

CDM Inspection report for the Medicare Claims Synthetic Public Use Files database

Package Version: 1.0.0

Date: Sat Jan 16 16:36:42 2021

Authors: My Name

# Table of content

[1 Table of content 1](#_Toc61707426)

[2 General Information 2](#_Toc61707427)

[2.1 Contact Details 2](#_Toc61707428)

[2.2 Database Description 2](#_Toc61707429)

[2.3 SME Role 2](#_Toc61707430)

[3 ETL Development General 3](#_Toc61707431)

[3.1 ETL Documentation 3](#_Toc61707432)

[3.2 ETL Implementation 3](#_Toc61707433)

[3.3 Record counts data tables 3](#_Toc61707434)

[3.4 Data density plots 5](#_Toc61707435)

[3.5 Concepts per person 6](#_Toc61707436)

[4 Vocabulary Mapping 6](#_Toc61707437)

[4.1 Vocabularies 6](#_Toc61707438)

[4.2 Table counts 8](#_Toc61707439)

[4.3 Mapping Completeness 8](#_Toc61707440)

[4.4 Drug Mappings 10](#_Toc61707441)

[4.5 Unmapped Codes 10](#_Toc61707442)

[4.6 Mapped Codes 13](#_Toc61707443)

[4.7 Source to concept map 18](#_Toc61707444)

[5 Technical Infrastructure 18](#_Toc61707445)

[5.1 CDM Source Table 18](#_Toc61707446)

[5.2 HADES packages 18](#_Toc61707447)

[5.3 System Information 19](#_Toc61707448)

[5.4 Vocabulary Query Performance 19](#_Toc61707449)

[5.5 Achilles Query Performance 19](#_Toc61707450)

[6 Scientific Preparedness 27](#_Toc61707451)

[6.1 Staff training 27](#_Toc61707452)

[6.2 Study execution 27](#_Toc61707453)

[7 Quality Control 27](#_Toc61707454)

[8 Maintenance 27](#_Toc61707455)

[9 Checklist 28](#_Toc61707456)

# General Information

The goal of the inspection report is to provide insight into the completeness, transparency and quality of the performed Extraction Transform, and Load (ETL) process and the readiness of the data source to be onboarded in the data network to participate in research studies.

## Contact Details

Fill in the table below

| **items** | **answers** |
| --- | --- |
| Data Partner |  |
| Database fullname | Medicare Claims Synthetic Public Use Files |
| Database acronym | SYNPUF |
| Contact Person |  |
| Email |  |
| SME |  |
| Contact Person |  |
| Email SME |  |

## Database Description

The CMS Linkable 2008–2010 Medicare DE-SynPUF originated from a disjoint (mutually exclusive from existing samples) 5% random sample of beneficiaries from the 100% Beneficiary Summary File for 2008. To exclude any overlap with the beneficiaries in the existing 5% CMS research sample, 3 the beneficiaries in that other sample were excluded, and a 5-in-95 random draw was made with the remaining 95% of beneficiaries. A variety of statistical disclosure limitation techniques were used to protect the confidentiality of Medicare beneficiaries in the CMS Linkable 2008–2010 Medicare DE-SynPUF. The DE-SynPUF was created by starting with an actual beneficiary as a “seed” for a synthetic beneficiary. Synthetic beneficiaries and their claims are based on actual seed beneficiaries. Disclosure is reduced through multiple deterministically or stochastically applied treatment mechanisms. First, hot decking based procedures are used to find donors for beneficiary-level variables and individual claims. Second, other synthetic processes are used to protect other elements of the data. Disclosure limitation methods used in the process include variable reduction, suppression, substitution, synthesis, date perturbation, and coarsening. Please refer to the CMS Linkable 2008–2010 Medicare Data Entrepreneurs’ Synthetic Public Use File (DE-SynPUF) User Manual for details regarding how DE-SynPUF was created

## SME Role

Describe the involvement of the SME in the ETL Delopment process

# ETL Development General

This section decribes the ETL development steps and discusses the quality control steps performed by the SME

## ETL Documentation

Perform the following checks and discuss the findings here:

Approve the quality of the ETL documentation with respect to its completeness and level of detail per data domain. Ideally it is based on the Rabbit-in-a-Hat mapping definition document. If a staging table approach is used, its creation needs to be described in detail.

Does it contain enough detail on the applied business rules and are the THEMIS rules followed?

Compare the ETL documentation with the shared ETL code to make sure it is a correct representation of the implementation. Ideally, end-to-end tests using the Rabbit-in-a-hat testFramework.R is implemented and results are shared. If this is not available explain the quality control mechanism that is applied

Is the ETL code executable fully automatically or are there manual steps? If there are manual steps these need to be explained.

## ETL Implementation

Described the technology used for implementing the ETL (SQL,R, Python etc).

Provide feedback on the level of commenting and code structure. The minimum level of commenting contains an explanation of the sql query, R function, etc. See also the guidance provided by OHDSI. Code structure refers to a logical structure of the SQL/R files. We recommend that the files are name as their target table and contain all code related to that domain, e.g. insert\_person.sql, insert\_condition\_occurence.sql. If another method is applied provide there details.

Is there a version control mechanism in place?

## Record counts data tables

Table 1. Shows the number of records in all clinical data tables

| TABLENAME | COUNT |
| --- | --- |
| cost | 37,169,949 |
| condition\_occurrence | 14,803,406 |
| procedure\_occurrence | 13,857,192 |
| condition\_era | 10,150,726 |
| drug\_exposure | 6,294,108 |
| drug\_era | 6,251,166 |
| visit\_occurrence | 5,579,542 |
| measurement | 3,387,775 |
| observation | 1,925,159 |
| provider | 631,920 |
| payer\_plan\_period | 389,231 |
| care\_site | 239,153 |
| device\_exposure | 224,499 |
| person | 116,352 |
| observation\_period | 104,891 |
| death | 5,461 |
| location | 3,088 |
| dose\_era | 0 |
| note | 0 |
| specimen | 0 |
| visit\_details | 0 |

Query executed in 10.93 secs

## Data density plots

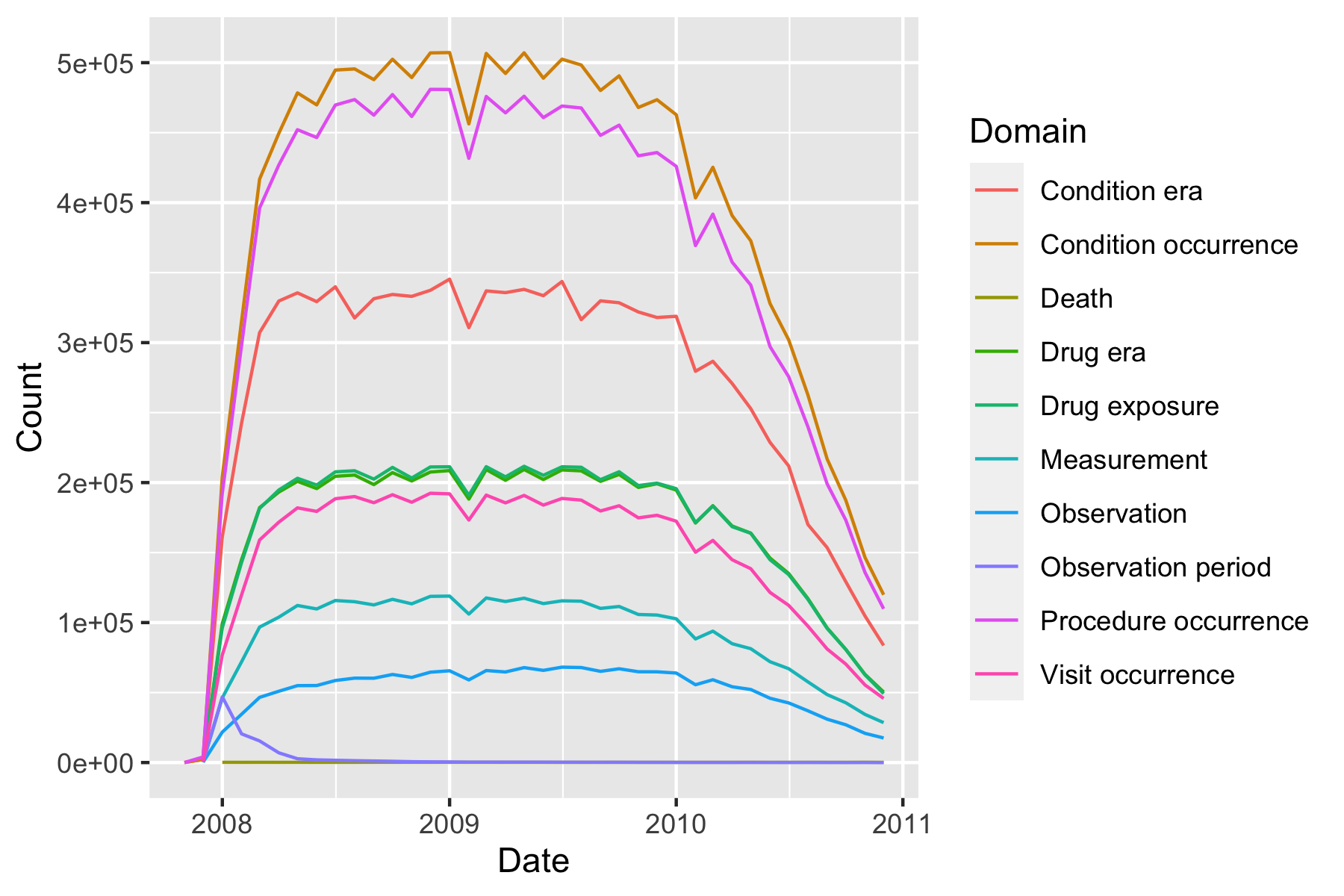


Figure 1. Total record count over time per data domain

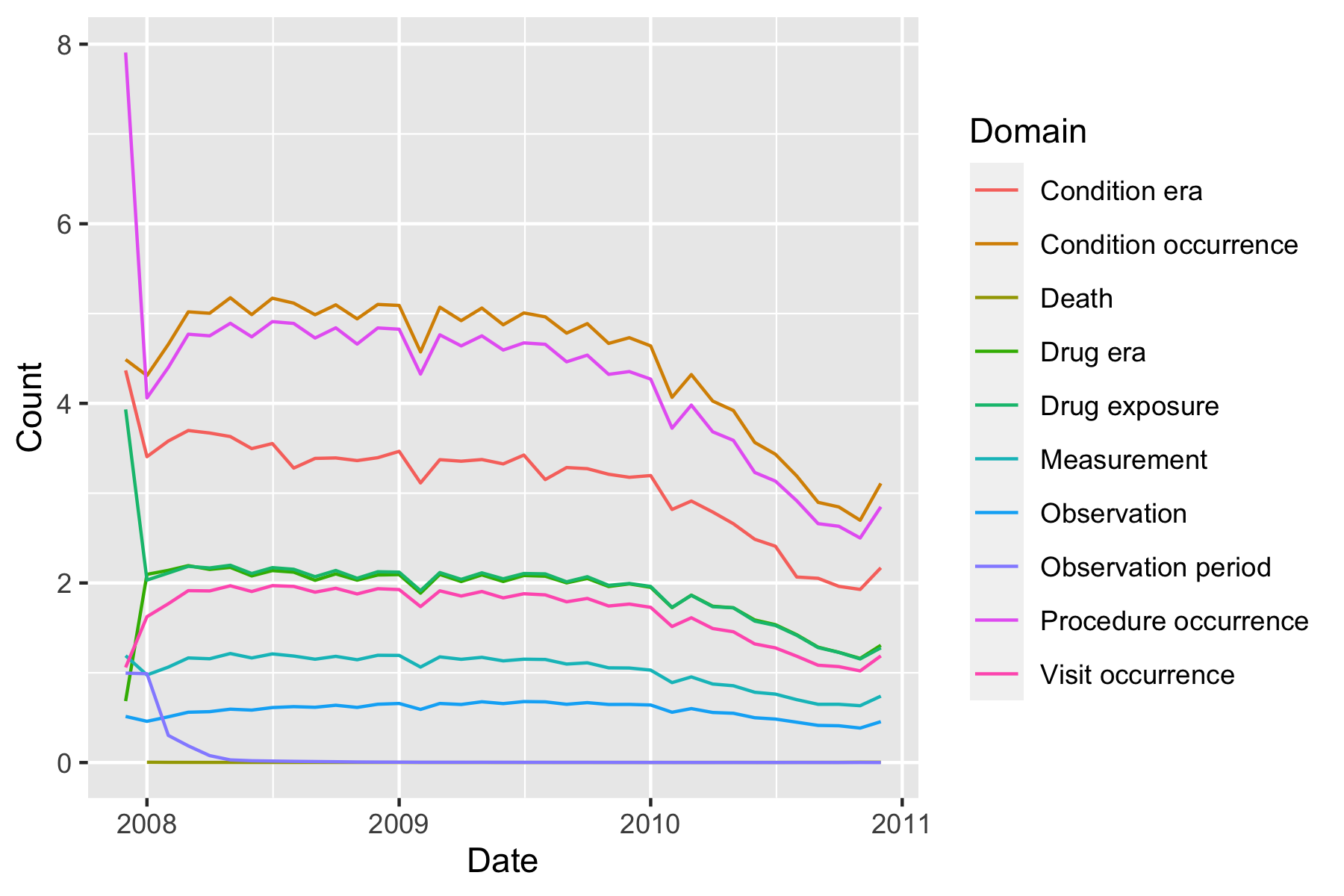


Figure 2. Number of records per person over time per data domain

## Concepts per person

Table 2. Shows the number of records per person for all data domains

| Domain | Min | P10 | P25 | MEDIAN | P75 | P90 | Max |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Condition era | 1 | 8 | 38 | 85 | 136 | 188 | 353 |
| Condition occurrence | 1 | 8 | 38 | 86 | 138 | 190 | 357 |
| Drug era | 1 | 4 | 9 | 39 | 78 | 110 | 200 |
| Drug exposure | 1 | 5 | 9 | 41 | 87 | 128 | 241 |
| Measurement | 1 | 4 | 12 | 21 | 31 | 40 | 93 |
| Observation | 1 | 2 | 6 | 11 | 17 | 23 | 51 |
| Procedure occurrence | 1 | 7 | 34 | 68 | 103 | 136 | 245 |

# Vocabulary Mapping

Describe how the vocabulary mapping process was implemented, and what the quality control mechanism are.

All the custom mappings need to be shared with the report as Excel file or as source\_to\_concept map, to allow for random checks by EHDEN’s vocabulary team. Ideally these lists are sorted descending by source code frequency.

## Vocabularies

Vocabulary version: v5.0 05-NOV-17

Table 3. The vocabularies available in the CDM with concept count. Note that this does not reflect which concepts are actually used in the clinical CDM tables. S=Standard, C=Classification and '-'=Non-standard

| ID | NAME | VERSION | S | C | - |
| --- | --- | --- | --- | --- | --- |
| ABMS | Provider Specialty (American Board of Medical Specialties) | NA | 89 | 0 | 0 |
| ATC | WHO Anatomic Therapeutic Chemical Classification | RxNorm Full 20170807 | 0 | 6,082 | 47 |
| Cohort Type | OMOP Cohort Type | NA | 0 | 0 | 1 |
| Concept Class | OMOP Concept Class | NA | 0 | 0 | 300 |
| Condition Type | OMOP Condition Occurrence Type | NA | 100 | 0 | 0 |
| Cost Type | OMOP Cost Type | NA | 3 | 0 | 0 |
| CPT4 | Current Procedural Terminology version 4 (AMA) | 2017AA | 10,524 | 3,262 | 1,660 |
| Currency | International Currency Symbol (ISO 4217) | 2008 | 180 | 0 | 0 |
| Death Type | OMOP Death Type | NA | 14 | 0 | 0 |
| Device Type | OMOP Device Type | NA | 3 | 0 | 0 |
| Domain | OMOP Domain | NA | 0 | 0 | 55 |
| Drug Type | OMOP Drug Exposure Type | NA | 14 | 0 | 0 |
| Ethnicity | OMOP Ethnicity | NA | 2 | 0 | 0 |
| Gender | OMOP Gender | NA | 2 | 0 | 3 |
| HCPCS | Healthcare Common Procedure Coding System (CMS) | 2017 Alpha Numeric HCPCS File | 5,294 | 0 | 2,835 |
| ICD10CM | International Classification of Diseases, Tenth Revision, Clinical Modification (NCHS) | ICD10CM FY2017 code descriptions | 0 | 0 | 108,971 |
| ICD9CM | International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 1 and 2 (NCHS) | ICD9CM v32 master descriptions | 0 | 0 | 18,672 |
| ICD9Proc | International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 3 (NCHS) | ICD9CM v32 master descriptions | 4,651 | 0 | 6 |
| LOINC | Logical Observation Identifiers Names and Codes (Regenstrief Institute) | LOINC 2.61 | 93,220 | 38,458 | 4,221 |
| Meas Type | OMOP Measurement Type | NA | 6 | 0 | 0 |
| NDC | National Drug Code (FDA and manufacturers) | NDC 20171023 | 0 | 0 | 769,530 |
| NDFRT | National Drug File - Reference Terminology (VA) | RxNorm Full 20170807 | 0 | 18,542 | 18,869 |
| None | OMOP Standardized Vocabularies | v5.0 05-NOV-17 | 0 | 0 | 1 |
| Note Type | OMOP Note Type | NA | 10 | 0 | 0 |
| NUCC | National Uniform Claim Committee Health Care Provider Taxonomy Code Set (NUCC) | 2013-07-01 | 829 | 0 | 0 |
| Obs Period Type | OMOP Observation Period Type | NA | 6 | 0 | 0 |
| Observation Type | OMOP Observation Type | NA | 13 | 0 | 0 |
| Place of Service | Place of Service Codes for Professional Claims (CMS) | 2009-01-11 | 59 | 0 | 0 |
| Procedure Type | OMOP Procedure Occurrence Type | NA | 95 | 0 | 0 |
| Race | Race and Ethnicity Code Set (USBC) | Version 1.0 | 50 | 0 | 3 |
| Relationship | OMOP Relationship | NA | 12 | 0 | 416 |
| Revenue Code | UB04/CMS1450 Revenue Codes (CMS) | 2010 Release | 538 | 0 | 0 |
| RxNorm | RxNorm (NLM) | RxNorm Full 20170807 | 146,228 | 37,172 | 95,169 |
| RxNorm Extension | RxNorm Extension (OMOP) | RxNorm Extension 15-OCT-17 | 1,468,309 | 0 | 193,853 |
| SNOMED | Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO) | SnomedCT Release 20170401 | 331,235 | 0 | 454,477 |
| Specialty | Medicare provider/supplier specialty codes (CMS) | NA | 107 | 0 | 4 |
| SPL | Structured Product Labeling (FDA) | NDC 20171023 | 0 | 224,583 | 14,854 |
| UCUM | Unified Code for Units of Measure (Regenstrief Institute) | Version 1.8.2 | 883 | 0 | 89 |
| VA Class | VA National Drug File Class (VA) | RxNorm Full 20170807 | 0 | 486 | 0 |
| Visit | OMOP Visit | NA | 5 | 0 | 0 |
| Visit Type | OMOP Visit Type | NA | 3 | 0 | 0 |
| Vocabulary | OMOP Vocabulary | NA | 0 | 0 | 82 |

Query executed in 4.43 secs

## Table counts

Table 4. Shows the number of records in all vocabulary tables

| TABLENAME | COUNT |
| --- | --- |
| concept | 4,075,186 |
| concept\_ancestor | 121,111,843 |
| concept\_class | 300 |
| concept\_relationship | 27,457,992 |
| concept\_synonym | 5,491,313 |
| domain | 39 |
| drug\_strength | 2,430,381 |
| vocabulary | 42 |

Query executed in 17.75 secs

## Mapping Completeness

Table 5. Shows the percentage of codes that are mapped to the standardized vocabularies as well as the percentage of records.

| Domain | #Codes Source | #Codes Mapped | %Codes Mapped | #Records Source | #Records Mapped | %Records Mapped |
| --- | --- | --- | --- | --- | --- | --- |
| condition | 10,313 | 10,292 | 99.80 | 14,803,406 | 14,775,949 | 99.81 |
| procedure | 8,607 | 8,018 | 93.16 | 13,857,192 | 13,246,809 | 95.60 |
| device | 811 | 760 | 93.71 | 224,499 | 222,163 | 98.96 |
| drug | 269,187 | 265,251 | 98.54 | 6,294,108 | 6,219,065 | 98.81 |
| observation | 1,759 | 1,612 | 91.64 | 1,925,159 | 1,765,558 | 91.71 |
| measurement | 998 | 955 | 95.69 | 3,387,775 | 3,358,281 | 99.13 |
| visit\_occurrence | 5,579,542 | 844,593 | 15.14 | 5,579,542 | 844,593 | 15.14 |
| measurement-unit | 0 | NA | NA | NA | NA | NA |
| observation-unit | 0 | NA | NA | NA | NA | NA |
| measurement-value | 1 | 0 | 0.00 | 3,387,775 | 0 | 0.00 |
| observation-value | 1 | 0 | 0.00 | 1,925,159 | 0 | 0.00 |

Query executed in 25.54 secs

## Drug Mappings

Table 6. The level of the drug mappings

| CLASS | #RECORDS | #PATIENTS | #SOURCE CODES |
| --- | --- | --- | --- |
| Clinical Drug | 4,402,549 | 100,478 | 215,778 |
| Branded Drug | 842,564 | 79,921 | 34,446 |
| Quant Clinical Drug | 186,097 | 59,496 | 9,088 |
| Quant Branded Drug | 133,461 | 53,189 | 5,444 |
| Undefined | 75,043 | 44,189 | 3,936 |
| Ingredient | 248,640 | 49,172 | 116 |
| Clinical Drug Comp | 192,172 | 45,198 | 104 |
| Clinical Drug Form | 47,409 | 28,118 | 93 |
| HCPCS | 62,084 | 22,791 | 81 |
| CPT4 | 101,888 | 53,606 | 48 |
| Clinical Pack | 394 | 391 | 36 |
| Branded Pack | 167 | 167 | 12 |
| Branded Drug Form | 1,640 | 1,556 | 5 |

Query executed in 31.84 secs

## Unmapped Codes

Table 7. Top 25 of unmapped drugs. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | SOURCE VALUE | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | G9141 | 6,200 | 5,900 |
| 2 | J3487 | 2,100 | 2,000 |
| 3 | G9142 | 1,100 | 1,000 |
| 4 | J3488 | 900 | 800 |
| 5 | 90703 | 600 | 600 |
| 6 | 00047200820 | 200 | 200 |
| 7 | 00047200922 | 200 | 200 |
| 8 | 00047200920 | 200 | 200 |
| 9 | 00047200822 | 200 | 200 |
| 10 | 00703408511 | 100 | 100 |
| 11 | 17088266101 | 100 | 100 |
| 12 | 54087032518 | 100 | 100 |
| 13 | 40042001710 | 100 | 100 |
| 14 | 17088260901 | 100 | 100 |
| 15 | 17088260601 | 100 | 100 |
| 16 | 17088260101 | 100 | 100 |
| 17 | 00555119614 | 100 | 100 |
| 18 | 54087035618 | 100 | 100 |
| 19 | 58364119604 | 100 | 100 |
| 20 | 54087032618 | 100 | 100 |
| 21 | 15210040211 | 100 | 100 |
| 22 | 00083260901 | 100 | 100 |
| 23 | J1785 | 100 | 100 |
| 24 | 17088009201 | 100 | 100 |
| 25 | 00083260104 | 100 | 100 |

Query executed in 1.09 secs

Table 8. All 15 unmapped conditions. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | SOURCE VALUE | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | 7395 | 13,700 | 10,200 |
| 2 | 7390 | 4,400 | 3,600 |
| 3 | 7396 | 3,100 | 3,000 |
| 4 | 7397 | 2,800 | 2,600 |
| 5 | 7398 | 2,600 | 2,500 |
| 6 | 32735 | 400 | 400 |
| 7 | 99559 | 200 | 200 |
| 8 | 99552 | 200 | 200 |
| 9 | 0770 | 200 | 200 |
| 10 | 37641 | 100 | 100 |
| 11 | 5226 | 100 | 100 |
| 12 | 1398 | 100 | 100 |
| 13 | 65583 | 100 | 100 |
| 14 | 65593 | 100 | 100 |
| 15 | 74446 | 100 | 100 |

Query executed in 0.98 secs

Table 9. Top 25 of unmapped measurements. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | SOURCE VALUE | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | 80101 | 8,000 | 6,200 |
| 2 | G0431 | 3,800 | 2,700 |
| 3 | 82055 | 2,800 | 2,600 |
| 4 | 83925 | 2,300 | 1,600 |
| 5 | 87621 | 1,900 | 1,800 |
| 6 | 80154 | 1,700 | 1,500 |
| 7 | 82145 | 1,400 | 1,100 |
| 8 | 80100 | 1,200 | 1,200 |
| 9 | 83840 | 1,100 | 800 |
| 10 | 80102 | 800 | 800 |
| 11 | 82205 | 800 | 600 |
| 12 | 82003 | 700 | 700 |
| 13 | 80196 | 700 | 700 |
| 14 | 82520 | 700 | 600 |
| 15 | 82491 | 600 | 600 |
| 16 | 83805 | 600 | 500 |
| 17 | G0430 | 300 | 300 |
| 18 | 82492 | 300 | 200 |
| 19 | 82541 | 300 | 300 |
| 20 | 88347 | 200 | 200 |
| 21 | 82544 | 100 | 100 |
| 22 | 82651 | 100 | 100 |
| 23 | 83887 | 100 | 100 |
| 24 | 80182 | 100 | 100 |
| 25 | 82980 | 100 | 100 |

Query executed in 1.73 secs

Table 10. Top 25 of unmapped observations. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | SOURCE VALUE | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | G8445 | 56,100 | 35,600 |
| 2 | G8447 | 33,900 | 25,500 |
| 3 | G8443 | 15,900 | 13,700 |
| 4 | G8446 | 13,700 | 12,100 |
| 5 | G8553 | 10,100 | 9,000 |
| 6 | G8457 | 6,000 | 5,600 |
| 7 | G8448 | 5,500 | 5,200 |
| 8 | Q1003 | 5,300 | 4,700 |
| 9 | G0317 | 3,600 | 1,100 |
| 10 | G8440 | 1,300 | 1,200 |
| 11 | G8429 | 1,000 | 1,000 |
| 12 | M0064 | 900 | 900 |
| 13 | G8470 | 500 | 500 |
| 14 | G8455 | 500 | 500 |
| 15 | G0323 | 400 | 400 |
| 16 | V8905 | 400 | 400 |
| 17 | G0319 | 400 | 300 |
| 18 | G8590 | 400 | 400 |
| 19 | V8909 | 300 | 300 |
| 20 | G0327 | 300 | 100 |
| 21 | G8593 | 300 | 300 |
| 22 | G8437 | 300 | 300 |
| 23 | V8901 | 300 | 300 |
| 24 | G8588 | 300 | 300 |
| 25 | G8449 | 300 | 300 |

Query executed in 1.12 secs

Table 11. Top 25 of unmapped procedures. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | SOURCE VALUE | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | 77052 | 51,100 | 28,700 |
| 2 | 92135 | 42,500 | 26,700 |
| 3 | 90806 | 32,500 | 16,900 |
| 4 | 90862 | 28,900 | 17,700 |
| 5 | 97001 | 25,200 | 18,800 |
| 6 | 77057 | 20,600 | 16,000 |
| 7 | 73510 | 19,300 | 15,700 |
| 8 | 78478 | 15,400 | 10,000 |
| 9 | 78480 | 15,200 | 10,000 |
| 10 | 90772 | 12,800 | 10,300 |
| 11 | 77418 | 12,400 | 6,800 |
| 12 | 78465 | 11,500 | 9,600 |
| 13 | 77051 | 9,300 | 7,200 |
| 14 | 62311 | 8,600 | 7,800 |
| 15 | 93545 | 8,600 | 6,400 |
| 16 | 93556 | 8,500 | 6,300 |
| 17 | 77413 | 8,500 | 4,700 |
| 18 | 90805 | 8,100 | 6,800 |
| 19 | 90801 | 8,100 | 7,000 |
| 20 | 90816 | 7,400 | 5,800 |
| 21 | 90818 | 7,400 | 5,900 |
| 22 | 76645 | 7,300 | 6,400 |
| 23 | 77414 | 7,200 | 4,300 |
| 24 | 93543 | 7,100 | 5,500 |
| 25 | 93555 | 6,900 | 5,300 |

Query executed in 1.74 secs

Table 12. All 11 unmapped devices. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | SOURCE VALUE | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | C1300 | 2,000 | 1,400 |
| 2 | C1879 | 200 | 200 |
| 3 | E0230 | 100 | 100 |
| 4 | L1800 | 100 | 100 |
| 5 | L1825 | 100 | 100 |
| 6 | L3651 | 100 | 100 |
| 7 | D1110 | 100 | 100 |
| 8 | A6200 | 100 | 100 |
| 9 | L3909 | 100 | 100 |
| 10 | E0238 | 100 | 100 |
| 11 | L3701 | 100 | 100 |

Query executed in 0.44 secs

## Mapped Codes

Table 13. Top 25 of mapped drugs. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Epoetin Alfa | 139,300 | 9,500 |
| 2 | "" | 101,900 | 53,700 |
| 3 | paricalcitol 0.001 MG | 75,800 | 2,800 |
| 4 | Oxygen 99 % Gas for Inhalation | 61,500 | 37,400 |
| 5 | Gemfibrozil 600 MG Oral Tablet | 37,700 | 27,000 |
| 6 | Omeprazole 20 MG Delayed Release Oral Capsule | 37,000 | 24,800 |
| 7 | Lovastatin 20 MG Oral Tablet | 32,100 | 24,100 |
| 8 | Simvastatin 40 MG Oral Tablet | 29,900 | 23,000 |
| 9 | Simvastatin 20 MG Oral Tablet | 27,900 | 21,600 |
| 10 | Lovastatin 40 MG Oral Tablet | 27,300 | 21,300 |
| 11 | Furosemide 40 MG Oral Tablet | 25,100 | 19,900 |
| 12 | Metformin hydrochloride 500 MG Oral Tablet | 24,400 | 17,600 |
| 13 | Hydrochlorothiazide 50 MG Oral Tablet | 24,300 | 19,500 |
| 14 | Injection, iron sucrose, 1 mg | 24,100 | 3,300 |
| 15 | Glyburide 5 MG Oral Tablet | 23,300 | 17,100 |
| 16 | Doxercalciferol 0.001 MG | 23,000 | 2,300 |
| 17 | Sodium Chloride Injectable Solution | 22,900 | 16,800 |
| 18 | Dipyridamole 25 MG Oral Tablet | 22,700 | 15,500 |
| 19 | Hydrochlorothiazide 25 MG Oral Tablet | 22,100 | 18,000 |
| 20 | Simvastatin 10 MG Oral Tablet | 21,900 | 18,000 |
| 21 | Lovastatin 10 MG Oral Tablet | 21,500 | 17,800 |
| 22 | Dipyridamole 50 MG Oral Tablet | 21,100 | 14,700 |
| 23 | Furosemide 20 MG Oral Tablet | 21,000 | 17,200 |
| 24 | Glipizide 10 MG Oral Tablet | 20,000 | 15,300 |
| 25 | Dipyridamole 75 MG Oral Tablet | 19,800 | 14,000 |

Query executed in 13.84 secs

Table 14. Top 25 of mapped conditions. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Type 2 diabetes mellitus | 612,900 | 78,700 |
| 2 | Atrial fibrillation | 299,000 | 61,100 |
| 3 | Chest pain | 246,700 | 57,400 |
| 4 | Pure hypercholesterolemia | 224,500 | 67,700 |
| 5 | Anemia | 194,300 | 57,200 |
| 6 | Coronary arteriosclerosis in native artery | 184,100 | 53,400 |
| 7 | Hypothyroidism | 183,000 | 62,400 |
| 8 | Malaise and fatigue | 171,400 | 57,600 |
| 9 | Congestive heart failure | 166,500 | 51,700 |
| 10 | Urinary tract infectious disease | 164,100 | 54,100 |
| 11 | Low back pain | 154,900 | 52,300 |
| 12 | Coronary arteriosclerosis | 128,600 | 51,300 |
| 13 | Dyspnea | 118,100 | 46,300 |
| 14 | Pain in limb | 115,400 | 52,100 |
| 15 | Abdominal pain | 100,100 | 43,000 |
| 16 | Type II diabetes mellitus uncontrolled | 94,700 | 36,400 |
| 17 | End stage renal disease | 93,400 | 21,400 |
| 18 | Actinic keratosis | 85,700 | 33,600 |
| 19 | Gastroesophageal reflux disease | 83,600 | 47,200 |
| 20 | Iron deficiency anemia | 83,000 | 26,500 |
| 21 | Onychomycosis due to dermatophyte | 82,800 | 41,700 |
| 22 | Nuclear senile cataract | 80,600 | 37,900 |
| 23 | Conduction disorder of the heart | 75,100 | 36,900 |
| 24 | Arthralgia of the lower leg | 74,200 | 33,800 |
| 25 | Respiratory symptom | 72,700 | 35,000 |

Query executed in 19.66 secs

Table 15. Top 25 of mapped measurements. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | "" | 3,290,600 | 91,900 |
| 2 | Prostate cancer screening; prostate specific antigen test (psa) | 16,900 | 15,000 |
| 3 | Laboratory test | 16,900 | 13,200 |
| 4 | Screening cytopathology, cervical or vaginal (any reporting system), collected in preservative fluid, automated thin layer preparation, with screening by automated system and manual rescreening under physician supervision | 5,600 | 5,300 |
| 5 | Screening cytopathology, cervical or vaginal (any reporting system), collected in preservative fluid, automated thin layer preparation, screening by cytotechnologist under physician supervision | 3,900 | 3,800 |
| 6 | Immunology laboratory test | 3,100 | 2,600 |
| 7 | Colorectal cancer screening; fecal occult blood test, immunoassay, 1-3 simultaneous | 3,100 | 3,000 |
| 8 | Hematology screening test | 2,600 | 2,500 |
| 9 | Microscopic examination of vaginal Papanicolaou smear | 2,500 | 2,500 |
| 10 | Genetic test | 2,500 | 2,400 |
| 11 | Microscopic examination of cervical Papanicolaou smear | 1,700 | 1,600 |
| 12 | Blood group typing | 1,600 | 1,600 |
| 13 | Phenylketonuria screening test | 1,200 | 1,200 |
| 14 | Antenatal RhD antibody screening | 1,200 | 1,200 |
| 15 | Sickle cell disease screening test | 1,200 | 1,200 |
| 16 | Detection of parasite | 800 | 800 |
| 17 | Complete cbc, automated (hgb, hct, rbc, wbc, without platelet count) and automated wbc differential count | 700 | 700 |
| 18 | Screening papanicolaou smear, cervical or vaginal, up to three smears, by technician under physician supervision | 600 | 600 |
| 19 | Screening cytopathology, cervical or vaginal (any reporting system), collected in preservative fluid, automated thin layer preparation, requiring interpretation by physician | 500 | 500 |
| 20 | Complete (cbc), automated (hgb, hct, rbc, wbc; without platelet count) | 400 | 400 |
| 21 | Wet mounts, including preparations of vaginal, cervical or skin specimens | 400 | 400 |
| 22 | Type 1 hypersensitivity skin test | 400 | 400 |
| 23 | All potassium hydroxide (koh) preparations | 200 | 200 |
| 24 | Screening cytopathology, cervical or vaginal (any reporting system), collected in preservative fluid, automated thin layer preparation, with manual screening and rescreening by cytotechnologist under physician supervision | 200 | 200 |
| 25 | Screening cytopathology, cervical or vaginal (any reporting system), collected in preservative fluid, automated thin layer preparation, with screening by automated system, under physician supervision | 100 | 100 |

Query executed in 6.50 secs

Table 16. Top 25 of mapped observations. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Vaccination required | 269,300 | 66,300 |
| 2 | "" | 220,800 | 68,000 |
| 3 | High risk drug monitoring status | 184,900 | 53,100 |
| 4 | History of clinical finding in subject | 178,700 | 63,600 |
| 5 | Ground mileage, per statute mile | 82,600 | 34,700 |
| 6 | Postoperative care | 50,900 | 31,800 |
| 7 | H/O: artificial joint | 42,900 | 24,100 |
| 8 | Ambulance service, basic life support, non-emergency transport, (bls) | 38,300 | 23,200 |
| 9 | Diabetic on insulin | 34,100 | 25,300 |
| 10 | Family history of clinical finding | 31,800 | 24,900 |
| 11 | Aftercare | 30,700 | 21,700 |
| 12 | Ambulance service, advanced life support, emergency transport, level 1 (als 1 - emergency) | 29,700 | 22,500 |
| 13 | H/O: artificial eye lens | 27,200 | 19,800 |
| 14 | Antineoplastic chemotherapy regimen | 24,400 | 12,200 |
| 15 | Low osmolar contrast material, 300-399 mg/ml iodine concentration, per ml | 21,700 | 16,400 |
| 16 | Abnormal weight loss | 21,600 | 15,500 |
| 17 | Cardiac pacemaker in situ | 20,700 | 16,300 |
| 18 | Dialysis finding | 20,300 | 9,000 |
| 19 | History of renal transplant | 18,300 | 8,500 |
| 20 | Ambulance service, basic life support, emergency transport (bls-emergency) | 16,200 | 14,000 |
| 21 | H/O: artificial heart valve | 12,900 | 9,700 |
| 22 | Fall | 12,400 | 10,900 |
| 23 | Follow-up encounter | 12,300 | 10,600 |
| 24 | H/O: surgery | 10,800 | 9,400 |
| 25 | Set-up portable x-ray equipment | 10,400 | 8,500 |

Query executed in 5.03 secs

Table 17. Top 25 of mapped procedures. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | "" | 6,710,200 | 97,900 |
| 2 | Other diagnostic procedures on lymphatic structures | 658,600 | 84,500 |
| 3 | Biopsy of lymphatic structure | 517,300 | 81,300 |
| 4 | Biopsy of mouth, unspecified structure | 400,400 | 78,200 |
| 5 | Long-term drug therapy | 217,200 | 65,300 |
| 6 | Excision of anus | 137,000 | 51,800 |
| 7 | Biopsy of uvula and soft palate | 116,200 | 50,400 |
| 8 | Screening for malignant neoplasm of breast | 103,900 | 37,200 |
| 9 | Incision of cervix to assist delivery | 101,200 | 36,000 |
| 10 | Surgical correction of prominent ear | 98,800 | 23,800 |
| 11 | Other resection of rectum | 90,100 | 33,600 |
| 12 | Administration of influenza virus vaccine | 82,500 | 45,800 |
| 13 | Bone graft, humerus | 78,200 | 28,400 |
| 14 | External version assisting delivery | 74,600 | 30,800 |
| 15 | Removal of implanted devices from bone, humerus | 65,400 | 37,300 |
| 16 | High forceps operation with episiotomy | 65,200 | 32,500 |
| 17 | Electrical stimulation (unattended), to one or more areas for indication(s) other than wound care, as part of a therapy plan of care | 60,400 | 29,900 |
| 18 | Other gastroenterostomy without gastrectomy | 56,700 | 33,300 |
| 19 | Application of external fixator device, humerus | 55,700 | 27,500 |
| 20 | Limb shortening procedures, radius and ulna | 55,600 | 33,900 |
| 21 | Temporary tracheostomy | 52,800 | 32,400 |
| 22 | Bone graft, carpals and metacarpals | 52,200 | 31,200 |
| 23 | Other and unspecified robotic assisted procedure | 51,500 | 16,100 |
| 24 | Other repair of urethra | 51,000 | 23,900 |
| 25 | Other diagnostic procedures on orbit and eyeball | 46,800 | 14,500 |

Query executed in 22.10 secs

Table 18. Top 25 of mapped devices. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

| ROW\_NUM | CONCEPT NAME | #RECORDS | #SUBJECTS |
| --- | --- | --- | --- |
| 1 | Syringe, with or without needle, each | 129,500 | 2,600 |
| 2 | Non-covered item or service | 15,400 | 12,500 |
| 3 | Technetium tc-99m sestamibi, diagnostic, per study dose | 7,100 | 6,000 |
| 4 | Technetium tc-99m tetrofosmin, diagnostic, per study dose | 5,400 | 4,700 |
| 5 | Injection, gadolinium-based magnetic resonance contrast agent, not otherwise specified (nos), per ml | 5,300 | 4,700 |
| 6 | Miscellaneous dialysis supplies, not otherwise specified | 5,300 | 1,300 |
| 7 | Guide wire | 4,300 | 4,000 |
| 8 | Fluorodeoxyglucose f-18 fdg, diagnostic, per study dose, up to 45 millicuries | 2,700 | 2,400 |
| 9 | Thallium tl-201 thallous chloride, diagnostic, per millicurie | 2,700 | 2,500 |
| 10 | Introducer/sheath, other than guiding, other than intracardiac electrophysiological, non-laser | 2,500 | 2,400 |
| 11 | Surgical trays | 2,400 | 2,300 |
| 12 | Technetium tc-99m medronate, diagnostic, per study dose, up to 30 millicuries | 2,000 | 1,700 |
| 13 | Catheter, guiding (may include infusion/perfusion capability) | 1,700 | 1,600 |
| 14 | Catheter, transluminal angioplasty, non-laser (may include guidance, infusion/perfusion capability) | 1,600 | 1,600 |
| 15 | Closure device, vascular (implantable/insertable) | 1,400 | 1,400 |
| 16 | Catheter, infusion, inserted peripherally, centrally or midline (other than hemodialysis) | 900 | 900 |
| 17 | Gauze, non-impregnated, sterile, pad size 16 sq. in. or less, without adhesive border, each dressing | 800 | 600 |
| 18 | Posterior chamber intraocular lens | 800 | 800 |
| 19 | Lens, intraocular (new technology) | 700 | 700 |
| 20 | Gauze, non-impregnated, non-sterile, pad size 16 sq. in. or less, without adhesive border, each dressing | 600 | 300 |
| 21 | Mesh (implantable) | 600 | 600 |
| 22 | Catheter, drainage | 600 | 600 |
| 23 | Anchor/screw for opposing bone-to-bone or soft tissue-to-bone (implantable) | 500 | 500 |
| 24 | Technetium tc-99m sulfur colloid, diagnostic, per study dose, up to 20 millicuries | 500 | 500 |
| 25 | Surgical supply; miscellaneous | 500 | 500 |

Query executed in 1.21 secs

## Source to concept map

If you did not use the source\_to\_concept\_map table in the ETL the table below will be empty. In that case provide your custom mappings in an Excel file.

Table 19. Source to concept map breakdown

| SOURCE\_VOCABULARY\_ID | TARGET\_VOCABULARY\_ID | COUNT |
| --- | --- | --- |
|  |  |  |

Query executed in 0.50 secs

Note that the full source\_to\_concept\_map table is added in the results.zip

# Technical Infrastructure

Check that the following tools are available and functional: ATLAS, ACHILLES report. Functionality needs to be tested by design of cohort in Atlas, generation of cohort counts in ATLAS, execution of a simple cohort characterisation in ATLAS.

Is the data source added in the EHDEN Database Catalogue and has the CatalogUeExport results been uploaded for the visualizations? Also describe if a process has been agreed for updating this information regularly.

Add additional relevant information about the local infrastructure here, such as backup facilities, specifications webserver hosting ATLAS, testing environment if available etc.

## CDM Source Table

Table 20. cdm\_source table content

| field | 1 |
| --- | --- |
| CDM\_SOURCE\_NAME | synpuf |
| CDM\_SOURCE\_ABBREVIATION | synpuf |
| CDM\_HOLDER | ohdsi |
| SOURCE\_DESCRIPTION | simulated data |
| SOURCE\_DOCUMENTATION\_REFERENCE | NA |
| CDM\_ETL\_REFERENCE | NA |
| SOURCE\_RELEASE\_DATE | 2020-07-01 |
| CDM\_RELEASE\_DATE | 2020-07-07 |
| CDM\_VERSION | 5.3.1 |
| VOCABULARY\_VERSION | v5.0 05-NOV-17 |

## HADES packages

Table 21. Versions of all installed HADES R packages

| Package | Version |
| --- | --- |
| Andromeda | 0.4.0 |
| CohortDiagnostics | 2.0.0 |
| CohortMethod | 4.0.1 |
| Cyclops | 3.1.0 |
| DatabaseConnector | 3.0.0 |
| DatabaseConnectorJars | 1.1.0 |
| EmpiricalCalibration | 2.0.2 |
| Eunomia | 1.0.1 |
| EvidenceSynthesis | 0.2.2 |
| FeatureExtraction | 3.1.0 |
| Hydra | 0.1.1 |
| MethodEvaluation | 2.0.0 |
| OhdsiSharing | 0.2.2 |
| ParallelLogger | 2.0.1 |
| PatientLevelPrediction | 4.1.0 |
| ROhdsiWebApi | 1.1.3 |
| SelfControlledCaseSeries | 2.0.0 |
| SelfControlledCohort | 1.5.0 |
| SqlRender | 1.6.8 |

All HADES packages were available

## System Information

Installed R version: R version 4.0.3 (2020-10-10)

System CPU vendor: GenuineIntel

System CPU model: Intel(R) Core(TM) i9-9980HK CPU @ 2.40GHz

System CPU number of cores: 16

System RAM: 68.72 GB

DBMS: sql server

WebAPI version: 2.8.0

## Vocabulary Query Performance

The number of 'Maps To' relations is equal to 3106724. This query was executed in 8.31 secs

## Achilles Query Performance

Table 22. Execution time of queries of the Achilles R-Package

| ID | NAME | DURATION |
| --- | --- | --- |
| 0 | Source name | 0.091678 secs |
| 1 | Number of persons | 0.041696 secs |
| 2 | Number of persons by gender | 0.183182 secs |
| 3 | Number of persons by year of birth | 0.058228 secs |
| 4 | Number of persons by race | 0.059089 secs |
| 5 | Number of persons by ethnicity | 0.070693 secs |
| 7 | Number of persons with invalid provider\_id | 0.056060 secs |
| 8 | Number of persons with invalid location\_id | 0.096566 secs |
| 9 | Number of persons with invalid care\_site\_id | 0.054500 secs |
| 10 | Number of all persons by year of birth by gender | 0.072196 secs |
| 11 | Number of non-deceased persons by year of birth by gender | 0.099605 secs |
| 12 | Number of persons by race and ethnicity | 0.077321 secs |
| 101 | Number of persons by age, with age at first observation period | 0.176635 secs |
| 102 | Number of persons by gender by age, with age at first observation period | 0.159337 secs |
| 108 | Number of persons by length of observation period, in 30d increments | 0.217217 secs |
| 109 | Number of persons with continuous observation in each year | 0.795322 secs |
| 110 | Number of persons with continuous observation in each month | 10.731043 secs |
| 111 | Number of persons by observation period start month | 0.216336 secs |
| 112 | Number of persons by observation period end month | 0.157835 secs |
| 113 | Number of persons by number of observation periods | 0.074423 secs |
| 114 | Number of persons with observation period before year-of-birth | 0.126115 secs |
| 115 | Number of persons with observation period end < observation period start | 0.056148 secs |
| 116 | Number of persons with at least one day of observation in each year by gender and age decile | 1.783371 secs |
| 117 | Number of persons with at least one day of observation in each month | 11.326960 secs |
| 118 | Number of observation periods with invalid person\_id | 0.082853 secs |
| 119 | Number of observation period records by period\_type\_concept\_id | 0.056829 secs |
| 200 | Number of persons with at least one visit occurrence, by visit\_concept\_id | 2.925766 secs |
| 201 | Number of visit occurrence records, by visit\_concept\_id | 2.710461 secs |
| 202 | Number of persons by visit occurrence start month, by visit\_concept\_id | 5.914636 secs |
| 204 | Number of persons with at least one visit occurrence, by visit\_concept\_id by calendar year by gender by age decile | 5.323025 secs |
| 207 | Number of visit records with invalid person\_id | 1.546856 secs |
| 208 | Number of visit records outside valid observation period | 2.198143 secs |
| 209 | Number of visit records with end date < start date | 0.513847 secs |
| 210 | Number of visit records with invalid care\_site\_id | 2.045144 secs |
| 212 | Number of persons with at least one visit occurrence, by calendar year by gender by age decile | 4.888385 secs |
| 220 | Number of visit occurrence records by visit occurrence start month | 3.809848 secs |
| 221 | Number of persons by visit start year | 3.359340 secs |
| 225 | Number of visit\_occurrence records by visit\_source\_concept\_id | 1.327818 secs |
| 300 | Number of providers | 0.144437 secs |
| 301 | Number of providers by specialty concept\_id | 1.195023 secs |
| 302 | Number of providers with invalid care site id | 0.330636 secs |
| 325 | Number of provider records by specialty\_source\_concept\_id | 0.168504 secs |
| 400 | Number of persons with at least one condition occurrence, by condition\_concept\_id | 5.380062 secs |
| 401 | Number of condition occurrence records, by condition\_concept\_id | 2.679631 secs |
| 402 | Number of persons by condition occurrence start month, by condition\_concept\_id | 18.999160 secs |
| 404 | Number of persons with at least one condition occurrence, by condition\_concept\_id by calendar year by gender by age decile | 20.714319 secs |
| 405 | Number of condition occurrence records, by condition\_concept\_id by condition\_type\_concept\_id | 5.871664 secs |
| 409 | Number of condition occurrence records with invalid person\_id | 4.036531 secs |
| 410 | Number of condition occurrence records outside valid observation period | 6.051113 secs |
| 411 | Number of condition occurrence records with end date < start date | 2.057181 secs |
| 412 | Number of condition occurrence records with invalid provider\_id | 5.043638 secs |
| 413 | Number of condition occurrence records with invalid visit\_id | 7.481834 secs |
| 420 | Number of condition occurrence records by condition occurrence start month | 10.181705 secs |
| 424 | Number of co-occurring condition\_occurrence condition\_concept\_id pairs | 2.300317 hours |
| 425 | Number of condition\_occurrence records by condition\_source\_concept\_id | 5.263985 secs |
| 500 | Number of persons with death, by cause\_concept\_id | 0.033817 secs |
| 501 | Number of records of death, by cause\_concept\_id | 0.023165 secs |
| 502 | Number of persons by death month | 0.067278 secs |
| 504 | Number of persons with a death, by calendar year by gender by age decile | 0.071771 secs |
| 505 | Number of death records, by death\_type\_concept\_id | 0.030138 secs |
| 509 | Number of death records with invalid person\_id | 0.049869 secs |
| 510 | Number of death records outside valid observation period | 0.055249 secs |
| 525 | Number of death records by cause\_source\_concept\_id | 0.026784 secs |
| 600 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id | 5.108358 secs |
| 601 | Number of procedure occurrence records, by procedure\_concept\_id | 2.324662 secs |
| 602 | Number of persons by procedure occurrence start month, by procedure\_concept\_id | 16.993680 secs |
| 604 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id by calendar year by gender by age decile | 17.071125 secs |
| 605 | Number of procedure occurrence records, by procedure\_concept\_id by procedure\_type\_concept\_id | 4.307118 secs |
| 609 | Number of procedure occurrence records with invalid person\_id | 3.295144 secs |
| 610 | Number of procedure occurrence records outside valid observation period | 5.463451 secs |
| 612 | Number of procedure occurrence records with invalid provider\_id | 4.679316 secs |
| 613 | Number of procedure occurrence records with invalid visit\_id | 6.219301 secs |
| 620 | Number of procedure occurrence records by procedure occurrence start month | 8.760135 secs |
| 624 | Number of co-occurring procedure\_occurrence procedure\_concept\_id pairs | 55.061515 mins |
| 625 | Number of procedure\_occurrence records by procedure\_source\_concept\_id | 4.217861 secs |
| 691 | Percentage of total persons that have at least x procedures | 6.357057 secs |
| 700 | Number of persons with at least one drug exposure, by drug\_concept\_id | 3.079393 secs |
| 701 | Number of drug exposure records, by drug\_concept\_id | 1.134527 secs |
| 702 | Number of persons by drug exposure start month, by drug\_concept\_id | 11.885105 secs |
| 704 | Number of persons with at least one drug exposure, by drug\_concept\_id by calendar year by gender by age decile | 9.699256 secs |
| 705 | Number of drug exposure records, by drug\_concept\_id by drug\_type\_concept\_id | 2.715549 secs |
| 709 | Number of drug exposure records with invalid person\_id | 1.611035 secs |
| 710 | Number of drug exposure records outside valid observation period | 2.323260 secs |
| 711 | Number of drug exposure records with end date < start date | 1.095796 secs |
| 712 | Number of drug exposure records with invalid provider\_id | 1.582338 secs |
| 713 | Number of drug exposure records with invalid visit\_id | 1.500728 secs |
| 720 | Number of drug exposure records by drug exposure start month | 4.051641 secs |
| 724 | Number of co-occurring drug\_exposure drug\_concept\_id pairs | 1.010575 hours |
| 725 | Number of drug\_exposure records by drug\_source\_concept\_id | 4.621935 secs |
| 791 | Percentage of total persons that have at least x drug exposures | 4.312996 secs |
| 800 | Number of persons with at least one observation occurrence, by observation\_concept\_id | 0.702413 secs |
| 801 | Number of observation occurrence records, by observation\_concept\_id | 0.364348 secs |
| 802 | Number of persons by observation occurrence start month, by observation\_concept\_id | 2.353778 secs |
| 804 | Number of persons with at least one observation occurrence, by observation\_concept\_id by calendar year by gender by age decile | 2.399461 secs |
| 805 | Number of observation occurrence records, by observation\_concept\_id by observation\_type\_concept\_id | 0.548361 secs |
| 807 | Number of observation occurrence records, by observation\_concept\_id and unit\_concept\_id | 0.542313 secs |
| 809 | Number of observation records with invalid person\_id | 0.525063 secs |
| 810 | Number of observation records outside valid observation period | 0.753375 secs |
| 812 | Number of observation records with invalid provider\_id | 0.947995 secs |
| 813 | Number of observation records with invalid visit\_id | 2.219671 secs |
| 814 | Number of observation records with no value (numeric, string, or concept) | 0.193324 secs |
| 820 | Number of observation records by observation start month | 1.674066 secs |
| 822 | Number of observation records, by observation\_concept\_id and value\_as\_concept\_id | 0.687462 secs |
| 823 | Number of observation records, by observation\_concept\_id and qualifier\_concept\_id | 0.621082 secs |
| 824 | Number of co-occurring observation observation\_concept\_id pairs | 43.725993 secs |
| 825 | Number of observation records by observation\_source\_concept\_id | 0.488202 secs |
| 826 | Number of observation records by value\_as\_concept\_id | 0.484036 secs |
| 827 | Number of observation records by unit\_concept\_id | 0.404343 secs |
| 891 | Percentage of total persons that have at least x observations | 1.212092 secs |
| 900 | Number of persons with at least one drug era, by drug\_concept\_id | 2.448727 secs |
| 901 | Number of drug era records, by drug\_concept\_id | 1.043789 secs |
| 902 | Number of persons by drug era start month, by drug\_concept\_id | 9.515116 secs |
| 904 | Number of persons with at least one drug era, by drug\_concept\_id by calendar year by gender by age decile | 9.251103 secs |
| 908 | Number of drug eras without valid person | 1.686578 secs |
| 909 | Number of drug eras outside valid observation period | 2.768495 secs |
| 910 | Number of drug eras with end date < start date | 0.655557 secs |
| 920 | Number of drug era records by drug era start month | 4.186234 secs |
| 1,000 | Number of persons with at least one condition era, by condition\_concept\_id | 4.415717 secs |
| 1,001 | Number of condition era records, by condition\_concept\_id | 1.921693 secs |
| 1,002 | Number of persons by condition era start month, by condition\_concept\_id | 15.667716 secs |
| 1,004 | Number of persons with at least one condition era, by condition\_concept\_id by calendar year by gender by age decile | 15.305212 secs |
| 1,008 | Number of condition eras without valid person | 2.640580 secs |
| 1,009 | Number of condition eras outside valid observation period | 3.393941 secs |
| 1,010 | Number of condition eras with end date < start date | 0.854347 secs |
| 1,020 | Number of condition era records by condition era start month | 6.440133 secs |
| 1,100 | Number of persons by location 3-digit zip | 0.053521 secs |
| 1,101 | Number of persons by location state | 0.077448 secs |
| 1,102 | Number of care sites by location 3-digit zip | 0.119123 secs |
| 1,103 | Number of care sites by location state | 0.069172 secs |
| 1,200 | Number of persons by place of service | 0.043172 secs |
| 1,201 | Number of visits by place of service | 2.278594 secs |
| 1,202 | Number of care sites by place of service | 0.139674 secs |
| 1,203 | Number of visits by place of service discharge type | 0.438741 secs |
| 1,408 | Number of persons by length of payer plan period, in 30d increments | 0.715718 secs |
| 1,409 | Number of persons with continuous payer plan in each year | 3.582449 secs |
| 1,410 | Number of persons with continuous payer plan in each month | 7.786993 secs |
| 1,411 | Number of persons by payer plan period start month | 0.432426 secs |
| 1,412 | Number of persons by payer plan period end month | 0.487946 secs |
| 1,413 | Number of persons by number of payer plan periods | 0.180353 secs |
| 1,414 | Number of persons with payer plan period before year-of-birth | 0.245990 secs |
| 1,415 | Number of persons with payer plan period end < payer plan period start | 0.088444 secs |
| 1,800 | Number of persons with at least one measurement occurrence, by measurement\_concept\_id | 1.195491 secs |
| 1,801 | Number of measurement occurrence records, by measurement\_concept\_id | 0.506576 secs |
| 1,802 | Number of persons by measurement occurrence start month, by measurement\_concept\_id | 4.277425 secs |
| 1,804 | Number of persons with at least one mesurement occurrence, by measurement\_concept\_id by calendar year by gender by age decile | 4.519690 secs |
| 1,805 | Number of measurement occurrence records, by measurement\_concept\_id by measurement\_type\_concept\_id | 0.957127 secs |
| 1,807 | Number of measurement occurrence records, by measurement\_concept\_id and unit\_concept\_id | 0.953142 secs |
| 1,809 | Number of measurement records with invalid person\_id | 0.920115 secs |
| 1,810 | Number of measurement records outside valid observation period | 1.362185 secs |
| 1,812 | Number of measurement records with invalid provider\_id | 1.399858 secs |
| 1,813 | Number of measurement records with invalid visit\_id | 3.019805 secs |
| 1,814 | Number of measurement records with no value (numeric, string, or concept) | 0.319466 secs |
| 1,818 | Number of measurement records below/within/above normal range, by measurement\_concept\_id and unit\_concept\_id | 0.420035 secs |
| 1,820 | Number of measurement records by measurement start month | 1.419406 secs |
| 1,821 | Number of measurement records with no numeric value | 0.488336 secs |
| 1,822 | Number of measurement records, by measurement\_concept\_id and value\_as\_concept\_id | 0.811626 secs |
| 1,823 | Number of measurement records, by measurement\_concept\_id and operator\_concept\_id | 0.921060 secs |
| 1,824 | Number of co-occurring measurement measurement\_concept\_id pairs | 3.858810 mins |
| 1,825 | Number of measurement records by measurement\_source\_concept\_id | 0.831013 secs |
| 1,826 | Number of measurement records by value\_as\_concept\_id | 0.894655 secs |
| 1,827 | Number of measurement records by unit\_concept\_id | 0.921372 secs |
| 1,891 | Percentage of total persons that have at least x measurements | 1.705694 secs |
| 1,900 | Source values mapped to concept\_id 0 by table, by source\_value | 8.943530 secs |
| 2,000 | Number of patients with at least 1 Dx and 1 Rx | 4.858068 secs |
| 2,001 | Number of patients with at least 1 Dx and 1 Proc | 5.085122 secs |
| 2,002 | Number of patients with at least 1 Meas, 1 Dx and 1 Rx | 1.501264 secs |
| 2,003 | Number of patients with at least 1 Visit | 1.083741 secs |
| 2,100 | Number of persons with at least one device exposure, by device\_concept\_id | 0.129376 secs |
| 2,101 | Number of device exposure records, by device\_concept\_id | 0.071823 secs |
| 2,102 | Number of persons by device records start month, by device\_concept\_id | 0.236976 secs |
| 2,104 | Number of persons with at least one device exposure, by device\_concept\_id by calendar year by gender by age decile | 0.313716 secs |
| 2,105 | Number of device exposure records, by device\_concept\_id by device\_type\_concept\_id | 0.111327 secs |
| 2,125 | Number of device\_exposure records by device\_source\_concept\_id | 0.096454 secs |
| 2,200 | Number of persons with at least one note by note\_type\_concept\_id | 0.036714 secs |
| 2,201 | Number of note records, by note\_type\_concept\_id | 0.020697 secs |

Query executed in 0.75 secs

# Scientific Preparedness

This section contains several items related to the interaction with the EHDEN/OHDSI community and training after the mapping process.

## Staff training

Describe how the Data Partner will train and educate the different users of the system in their organizaton and what the current status is of the expertise in the team.

## Study execution

Describe how the Data Partner will be able to execute the ongoing OHDSI/EHDEN network studies, e.g. are there governance issues, lack of resources, etc.

Are there plans to initiate research studies?

Are there plans to participate in OHDSI Working Groups?

# Quality Control

Show that the Data Quality Dashboard results are 100% and check if the thresholds have been changed by doing a diff with the default.

Discuss with the Data Partner why the thresholds have been changed and share this information.

Have the Achilles results been reviewed by the Data Partner?

How is the ETL code tested? Discuss the quality controls steps or ideally share the code that executes this. Have all checks been passed? For example, is there a comparison available of the person count on the source and CDM and are the differences explained?

# Maintenance

Describe briefly the process the Data Partner implemented to keep the data in the OMOP CDM up-to-date when new source data will become available, if the local coding systems are updated, or if new versions of the CDM will be released. Describe how versions of the CDM will be maintained over time.

Describe the maintenance process put in place by the data partner for the tool updates.

# Checklist

Have the following checks been performed?

[ ] ATLAS cohort creation, e.g. Type 2 Diabetes

[ ] Check of Achilles results

Comments:

Check that all the items mentioned below are shared with EHDEN in addition to this inspection report. If items cannot be shared, add an explanation in the comments section.

[ ] ETL Documentation

[ ] ETL Code

[ ] DQD dashboard json file

[ ] White Rabbit output

[ ] CdmInspection results.zip

Comments: