE_OBS_CODE	Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule: Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'OBSERVATION' SOURCE_VOCABULARY_CODE corresponding to 'LOINC' and SOURCE_CODE matching SOURCE_OBS_CODE If no matching values found, store the value zero (0).	
OBS_VALUE_AS_NUMBER	[LAB_CLAIMS].LAB_RESULTS	If [LAB_CLAIMS].LAB_RESULTS is numeric, use [LAB_CLAIMS].LAB_RESULTS. Otherwise, set as missing.	
OBS_DATE	[LAB_CLAIMS].COLLECTION_DATE		
OBS_RANGE_LOW	[LAB_CLAIMS].NORMALS_DEC_LOW	If [LAB_CLAIMS].NORMALS_DEC_LOW = [LAB_CLAIMS].NORMALS_DEC_HIGH or Observation results are text, set as missing.	
OBS_RANGE_HIGH	[LAB_CLAIMS].NORMALS_DEC_HIGH	If [LAB_CLAIMS].NORMALS_DEC_LOW = [LAB_CLAIMS].NORMALS_DEC_HIGH or Observation results are text, set as missing.	
OBS_TYPE		If OBS_VALUE_AS_NUMBER is not missing, set as 'LON'; otherwise set as 'LOT'	

Destination Field	Source Field	Applied Rule	Comment
OBS_VALUE_AS_STRING	[LAB_CLAIMS].LAB_RESULTS	If [LAB_CLAIMS].LAB_RESULTS is not numeric, use [LAB_CLAIMS].LAB_RESULTS. Otherwise, set as missing.	
OBS_VALUE_AS_CONCEPT_ID		NULL	
OBS_UNITS_CONCEPT_ID	[LAB_CLAIMS].RESULTS_UNITS	[LAB_CLAIMS].RESULTS_UNITS mapped to a standard concept identifier using a modified mapping table: 140 new entries added based on Humana LAB_CLAIMS. Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP_HUM table using the following rule: Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'UOM' SOURCE_VOCABULARY_CODE corresponding to 'THOMSON' and 'Humana' and SOURCE_CODE matching [LAB_CLAIMS].RESULTS_UNITS If no matching values found, store the value zero (0).	

3.1.7 TABLE NAME: OBSERVATION_PERIOD

Patient Status during an Observation Period in the Humana claims data is available from enrollment summary table MTH_MBR_COV_PLAN_PCP. This table includes records that indicate a person’s enrollment for each month for the period covered by the claims data.

Enrollment entries are consolidated by combining records that indicate continuous enrollment over a period of time with no change in **both drug and medical coverage** into a single entry with start and end dates to cover the period. If enrollment data indicates a person’s coverage did not extend through the entire duration covered by claims data, then multiple Observation Periods are recorded to capture all periods of coverage with corresponding start and end dates.

The consolidation is done through the following steps:

- Records for each person are sorted in an ascending order of COV_EFF_DATE (Start Date).

- Periods of continuous enrollment and no change in **both drug and medical coverage** status are consolidated by combining records and recording the Start Date (COV_EFF_DATE) for the first period as the Observation start date and end date (COV_END_DATE) for the last period as the Observation end date. Status is set to 'Active' for the duration of continuous enrollment.
- If the **drug coverage or medical coverage** indicator changed during the period of continuous enrollment, a new observation period is added with status set to 'Active' and the drug coverage or medical coverage indicator set to reflect the change.
- No separate records are added to cover gaps in coverage.
- Keep coverage period between 10/01/2006 and 05/31/2009 using the following steps:
 - (1) If Start Date < 10/01/2006, set 10/01/2006 as Start Date
 - (2) If End Date > DECSD_DATE, set DECSD_DATE as End Date
 - (3) If End Date > 05/31/2009, set 05/31/2009 as End Date.
 - (4) Remove any record with Start Date > End Date
- PERSON_STATUS is set as 'ACTIVE' for each record first. If member was recorded with DECSD_DATE, End Date will be set as DECSD_DATE -1 for the latest record. One separate record with both Start Date and End date equal to DECSD_DATE will be added, and PERSON_STATUS is set as 'DECEASED'.

Lessons learned, challenges and Humana modification:

- (1) Additional cleaning for MTH_MBR_COV_PLAN_PCP table: Refer to section 3.1.1
- (2) Derived and transformed fields: DX_DATA_AVAILABILITY and HOSPITAL_DATA_AVAILABILITY. See applied rules in the mapping table below.
- (3) Sorting and index: This table was sorted by PERSON_ID, OBSERVATION_START_DATE and OBSERVATION_END_DATE. We also created indexes on columns PERSON_ID, RX_DATA_AVAILABILITY and DX_DATA_AVAILABILITY.

The field mapping is as follows:

Destination Field	Source Field	Applied Rule	Comment
OBSERVATION_PERIOD_ID		System generated unique identifier	
OBSERVATION_PERIOD_START_DATE	[MTH_MBR_COV_PLAN_PCP].COV_EFF_DATE	Extracted as: Maximum value of [MTH_MBR_COV_PLAN_PCP].COV_EFF_DATE and '01OCT2006'	'01OCT2006' is the start date of Humana data.
OBSERVATION_PERIOD_END	[MTH_MBR_COV_PLAN_PCP].COV_END_DATE	Extracted as: Minimum value of [MTH_MBR_COV_PLAN_PCP].COV_EFF_DATE and '31MAY2009' and	'31MAY2009' is the end date of

Destination Field	Source Field	Applied Rule	Comment
_DATE	DATE	[MTH_MBR_COV_PLAN_PCP].DECSD_DATE	Humana data
PERSON_ID	[MTH_MBR_COV_PLAN_PCP].PERS_GEN_KEY		
PERSON_STATUS_CONCEPT_ID		Matching Concept ID extracted from the SOURCE_TO_CONCEPT_MAP table using the following rule: Extract TARGET_CONCEPT_ID for - MAPPING_TYPE of 'PATIENT STATUS' SOURCE_VOCABULARY_CODE corresponding to 'GE' and SOURCE_CODE_DESCRIPTION matching PATIENT_STATUS	Successive entries for a given person status are combined into one observation period with the start and end dates defined by both.
RX_DATA_AVAILABILITY	[MTH_MBR_COV_PLAN_PCP].RX_ELIGIBLE_IND		
DX_DATA_AVAILABILITY	[MTH_MBR_COV_PLAN_PCP].FIN_SUB_CD	If [MTH_MBR_COV_PLAN_PCP].FIN_SUB_CD in ('MDG','MDI') then DX_DATA_AVAILABILITY = 'N'; else 'Y'	
HOSPITAL_DATA_AVAILABILITY	[MTH_MBR_COV_PLAN_PCP].FIN_SUB_CD	If [MTH_MBR_COV_PLAN_PCP].FIN_SUB_CD in ('MDG','MDI') then DX_DATA_AVAILABILITY = 'N'; else 'Y'	
CONFIDENCE		NULL	

3.2 Source Independent Data Mapping

Unless otherwise specified in the sections below, Source Independent Data Mapping will follow specifications as defined in ETL Mapping Specification document (V2.0)

3.2.1 TABLE NAME: DRUG_ERA

The DRUG_ERA table is constructed through aggregation of individual drug exposures from the DRUG_EXPOSURE table. We followed the instructions in **OMOP ETL Mapping (V2.0)** for constructing Drug Eras.

Lessons learned, challenges and Humana modification:

- (1) Records without mapping DRUG_CONCEPT_ID: Since DRUG_CONCEPT_ID was set as 0 in DRUG_EXPOSURE table for these records, it's inappropriate to aggregate them. We excluded these records before constructing DRUG_ERA table.
- (2) Creating intermediate mapping table: Instead of using nested SQL queries provided by OMOP, we used in-line view (much faster).

```

proc sql;
  create table DATAOMOP.concept_drug_int as
  select a.ancestor_concept_id,a.descendant_concept_id
  from DATAOMOP.concept_ancestor a, DATAOMOP.concept b
  where b.concept_vocabulary_code = '08'
        and b.concept_level = 2
  and b.concept_id = a.ancestor_concept_id
  and a.descendant_concept_id in (select c.concept_id
                                  from DATAOMOP.concept c
                                  where c.concept_level = 1
                                  and c.concept_vocabulary_code IN
                                       '08', '54'))

  Union

  select c.concept_id as ancestor_concept_id,c.concept_id as
  descendant_concept_id
  from DATAOMOP.concept c
  where c.concept_vocabulary_code = '08'
        and c.concept_level = 2
        and c.CONCEPT_ID not in (select d.CONCEPT_ID_2
                                   from DATAOMOP.concept_relationship d
                                   where d.relationship_type = '003');

quit;

```

- (3) Sorting and Index: This table was sorted by PERSON_ID, DRUG_EXPOSURE_TYPE, DRUG_CONCEPT_ID and DRUG_ERA_START_DATE. We created indexes on columns PERSON_ID, DRUG_CONCEPT_ID and DRUG_EXPOSURE_TYPE.

All Drug Eras are recorded in the DRUG_ERA table based on the following field mapping:

Destination Field	Source Field	Applied Rule	Comment
DRUG_ERA_ID		System generated unique identifier	
DRUG_ERA_START_DATE	START_DATE	Extracted as the Minimum START_DATE among all the Drug exposures included in the Era	
DRUG_ERA_END_DATE	END_DATE	Extracted as the Maximum END_DATE among all the Drug exposures included in the Era	
PERSON_ID	DRUG_EXPOSURE.PERSON		

Destination Field	Source Field	Applied Rule	Comment
	_ID		
DRUG_EXPOSURE_TYPE	DRUG_EXPOSURE_TYPE	<p>The following settings apply:</p> <p>For Eras created with 30 day persistence window setting:</p> <p>DRUG_EXPOSURE_TYPE matching DRUG_EXPOSURE_TYPE_DESC of 'Drug Era – 30 day window'</p> <p>For Eras created with 0 day persistence window setting:</p> <p>DRUG_EXPOSURE_TYPE matching DRUG_EXPOSURE_TYPE_DESC of 'Drug Era – 0 day window'</p>	Drug Exposure Type for the Era created based on the persistence window setting.
DRUG_CONCEPT_ID	DRUG_EXPOSURE_ID.DRUG_CONCEPT_ID	<p>Concept identifier for the Drug concept used to create the Drug Eras.</p> <p>Drug Concept Identifier will contain Drug Ingredient layer of the Drug Hierarchy.</p> <p>If no matching values found, store the value zero (0).</p>	
DRUG_EXPOSURE_COUNT		Number of individual Drug Exposures consolidated into the Drug Era	

3.2.2 TABLE NAME: CONDITION_ERA

Condition Era table is constructed through an aggregation of individual Condition Occurrences recorded in the CONDITION_OCCURRENCE table. We followed the instructions in **OMOP ETL Mapping (V2.0)** for constructing Condition Eras.

Lessons learned, challenges and Humana modification:

- (1) Records without mapping CONDITION_CONCEPT_ID: Since CONDITION_CONCEPT_ID was set as 0 in CONDITION_OCCURRENCE table for these records, it's inappropriate to aggregate them. We excluded these records before constructing CONDITION_ERA table.
- (2) Sorting and index: This table was sorted by PERSON_ID, CONDITION_OCCURRENCE_TYPE, CONDITION_CONCEPT_ID, CONDITION_ERA_START_DATE and CONDITION_ERA_END_DATE. We also created indexes on columns PERSON_ID, CONDITION_CONCEPT_ID and CONDITION_OCCURRENCE_TYPE.

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 unique identifier	
CONDITION_ERA_START_DATE	CONDITION_START_DATE	Extracted as the minimum value of the CONDITION_START_DATE for all the occurrences included in the Condition Era.	
PERSON_ID	CONDITION_OCCURRENCE_PERSON_ID		Unique identifier for the person for whom the Condition Era has been constructed.
CONFIDENCE		NULL	
CONDITION_ERA_END_DATE	CONDITION_END_DATE CONDITION_START_DATE	If no End Date is available for the last occurrence, then it is the greatest Start Date or End Date of any occurrence.	
CONDITION_CONCEPT_ID	CONDITION_OCCURRENCE_CONDITION_CONCEPT_ID	If no matching values found, store the value zero (0).	Concept identifier for the Condition concept used to create the Condition Eras. This concept can be the same as the Condition concept related to the individual occurrences or a higher level concept from the Condition concept hierarchy in the vocabulary.
CONDITION_OCCURRENCE_TYPE	CONDITION_OCCURRENCE_TYPE	The following settings apply: For Eras created with 30 day persistence window setting: CONDITION_OCCURRENCE_TYPE matching CONDITION_OCCURRENCE_TYPE_DESC of 'Condition Era – 0 day window' For Eras created with 0 day persistence window setting: CONDITION_OCCURRENCE_TYPE	Condition Occurrence Type for the Era created based on the persistence window setting.

Destination Field	Source Field	Applied Rule	Comment
		matching CONDITION_OCCURRENCE_TYPE_DESC of 'Condition Era – 30 day window'	
CONDITION_OCCURRENCE_COUNT		Number of individual Condition Occurrences consolidated into the Condition Era	

3.3 Reference Tables

3.3.1 TABLE NAME: DRUG_EXPOSURE_REF

Drug Exposure Types are used to define the indicators from which exposures have been extracted. They also define the characteristics of the exposure and the level of aggregation. The following Drug Exposure Types are allowed.

Drug Exposure Type	Drug Exposure Type Description	Persistence Window (In Days)
1	Prescription Dispensed	
2	Prescription Written	
3	Medication List	
4	Physician Administered Drug (Identified as Procedure)	
5	Inpatient Administration	
6	Drug Era – 0 day window	0
7	Drug Era – 30 days window	30

3.3.2 TABLE NAME: CONDITION_OCCURRENCE_REF

Condition Occurrence Reference table serves as the reference listing of various types of Condition Occurrences recorded for analysis. The Condition Occurrence Type conveys the indicator(s) from which the Condition Occurrence was captured and defines the characteristic of the occurrence and the level of aggregation.

Humana modification:

(1) Add one row

48	Outpatient Header	Primary
----	-------------------	---------

(2) Delete one row

62	Outpatient Header	15
----	-------------------	----

(3) Change Condition Occurrence Type sequentially

This table is loaded based on a reference list of Occurrence types, descriptions and persistence window settings. The current listing is as follows:

Condition Occurrence Type	Condition Occurrence Type Description	Condition Occurrence Position	Persistence Window (in days)
1	Inpatient Detail	Primary	
2	Inpatient Detail	1	
3	Inpatient Detail	2	
4	Inpatient Detail	3	
5	Inpatient Detail	4	
6	Inpatient Detail	5	
7	Inpatient Detail	6	
8	Inpatient Detail	7	
9	Inpatient Detail	8	
10	Inpatient Detail	9	
11	Inpatient Detail	10	
12	Inpatient Detail	11	
13	Inpatient Detail	12	
14	Inpatient Detail	13	
15	Inpatient Detail	14	
16	Inpatient Detail	15	
17	Inpatient Header	Primary	
18	Inpatient Header	1	
19	Inpatient Header	2	
20	Inpatient Header	3	
21	Inpatient Header	4	
22	Inpatient Header	5	
23	Inpatient Header	6	
24	Inpatient Header	7	

Condition Occurrence Type	Condition Occurrence Type Description	Condition Occurrence Position	Persistence Window (in days)
25	Inpatient Header	8	
26	Inpatient Header	9	
27	Inpatient Header	10	
28	Inpatient Header	11	
29	Inpatient Header	12	
30	Inpatient Header	13	
31	Inpatient Header	14	
32	Inpatient Header	15	
33	Outpatient Detail	1	
34	Outpatient Detail	2	
35	Outpatient Detail	3	
36	Outpatient Detail	4	
37	Outpatient Detail	5	
38	Outpatient Detail	6	
39	Outpatient Detail	7	
40	Outpatient Detail	8	
41	Outpatient Detail	9	
42	Outpatient Detail	10	
43	Outpatient Detail	11	
44	Outpatient Detail	12	
45	Outpatient Detail	13	
46	Outpatient Detail	14	
47	Outpatient Detail	15	
48	Outpatient Header	Primary	
49	Outpatient Header	1	
50	Outpatient Header	2	
51	Outpatient Header	3	
52	Outpatient Header	4	

Condition Occurrence Type	Condition Occurrence Type Description	Condition Occurrence Position	Persistence Window (in days)
53	Outpatient Header	5	
54	Outpatient Header	6	
55	Outpatient Header	7	
56	Outpatient Header	8	
57	Outpatient Header	9	
58	Outpatient Header	10	
59	Outpatient Header	11	
60	Outpatient Header	12	
61	Outpatient Header	13	
62	Outpatient Header	14	
63	Problem List		
64	Condition Era		0
65	Condition Era		30
66	Death at Discharge		

3.3.3 TABLE NAME: PROC_OCCURRENCE_REF

Procedure Occurrence Reference table serves as the reference listing of various types of Procedure Occurrences recorded for analysis. The Procedure Occurrence Type conveys the indicator(s) from which the Procedure Occurrence was captured, and defines the characteristic of the occurrence.

Humana modification:

(1) Change one row

21	Outpatient Header	1
----	-------------------	---

(2) Change Procedure Occurrence Type sequentially

This table is loaded based on a reference list of occurrence types, position and descriptions. The current listing is as follows:

Procedure Occurrence Type	Procedure Occurrence Type Description	Procedure Occurrence Position
1	Inpatient Detail	Primary
2	Inpatient Detail	1

Procedure Occurrence Type	Procedure Occurrence Type Description	Procedure Occurrence Position
3	Inpatient Header	Primary
4	Inpatient Header	1
5	Inpatient Header	2
6	Inpatient Header	3
7	Inpatient Header	4
8	Inpatient Header	5
9	Inpatient Header	6
10	Inpatient Header	7
11	Inpatient Header	8
12	Inpatient Header	9
13	Inpatient Header	10
14	Inpatient Header	11
15	Inpatient Header	12
16	Inpatient Header	13
17	Inpatient Header	14
18	Inpatient Header	15
19	Outpatient Detail	Primary
20	Outpatient Detail	1
21	Outpatient Header	1
21	Outpatient Header	2
22	Outpatient Header	3
23	Outpatient Header	4
24	Outpatient Header	5
25	Outpatient Header	6
26	Outpatient Header	7
27	EMR Order	

3.3.4 TABLE NAME: OBSERVATION_TYPE_REF

Assignment of an Observation type is essential to determine the type of source data, level of standardization, and coding, as well as the type of result recorded for the Observation. The Observation Types include the following.

- Lab Observation
- EHR observations with text results (e.g., reason for visit)
- Chief Complaint

Data in the OBSERVATION_TYPE_REF table is as follows:

Observation Type	Observation Type Description
CHC	Chief Complaint
EHR	Observation recorded from Electronic Health Records
LAB	Lab Observation
PRL	Problem List from Electronic Health Records
TEM	Observation recorded from Electronic Health Records with text results

3.3.5 TABLE NAME: VOCABULARY_REF

The Vocabulary Reference entity includes a list of all standard terminologies from which Concepts have been extracted for observational analysis using the Common Data Model. The reference table is populated with a single record for each Vocabulary source and includes a descriptive name for the Vocabulary source.

Humana modification: Add one row.

400	Humana
-----	--------

Data in the VOCABULARY_REF table is as follows:

VOCABULARY_CODE	VOCABULARY_NAME
01	SNOMED
02	ICD9 CM
03	ICD9 Procedure

VOCABULARY_CODE	VOCABULARY_NAME
04	CPT
05	HCPCS
06	LOINC
07	NDFRT
08	RxNorm
09	NDC
52	THOMSON
400	Humana

3.3.6 TABLE NAME: RELATIONSHIP_TYPE

A Concept Relationship is standardized via the Relationship Type entity. The Relationship Type codes are adopted from SNOMED-CT. Where the relationships are hierarchical, the Relationship Type captures the “IS A” string that identifies it as a Subtype relationship. Where the relationship is an Object Attribute Value relationship, the Relationship Type holds the Concept that defines the Attribute.

For a complete listing of the RELATIONSHIP_TYPE table see Standard Terminology Specifications Document.