CDM Onboarding report for the Synthea generated for 20k persons database

<author\_1>,<author\_2>

1. Table of content
2. Execution details

| Detail | Value |
| --- | --- |
| CdmOnboarding package version | 3.1.0 |
| Database | postgresql |
| CDM version | 5.4 |
| Execution date | 2024-02-20 |
| Execution duration | 1m 26.5s |
| Achilles version | 1.7 |
| Achilles execution date | 2022-12-04 |

Symbols used in table/figure captions: Ⓞ=Computed directly from OMOP CDM data, Ⓐ=Computed from Achilles results, Ⓢ=Estimated from system tables.

* 1. CDM Source Table

1. Content of the OMOP cdm\_source table Ⓞ

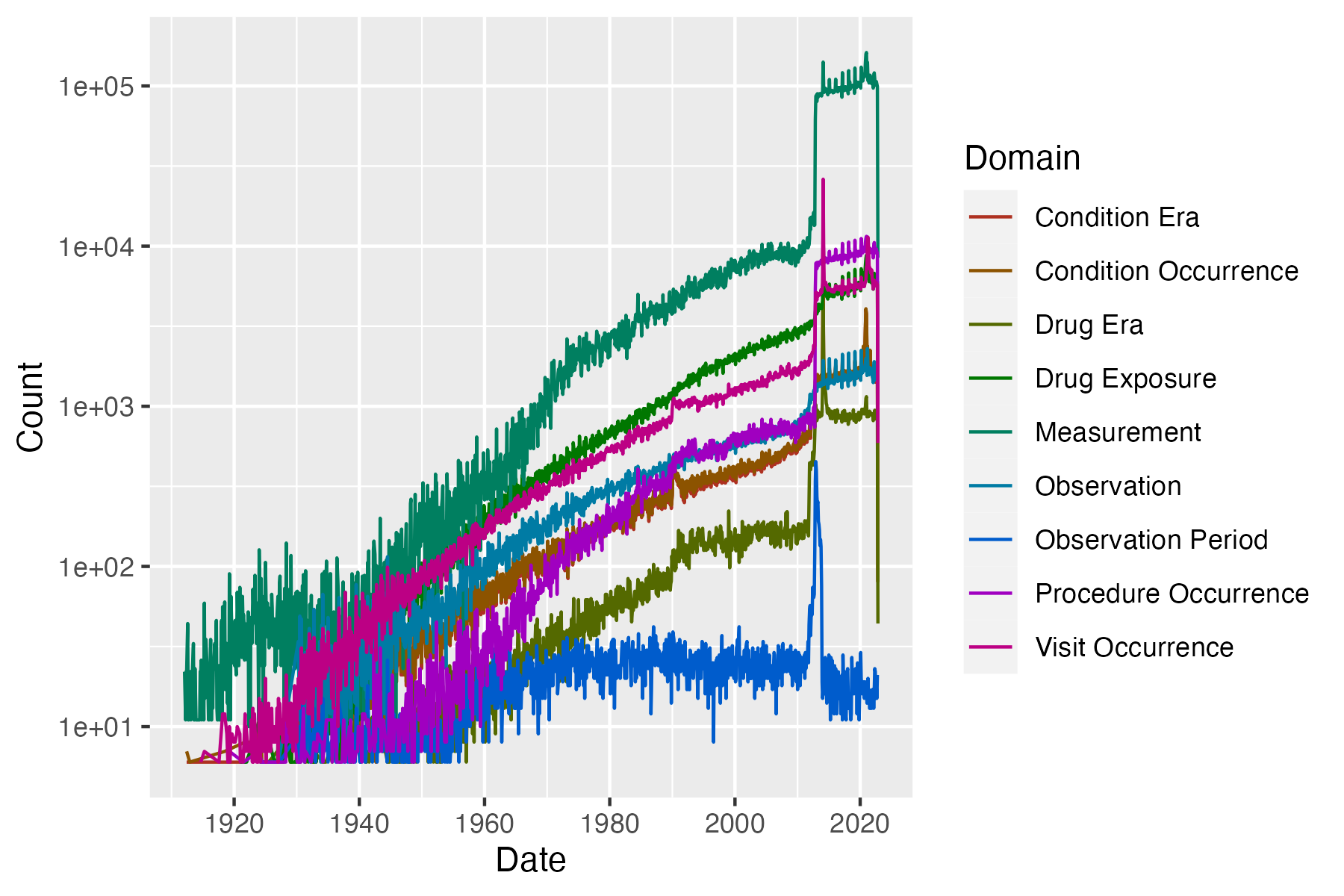
| field | Values |
| --- | --- |
| CDM\_SOURCE\_NAME | Synthea generated for 20k persons |
| CDM\_SOURCE\_ABBREVIATION | Synthea20k |
| CDM\_HOLDER | Maxim |
| SOURCE\_DESCRIPTION | Synthetic generated data for 20k alive patients |
| SOURCE\_DOCUMENTATION\_REFERENCE | https://synthetichealth.github.io/synthea/ |
| CDM\_ETL\_REFERENCE | https://github.com/OHDSI/ETL-Synthea |
| SOURCE\_RELEASE\_DATE | 2022-11-06 |
| CDM\_RELEASE\_DATE | 2022-11-06 |
| CDM\_VERSION | 5.4 |
| CDM\_VERSION\_CONCEPT\_ID | 756265 |
| VOCABULARY\_VERSION | v5.0 31-OCT-22 |

1. Clinical data
   1. Record counts per OMOP CDM table
2. The number of records in all clinical data tables Ⓞ

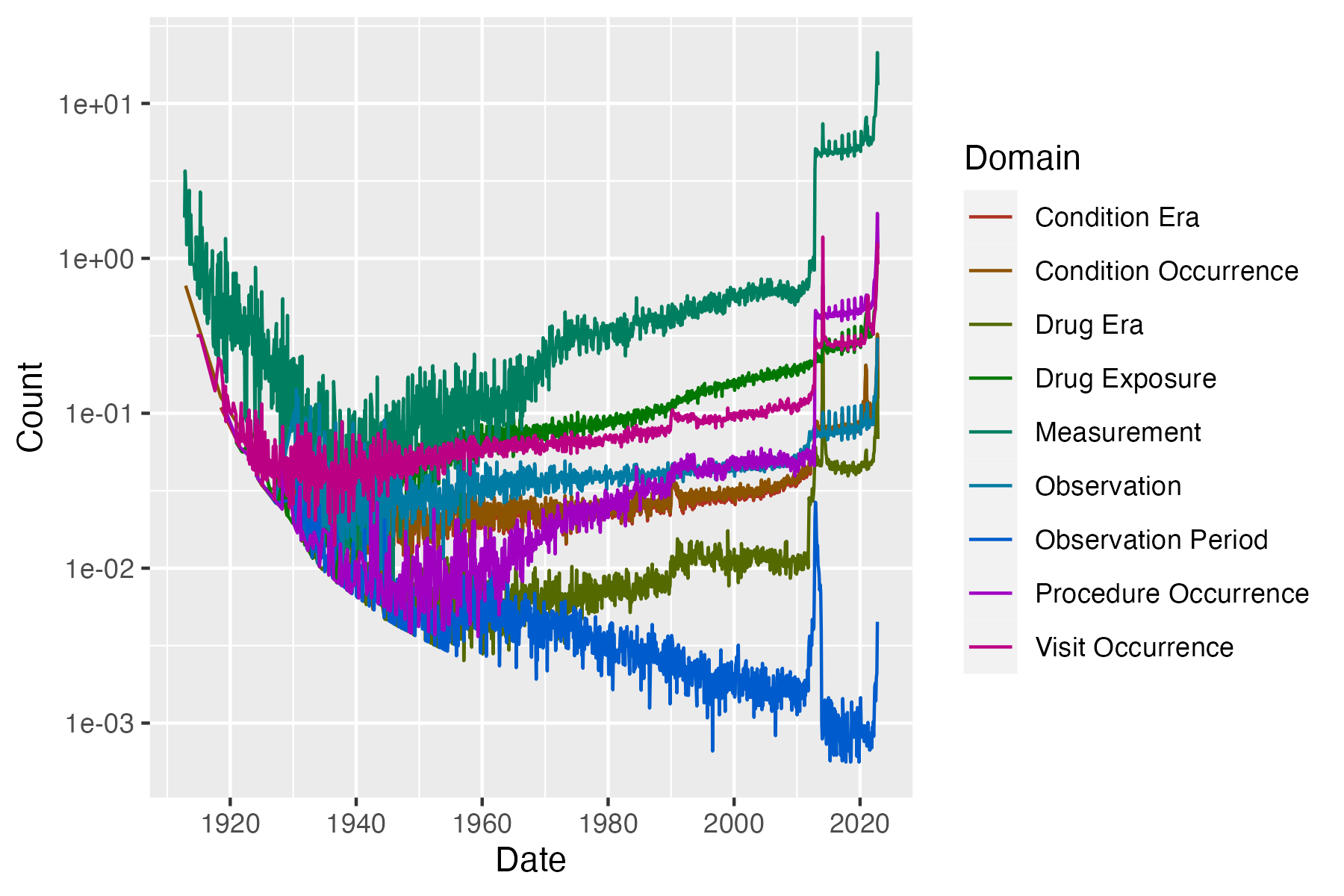
| Table | #Records | #Persons | %Persons |
| --- | --- | --- | --- |
| measurement | 15,072,717 | 23,517 | 100% |
| drug\_exposure | 1,483,660 | 22,478 | 95.6% |
| procedure\_occurrence | 1,296,930 | 23,484 | 99.9% |
| visit\_detail | 1,276,461 | 23,517 | 100% |
| visit\_occurrence | 1,276,461 | 23,517 | 100% |
| observation | 465,207 | 19,815 | 84.3% |
| condition\_occurrence | 402,721 | 23,243 | 98.8% |
| condition\_era | 389,210 | 23,243 | 98.8% |
| drug\_era | 178,000 | 22,478 | 95.6% |
| person | 23,517 | 23,517 | 100% |
| observation\_period | 23,517 | 23,517 | 100% |
| provider | 1,181 | NA | NA |
| cdm\_source | 1 | NA | NA |
| episode\_event | 0 | NA | NA |
| fact\_relationship | 0 | NA | NA |
| episode | 0 | 0 | 0% |
| cost | 0 | NA | NA |
| death | 0 | 0 | 0% |
| device\_exposure | 0 | 0 | 0% |
| dose\_era | 0 | 0 | 0% |
| note | 0 | 0 | 0% |
| payer\_plan\_period | 0 | 0 | 0% |
| specimen | 0 | 0 | 0% |
| care\_site | 0 | NA | NA |
| location | 0 | NA | NA |
| note\_nlp | 0 | NA | NA |
| metadata | 0 | NA | NA |

Query executed in 3.69 seconds

* 1. Data density plots



1. Total record count over time per OMOP data domain. Ⓐ



1. Number of records per person over time per OMOP data domain. Ⓐ

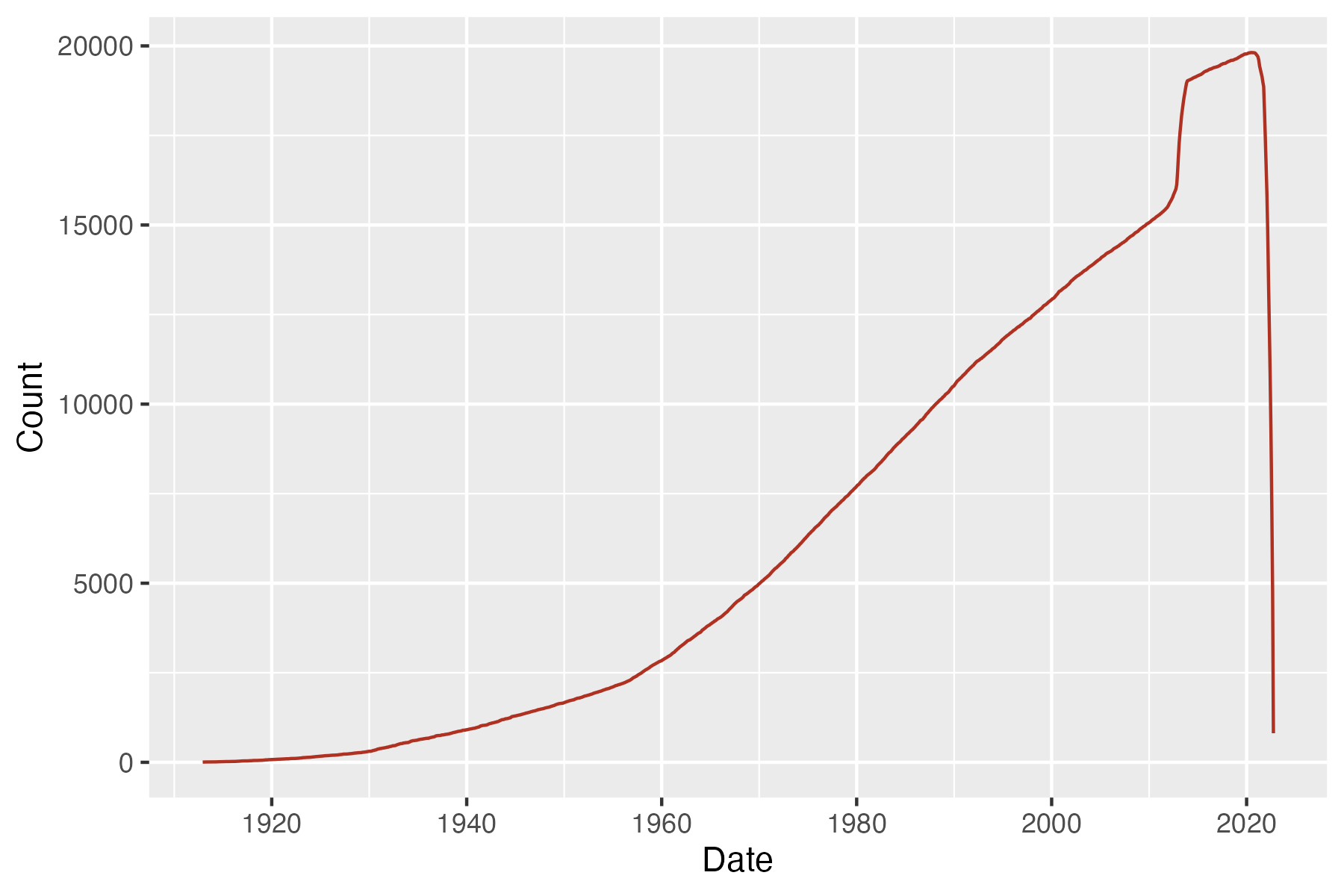
No death records found.

1. Number of deaths in each month. Overall mortality: 0%. Ⓐ
   1. Distinct concepts per person
2. The number of distinct concepts per person per OMOP data domains. Only persons with at least one record in that domain are included in the calculation. Ⓐ

| DOMAIN | N\_PERSONS | MIN | P10 | P25 | MEDIAN | P75 | P90 | MAX |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Visit Occurrence | 23,517 | 1 | 1 | 2 | 2 | 3 | 3 | 3 |
| Condition Occurrence | 23,242 | 1 | 4 | 7 | 10 | 14 | 18 | 45 |
| Procedure Occurrence | 23,484 | 1 | 3 | 6 | 8 | 13 | 19 | 36 |
| Drug exposure | 22,478 | 1 | 2 | 3 | 5 | 7 | 10 | 25 |
| Observation | 19,804 | 1 | 3 | 4 | 6 | 7 | 9 | 21 |
| Drug era | 22,478 | 1 | 2 | 3 | 5 | 8 | 11 | 28 |
| Condition era | 23,242 | 1 | 4 | 7 | 10 | 13 | 18 | 45 |
| Measurement | 23,517 | 1 | 2 | 5 | 6 | 7 | 22 | 36 |

Query executed in 0.03 seconds

* 1. Observation Period



1. Persons with continuous observation by month.Ⓐ In the last 6 months (before 2022-11-06), there are 13,332 persons with an active observation period.Ⓞ
2. Length of first observation period (years). Ⓐ

| AVG | STDEV | MIN | P10 | P25 | MEDIAN | P75 | P90 | MAX |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 30.8 | 21.4 | 0 | 8 | 9.9 | 28.1 | 46.5 | 60.8 | 109.7 |

Query executed in 0.03 seconds

| Field | Value | %Persons |
| --- | --- | --- |
| Persons with 1 observation period(s) | 23,517 | 100 |
| Persons with overlapping observation periods | 0 | NA |
| Number of overlapping observation periods | 0 | NA |

Queries executed in 0.03 seconds and 0.04 seconds

* 1. Date Range

1. Minimum and maximum event start date in each table, within an observation period and at least 5 records. Floored to the nearest month. Ⓐ

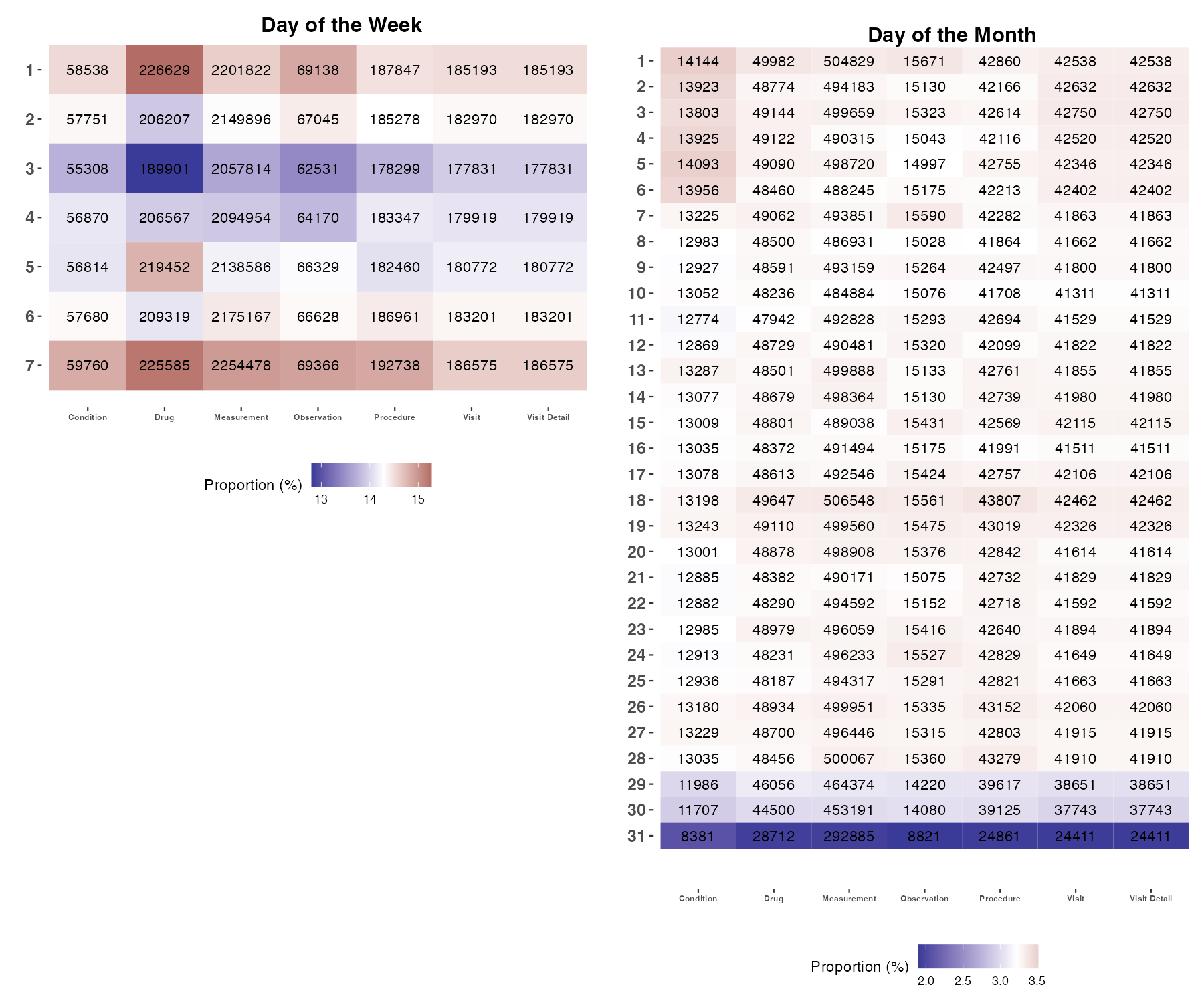
| Domain | N | Type | Start date [Min, Max] | End date [Min, Max] |
| --- | --- | --- | --- | --- |
| Condition | 402,721 | Prescription dispensed in pharmacy (NA) | [1912-06, 2022-11] | [1913-01, 2022-11] |
| Drug | 1,483,660 | Pharmacy claim (S) | [1912-06, 2022-11] | [1912-08, 2022-11] |
| Measurement | 15,072,717 | Outpatient detail - 1st position (NA) | [1912-02, 2022-11] | [NA, NA] |
| Obs. Period | 23,517 | Period covering healthcare encounters (NA) | [1912-02, 2022-10] | [1913-10, 2022-11] |
| Observation | 465,207 | Observation recorded from EHR (NA) | [1913-06, 2022-11] | [NA, NA] |
| Procedure | 1,296,930 | Outpatient detail - 1st position (NA) | [1912-02, 2022-11] | [1912-02, 2022-11] |
| Visit | 1,276,461 | Visit derived from encounter on claim (NA) | [1912-02, 2022-11] | [1912-02, 2022-11] |
| Visit Detail | 1,276,461 | Visit derived from encounter on claim (NA) | [1912-02, 2022-11] | [1912-02, 2022-11] |

Query executed in 3.05 seconds

1. Length of stay by visit concept. The length should be interpreted as number of nights, meaning a length of 0 is a same-day visit. Ⓐ

| Domain | Concept Name | AVG | STDEV | MIN | P10 | P25 | MEDIAN | P75 | P90 | MAX |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Visit | Inpatient Visit | 22.1 | 285.4 | 1 | 1 | 1 | 1 | 1 | 2 | 19,589 |
| Visit | Outpatient Visit | 0.7 | 65.2 | 0 | 0 | 0 | 0 | 0 | 0 | 19,957 |
| Visit | Emergency Room Visit | 2.7 | 210.0 | 0 | 0 | 0 | 0 | 0 | 0 | 28,202 |
| Visit Detail | Inpatient Visit | 22.1 | 285.4 | 1 | 1 | 1 | 1 | 1 | 2 | 19,589 |
| Visit Detail | Outpatient Visit | 0.7 | 65.2 | 0 | 0 | 0 | 0 | 0 | 0 | 19,957 |
| Visit Detail | Emergency Room Visit | 2.7 | 210.0 | 0 | 0 | 0 | 0 | 0 | 0 | 28,202 |

Query executed in 0.24 seconds



1. Day of the Week and Day of the Month distribution of event start dates after 1900-01-01 per domain. 1 = Monday ... 7 = Sunday. Ⓞ

Queries executed in 3.28 seconds and 4.41 seconds

* 1. Day, Month, Year of Birth

1. Distribution of day, month and year of birth of persons. Ⓞ

|  | %Missing | MIN | P10 | P25 | MEDIAN | P75 | P90 | MAX |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | 0% | 1912 | 1943 | 1959 | 1978 | 1999 | 2012 | 2022 |
| Month | 0% | 1 | 2 | 4 | 7 | 10 | 11 | 12 |
| Day | 0% | 1 | 4 | 8 | 16 | 23 | 28 | 31 |

Query executed in 0.04 seconds

1. Vocabulary mappings

Vocabulary version: v5.0 31-OCT-22

Pre-processing query executed in 15.65 seconds

* 1. Mapping Completeness

1. The number and percentage of codes and records that are mapped to an OMOP concept (not 0 and <2B). Note: for one-to-many mappings, the source codes will be counted multiple times so the reported total source codes could be bigger than actual number of unique source codes. Ⓞ

| Domain | #Codes Source | #Codes Mapped | %Codes Mapped | #Records Source | #Records Mapped | %Records Mapped |
| --- | --- | --- | --- | --- | --- | --- |
| Condition | 231 | 222 | 96.1% | 402,721 | 398,328 | 98.9% |
| Condition status | 1 | 0 | 0% | 402,721 | 0 | 0% |
| Death cause | 0 | NA | NA | NA | NA | NA |
| Device | 0 | NA | NA | NA | NA | NA |
| Drug | 256 | 256 | 100% | 1,483,660 | 1,483,660 | 100% |
| Drug Route | 1 | 0 | 0% | 1,483,660 | 0 | 0% |
| Measurement | 331 | 51 | 15.4% | 15,072,717 | 656,289 | 4.4% |
| Measurement Unit | 48 | 32 | 66.7% | 15,072,717 | 6,958,240 | 46.2% |
| Measurement value | 43,822 | 0 | 0% | 15,072,717 | 0 | 0% |
| Observation | 45 | 39 | 86.7% | 465,207 | 463,209 | 99.6% |
| Observation Unit | 1 | 0 | 0% | 465,207 | 0 | 0% |
| Observation value | NA | NA | NA | 465,207 | 0 | 0% |
| Procedure | 176 | 172 | 97.7% | 1,296,930 | 1,285,642 | 99.1% |
| Provider Specialty | 1 | 1 | 100% | 1,181 | 1,181 | 100% |
| Specimen | 0 | NA | NA | NA | NA | NA |
| Visit | 1,276,461 | 1,276,461 | 100% | 1,276,461 | 1,276,461 | 100% |
| Visit detail | 1,276,461 | 1,276,461 | 100% | 1,276,461 | 1,276,461 | 100% |

Query executed in 0.21 seconds

* 1. Drug Mappings

1. The level of the drug mappings Ⓞ

| Class | #Records | #Patients | #Codes | %Records |
| --- | --- | --- | --- | --- |
| Clinical Drug | 1,079,386 | 89,653 | 167 | 72.8% |
| Branded Drug | 169,419 | 7,271 | 19 | 11.4% |
| Quant Branded Drug | 113,674 | 1,622 | 12 | 7.7% |
| Quant Clinical Drug | 80,694 | 9,056 | 40 | 5.4% |
| Clinical Drug Comp | 25,227 | 983 | 2 | 1.7% |
| Branded Pack | 14,893 | 11,547 | 11 | 1.0% |
| Ingredient | 348 | 190 | 1 | 0.0% |
| Branded Drug Comp | 12 | 7 | 2 | 0.0% |
| Branded Drug Form | 7 | 5 | 1 | 0.0% |

Query executed in 0.02 seconds

* 1. Unmapped Codes

1. Omitted because no unmapped drugs were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. All 7 unmapped conditions. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 | 428251008 | 1,100 | 0.3% |
| 2 | 55680006 | 1,100 | 0.3% |
| 3 | 703151001 | 900 | 0.2% |
| 4 | 429007001 | 900 | 0.2% |
| 5 | 399211009 | 600 | 0.1% |
| 6 | 161622006 | 100 | 0.0% |
| 7 | 429280009 | 100 | 0.0% |

Query executed in 0.01 seconds

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

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 | 72514-3||Pain severity - 0-10 verbal numeric ratin | 406,900 | 2.7% |
| 2 | 8480-6||Systolic Blood Pressure | 324,600 | 2.2% |
| 3 | 8462-4||Diastolic Blood Pressure | 324,600 | 2.2% |
| 4 | 29463-7||Body Weight | 299,600 | 2.0% |
| 5 | 9279-1||Respiratory rate | 295,100 | 2.0% |
| 6 | 8867-4||Heart rate | 295,100 | 2.0% |
| 7 | 8302-2||Body Height | 288,900 | 1.9% |
| 8 | 72166-2||Tobacco smoking status | 287,900 | 1.9% |
| 9 | 39156-5||Body Mass Index | 265,000 | 1.8% |
| 10 | 6299-2||Urea Nitrogen | 245,300 | 1.6% |
| 11 | 6298-4||Potassium | 245,300 | 1.6% |
| 12 | 2339-0||Glucose | 245,300 | 1.6% |
| 13 | 38483-4||Creatinine | 245,300 | 1.6% |
| 14 | 20565-8||Carbon Dioxide | 245,300 | 1.6% |
| 15 | 2947-0||Sodium | 245,300 | 1.6% |
| 16 | 49765-1||Calcium | 245,300 | 1.6% |
| 17 | 2069-3||Chloride | 245,300 | 1.6% |
| 18 | 93033-9||Are you worried about losing your housing | 212,800 | 1.4% |
| 19 | 63586-2||During the past year what was the total | 212,800 | 1.4% |
| 20 | 63512-8||How many family members including yourse | 212,800 | 1.4% |
| 21 | 93038-8||Stress is when someone feels tense nervo | 212,800 | 1.4% |
| 22 | 67875-5||What is your current work situation? | 212,800 | 1.4% |
| 23 | 93035-4||At any point in the past 2 years has sea | 212,800 | 1.4% |
| 24 | 93034-7||Have you been discharged from the armed f | 212,800 | 1.4% |
| 25 | 93031-3||In the past year have you or any family | 212,800 | 1.4% |

Query executed in 0.01 seconds

1. All 6 unmapped observations. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 | 1191||Aspirin | 800 | 0.2% |
| 2 | 29046||Lisinopril | 500 | 0.1% |
| 3 | 7984||Penicillin V | 500 | 0.1% |
| 4 | 10831||Sulfamethoxazole / Trimethoprim | 200 | 0.0% |
| 5 | 25037||cefdinir | 100 | 0.0% |
| 6 | 5640||Ibuprofen | 100 | 0.0% |

Query executed in 0.01 seconds

1. All 3 unmapped procedures. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 | 399014008 | 6,300 | 0.5% |
| 2 | 66348005 | 5,100 | 0.4% |
| 3 | 71121004 | 100 | 0.0% |

Query executed in 0.01 seconds

1. Omitted because no unmapped devices were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. Omitted because no unmapped visits were found with a count >5. Ⓞ

Query executed in 0.08 seconds

1. Omitted because no unmapped visit details were found with a count >5. Ⓞ

Query executed in 0.08 seconds

1. All 16 unmapped measurement units. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 |  | 5,063,900 | 33.6% |
| 2 | {nominal} | 1,181,100 | 7.8% |
| 3 | {score} | 1,179,000 | 7.8% |
| 4 | {#} | 263,300 | 1.7% |
| 5 | /a | 212,800 | 1.4% |
| 6 | U/L | 143,900 | 1.0% |
| 7 | mL/min/{1.73\_m2} | 30,400 | 0.2% |
| 8 | kU/L | 24,600 | 0.2% |
| 9 | n/a | 7,100 | 0.0% |
| 10 | {INR} | 2,900 | 0.0% |
| 11 | [pH] | 2,600 | 0.0% |
| 12 | {T-score} | 1,700 | 0.0% |
| 13 | {count} | 1,600 | 0.0% |
| 14 | ratio | 200 | 0.0% |
| 15 | m[IU]/L | 100 | 0.0% |
| 16 | ng/dl | 100 | 0.0% |

Query executed in 0.01 seconds

1. All 1 unmapped observation units. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 |  | 465,300 | 100% |

Query executed in 0.01 seconds

1. Top 25 unmapped measurement values. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 | No | 1,506,900 | 10.0% |
| 2 |  | 656,300 | 4.4% |
| 3 | 1.0 | 516,500 | 3.4% |
| 4 | Yes | 376,900 | 2.5% |
| 5 | 0.0 | 273,000 | 1.8% |
| 6 | 2.0 | 264,700 | 1.8% |
| 7 | Never smoked tobacco (finding) | 214,100 | 1.4% |
| 8 | I have housing | 211,500 | 1.4% |
| 9 | English | 194,100 | 1.3% |
| 10 | 3.0 | 189,300 | 1.3% |
| 11 | White | 175,900 | 1.2% |
| 12 | I choose not to answer this question | 162,000 | 1.1% |
| 13 | 4.0 | 149,000 | 1.0% |
| 14 | Full-time work | 135,400 | 0.9% |
| 15 | More than high school | 124,000 | 0.8% |
| 16 | Cloudy urine (finding) | 97,800 | 0.6% |
| 17 | Urine glucose test = ++ (finding) | 97,500 | 0.6% |
| 18 | Urine nitrite negative (finding) | 97,500 | 0.6% |
| 19 | Finding of bilirubin in urine (finding) | 97,500 | 0.6% |
| 20 | Urine leukocyte test negative (finding) | 97,500 | 0.6% |
| 21 | Brown color (qualifier value) | 97,200 | 0.6% |
| 22 | Urine smell ammoniacal (finding) | 97,200 | 0.6% |
| 23 | Translucent (qualifier value) | 97,200 | 0.6% |
| 24 | Urine blood test = negative (finding) | 97,200 | 0.6% |
| 25 | 5.0 | 97,100 | 0.6% |

Query executed in 0.02 seconds

1. All 1 unmapped observation values. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 |  | 465,300 | 100% |

Query executed in 0.01 seconds

1. All 1 unmapped drug route. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Source Value | #Records | %Records |
| --- | --- | --- | --- |
| 1 |  | 1,483,700 | 100% |

Query executed in 0.01 seconds

1. Omitted because no unmapped specialty were found with a count >5. Ⓞ

Query executed in 0.01 seconds

* 1. Mapped Codes

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

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 19080128 | lisinopril 10 MG Oral Tablet | 228,200 | 15.4% |
| 2 | 19078106 | hydrochlorothiazide 25 MG Oral Tablet | 188,600 | 12.7% |
| 3 | 19073094 | amlodipine 2.5 MG Oral Tablet | 179,000 | 12.1% |
| 4 | 19009384 | insulin isophane, human 70 UNT/ML / insulin, regular, human 30 UNT/ML Injectable Suspension [Humulin] | 120,600 | 8.1% |
| 5 | 19030765 | 1 ML epoetin alfa 4000 UNT/ML Injection [Epogen] | 88,100 | 5.9% |
| 6 | 1511248 | NDA020503 200 ACTUAT albuterol 0.09 MG/ACTUAT Metered Dose Inhaler | 51,500 | 3.5% |
| 7 | 19126352 | nitroglycerin 0.4 MG/ACTUAT Mucosal Spray | 41,900 | 2.8% |
| 8 | 1539463 | simvastatin 10 MG Oral Tablet | 39,800 | 2.7% |
| 9 | 1539411 | simvastatin 20 MG Oral Tablet | 36,600 | 2.5% |
| 10 | 40163924 | 24 HR metformin hydrochloride 500 MG Extended Release Oral Tablet | 35,600 | 2.4% |
| 11 | 1332419 | amlodipine 5 MG Oral Tablet | 34,700 | 2.3% |
| 12 | 19075601 | clopidogrel 75 MG Oral Tablet | 32,900 | 2.2% |
| 13 | 40163554 | warfarin sodium 5 MG Oral Tablet | 25,900 | 1.7% |
| 14 | 19018935 | digoxin 0.125 MG Oral Tablet | 25,900 | 1.7% |
| 15 | 40171902 | verapamil hydrochloride 40 MG | 25,100 | 1.7% |
| 16 | 40173590 | alendronic acid 10 MG Oral Tablet | 20,300 | 1.4% |
| 17 | 40169217 | 120 ACTUAT fluticasone propionate 0.044 MG/ACTUAT Metered Dose Inhaler [Flovent] | 16,500 | 1.1% |
| 18 | 1154380 | albuterol 5 MG/ML Inhalation Solution | 15,900 | 1.1% |
| 19 | 40169281 | 60 ACTUAT fluticasone propionate 0.25 MG/ACTUAT / salmeterol 0.05 MG/ACTUAT Dry Powder Inhaler | 15,700 | 1.1% |
| 20 | 40162606 | acetaminophen 300 MG / hydrocodone bitartrate 5 MG Oral Tablet | 14,600 | 1.0% |
| 21 | 19098138 | budesonide 0.25 MG/ML Inhalation Suspension | 13,400 | 0.9% |
| 22 | 1127433 | acetaminophen 325 MG Oral Tablet | 12,600 | 0.8% |
| 23 | 19033893 | ibuprofen 400 MG Oral Tablet [Ibu] | 12,300 | 0.8% |
| 24 | 19041324 | acetaminophen 325 MG Oral Tablet [Tylenol] | 12,100 | 0.8% |
| 25 | 35604087 | cisplatin 50 MG Injection | 11,200 | 0.7% |

Query executed in 0.01 seconds

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

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 4251306 | Stress | 110,800 | 27.5% |
| 2 | 4309238 | Social isolation | 27,400 | 6.8% |
| 3 | 4172829 | Limited social contact | 27,300 | 6.8% |
| 4 | 40481087 | Viral sinusitis | 24,600 | 6.1% |
| 5 | 45768458 | Victim of intimate partner abuse | 19,100 | 4.7% |
| 6 | 4112343 | Acute viral pharyngitis | 14,100 | 3.5% |
| 7 | 260139 | Acute bronchitis | 11,600 | 2.9% |
| 8 | 4217975 | Normal pregnancy | 10,600 | 2.6% |
| 9 | 4311629 | Impaired glucose tolerance | 9,000 | 2.2% |
| 10 | 439777 | Anemia | 8,200 | 2.0% |
| 11 | 320128 | Essential hypertension | 8,100 | 2.0% |
| 12 | 4059650 | Risk activity involvement | 7,500 | 1.8% |
| 13 | 257012 | Chronic sinusitis | 5,000 | 1.2% |
| 14 | 436940 | Metabolic syndrome X | 4,300 | 1.1% |
| 15 | 4132891 | Chronic low back pain | 4,100 | 1.0% |
| 16 | 372328 | Otitis media | 4,000 | 1.0% |
| 17 | 4078393 | Miscarriage in first trimester | 3,800 | 0.9% |
| 18 | 28060 | Streptococcal sore throat | 3,400 | 0.8% |
| 19 | 140352 | Acute myeloid leukemia, disease | 3,100 | 0.8% |
| 20 | 4214746 | Severe anxiety (panic) | 3,000 | 0.7% |
| 21 | 43530622 | Chronic neck pain | 2,600 | 0.6% |
| 22 | 432867 | Hyperlipidemia | 2,500 | 0.6% |
| 23 | 81151 | Sprain of ankle | 2,400 | 0.6% |
| 24 | 4250734 | Febrile neutropenia | 2,400 | 0.6% |
| 25 | 192279 | Disorder of kidney due to diabetes mellitus | 2,300 | 0.6% |

Query executed in 0.02 seconds

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

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 4064377 | Depression screening | 189,600 | 1.3% |
| 2 | 765449 | Assessment using PHQ-2 questionnaire | 164,700 | 1.1% |
| 3 | 37016945 | Screening for drug abuse | 64,800 | 0.4% |
| 4 | 3661628 | Screening for domestic abuse | 64,200 | 0.4% |
| 5 | 37395717 | Depression screening using PHQ-9 (Patient Health Questionnaire 9) score | 25,300 | 0.2% |
| 6 | 4015052 | Hemogram, automated, with RBC, WBC, Hgb, Hct, Indices, Platelet count, and manual WBC differential | 12,700 | 0.1% |
| 7 | 4091461 | Standard pregnancy test | 9,700 | 0.1% |
| 8 | 4133840 | Spirometry | 8,200 | 0.1% |
| 9 | 4024958 | Throat culture | 6,600 | 0.0% |
| 10 | 4018595 | Measurement of Varicella-zoster virus antibody | 6,500 | 0.0% |
| 11 | 4102154 | Skin test for tuberculosis, Tine test | 6,500 | 0.0% |
| 12 | 4211845 | Urine protein measurement | 6,500 | 0.0% |
| 13 | 4191851 | Urine screening test for diabetes | 6,500 | 0.0% |
| 14 | 4136494 | HIV antigen test | 6,500 | 0.0% |
| 15 | 4149631 | Chlamydia antigen test | 6,500 | 0.0% |
| 16 | 4151408 | Syphilis infectious titer test | 6,500 | 0.0% |
| 17 | 4258677 | Blood group typing | 6,500 | 0.0% |
| 18 | 4018603 | Hepatitis C antibody, confirmatory test | 6,500 | 0.0% |
| 19 | 4014117 | Gonorrhea infection titer test | 6,500 | 0.0% |
| 20 | 4024509 | Urine culture | 6,500 | 0.0% |
| 21 | 4168615 | Hepatitis B surface antigen measurement | 6,500 | 0.0% |
| 22 | 4149386 | Urine screening for glucose | 6,300 | 0.0% |
| 23 | 4078285 | Alpha-fetoprotein test - antenatal | 6,300 | 0.0% |
| 24 | 4052083 | Measurement of respiratory function | 6,100 | 0.0% |
| 25 | 4304821 | Streptococcus pneumoniae group B antigen assay | 5,800 | 0.0% |

Query executed in 0.01 seconds

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

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 4053118 | Full-time employment | 305,500 | 65.6% |
| 2 | 4059634 | Part-time employment | 54,200 | 11.6% |
| 3 | 42538208 | Not in labor force | 24,400 | 5.2% |
| 4 | 4313322 | Reports of violence in the environment | 12,500 | 2.7% |
| 5 | 4076230 | Received higher education | 11,000 | 2.4% |
| 6 | 4060985 | Body mass index 30+ - obesity | 10,800 | 2.3% |
| 7 | 765457 | Received certificate of high school equivalency | 5,500 | 1.2% |
| 8 | 4145802 | Has a criminal record | 4,600 | 1.0% |
| 9 | 36712761 | Unhealthy alcohol drinking behavior | 4,000 | 0.9% |
| 10 | 4022661 | Housing problem | 3,000 | 0.6% |
| 11 | 4072733 | Only received primary school education | 2,300 | 0.5% |
| 12 | 37017222 | Lack of access to transportation | 2,300 | 0.5% |
| 13 | 4141789 | Transport problems | 2,200 | 0.5% |
| 14 | 4224654 | Mold | 2,200 | 0.5% |
| 15 | 4138133 | Animal dander | 2,100 | 0.4% |
| 16 | 37311060 | Suspected COVID-19 | 2,100 | 0.4% |
| 17 | 36684363 | Tree pollen | 1,700 | 0.4% |
| 18 | 4117479 | Grass pollen | 1,700 | 0.4% |
| 19 | 4251171 | Unemployed | 1,600 | 0.3% |
| 20 | 4125382 | House dust mite | 1,600 | 0.3% |
| 21 | 4074924 | Served in armed forces | 1,300 | 0.3% |
| 22 | 40488824 | Refugee | 1,100 | 0.2% |
| 23 | 42536288 | Shellfish | 900 | 0.2% |
| 24 | 4008070 | Latex | 700 | 0.1% |
| 25 | 42538933 | Peanut | 700 | 0.1% |

Query executed in 0.01 seconds

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

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 46272459 | Assessment of health and social care needs | 213,200 | 16.4% |
| 2 | 762506 | Assessment of substance use | 143,500 | 11.1% |
| 3 | 4326177 | Medication reconciliation | 136,500 | 10.5% |
| 4 | 46272472 | Assessment of anxiety | 113,500 | 8.7% |
| 5 | 4146536 | Renal dialysis | 103,300 | 8.0% |
| 6 | 35621997 | Assessment using AUDIT-C (Alcohol Use Disorders Identification Test - Consumption) | 65,300 | 5.0% |
| 7 | 4075363 | Auscultation of the fetal heart | 56,400 | 4.3% |
| 8 | 4168236 | Evaluation of uterine fundal height | 56,400 | 4.3% |
| 9 | 35621755 | Assessment using Morse Fall Scale | 50,300 | 3.9% |
| 10 | 4107731 | Subcutaneous immunotherapy | 31,100 | 2.4% |
| 11 | 4295880 | Intramuscular injection | 27,900 | 2.1% |
| 12 | 4078793 | Direct current cardioversion | 22,300 | 1.7% |
| 13 | 4240345 | Physical examination | 16,100 | 1.2% |
| 14 | 3655745 | Assessment using Car, Relax, Alone, Forget, Friends, Trouble Screening Test | 13,500 | 1.0% |
| 15 | 4016390 | Sleep apnea assessment | 13,400 | 1.0% |
| 16 | 45766298 | Combined chemotherapy and radiation therapy | 12,500 | 1.0% |
| 17 | 4249893 | Colonoscopy | 11,200 | 0.9% |
| 18 | 4298386 | Anticipatory guidance | 10,600 | 0.8% |
| 19 | 4060626 | Ultrasound scan for fetal viability | 9,700 | 0.7% |
| 20 | 4234541 | Cytopathology procedure, preparation of smear, genital source | 7,900 | 0.6% |
| 21 | 4162736 | Oxygen administration by mask | 7,700 | 0.6% |
| 22 | 4196006 | Placing subject in prone position | 7,500 | 0.6% |
| 23 | 4153050 | Development of individualized plan of care | 7,400 | 0.6% |
| 24 | 4151779 | Initial patient assessment | 7,400 | 0.6% |
| 25 | 4043071 | Cognitive and behavioral therapy | 7,300 | 0.6% |

Query executed in 0.02 seconds

1. Omitted because no mapped devices were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. All 3 mapped visits. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 9202 | Outpatient Visit | 1,198,800 | 93.9% |
| 2 | 9203 | Emergency Room Visit | 42,000 | 3.3% |
| 3 | 9201 | Inpatient Visit | 35,800 | 2.8% |

Query executed in 0.37 seconds

1. All 3 mapped visit details. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 9202 | Outpatient Visit | 1,198,800 | 93.9% |
| 2 | 9203 | Emergency Room Visit | 42,000 | 3.3% |
| 3 | 9201 | Inpatient Visit | 35,800 | 2.8% |

Query executed in 0.36 seconds

1. Top 25 mapped measurement units. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 8840 | milligram per deciliter | 1,823,500 | 12.1% |
| 2 | 8753 | millimole per liter | 1,018,800 | 6.8% |
| 3 | 8876 | millimeter mercury column | 654,500 | 4.3% |
| 4 | 8541 | per minute | 590,100 | 3.9% |
| 5 | 9448 | year | 425,300 | 2.8% |
| 6 | 9529 | kilogram | 402,900 | 2.7% |
| 7 | 8554 | percent | 383,300 | 2.5% |
| 8 | 8582 | centimeter | 323,700 | 2.1% |
| 9 | 9531 | kilogram per square meter | 265,000 | 1.8% |
| 10 | 8583 | femtoliter | 219,300 | 1.5% |
| 11 | 8713 | gram per deciliter | 215,100 | 1.4% |
| 12 | 8848 | thousand per microliter | 151,600 | 1.0% |
| 13 | 8795 | milliliter per minute | 127,100 | 0.8% |
| 14 | 8482 | pH | 97,500 | 0.6% |
| 15 | 8815 | million per microliter | 61,400 | 0.4% |
| 16 | 8564 | picogram | 61,400 | 0.4% |
| 17 | 8636 | gram per liter | 37,400 | 0.2% |
| 18 | 8723 | milligram per gram | 30,400 | 0.2% |
| 19 | 586323 | degree Celsius | 25,200 | 0.2% |
| 20 | 8845 | picogram per milliliter | 10,000 | 0.1% |
| 21 | 8837 | microgram per deciliter | 7,100 | 0.0% |
| 22 | 8842 | nanogram per milliliter | 6,800 | 0.0% |
| 23 | 8748 | microgram per liter | 6,500 | 0.0% |
| 24 | 9289 | degree Fahrenheit | 3,100 | 0.0% |
| 25 | 8751 | milligram per liter | 2,900 | 0.0% |

Query executed in 0.01 seconds

1. Omitted because no mapped observation units were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. Omitted because no mapped measurement values were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. Omitted because no mapped observation values were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. Omitted because no mapped drug route were found with a count >5. Ⓞ

Query executed in 0.01 seconds

1. All 1 mapped specialty. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted. Ⓞ

| # | Concept id | Concept Name | #Records | %Records |
| --- | --- | --- | --- | --- |
| 1 | 38004446 | NA | 1,200 | 100% |

Query executed in 0.01 seconds

* 1. Source to concept map

1. Source to concept map breakdown Ⓞ

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

Query executed in 0.02 seconds

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

1. Data Quality Dashboard

DataQualityDashboard Version: 2.0.0

DataQualityDashboard executed at 2022-12-05 11:52:57 in 8 mins.

1. Number of passed, failed and total DQD checks per category. For DQD v2, the checks with status 'NA' are not included.

| Category | Pass | Fail | Total | %Pass |
| --- | --- | --- | --- | --- |
| Plausibility | 141 | 8 | 2,500 | 5.6% |
| Conformance | 791 | 3 | 1,072 | 73.8% |
| Completeness | 202 | 55 | 464 | 43.5% |
| Total | 3,970 | 66 | 4,036 | 98.4% |

1. Drug Exposure Diagnostics
2. Drug Exposure Diagnostics results for selected ingredients. Executed with minCellCount = 5, sample = 1e+06, earliestStartDate = 2010-01-01. # = Number of records. Type (n,%) = Frequency and percentage of available drug types. Route (n,%) = Frequency and percentage of available routes. Dose Form present n (%) = Frequency and percentage with dose form present. Fixed amount dose form n (%) = Frequency and percentage of missing denominator unit concept id. Amount distrib. [null or missing] = Distribution of amount (median q05-q95), frequency and percentage of null or missing amount. Quantity distrib. [null or missing] = Distribution of quantity (median q05-q95), frequency and percentage of null or missing quantity. Exposure days distrib. [null or missing] = Distribution of exposure days (median q05-q95), frequency and percentage of null days\_supply or missing exposure dates. Neg. Days n (%) = Frequency and percentage of negative exposure days. Ⓞ

| Ingredient | Concept ID | # | Type (n,%) | Route (n,%) | Dose Form present n (%) | Fixed amount dose form n (%) | Amount distrib. [null or missing] | Quantity distrib. [null or missing] | Exposure days distrib. [null or missing] | Neg. Days n (%) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| acetaminophen | 1125315 | 68266 | Pharmacy claim (63928, 100%) | No matching concept (63928, 100%) | 0 (0%) | 0 (0%) | 300 (0-500) [7659, 7.7%] | 0 (0-0) [99853, 100%] | 36 (1-372) [15274, 15.3%] | 5 (0%) |
| albuterol | 1154343 | 67308 | Pharmacy claim (25438, 100%) | No matching concept (25438, 100%) | 0 (0%) | 0 (0%) | 0 (0-0) [25438, 100%] | 0 (0-0) [25438, 100%] | 239 (1-372) [4673, 18.4%] | 2 (0%) |

Query executed in 44.96 seconds

1. Technical Infrastructure
   1. R packages
2. Versions of all installed R packages from DARWIN EU® and the OHDSI Health Analytics Data-to-Evidence Suite (HADES). Packages can be installed from CRAN (install.packages("<package\_name>")) or Github (remotes::install\_github("<organisation>/<package>")) Ⓢ

| Organisation | Package | Version |
| --- | --- | --- |
| DARWIN EU® | CDMConnector | 1.3.0 |
| DARWIN EU® | CdmOnboarding | 3.1.0 |
| DARWIN EU® | DashboardExport | 1.2.0 |
| DARWIN EU® | DrugExposureDiagnostics | 1.0.3 |
| DARWIN EU® | TreatmentPatterns | 2.6.2 |
| DARWIN EU® | CodelistGenerator | Not installed |
| DARWIN EU® | CohortSurvival | Not installed |
| DARWIN EU® | DrugUtilisation | Not installed |
| DARWIN EU® | IncidencePrevalence | Not installed |
| DARWIN EU® | OMOPGenerics | Not installed |
| DARWIN EU® | PaRe | Not installed |
| DARWIN EU® | PatientProfiles | Not installed |
| DARWIN EU® | ReportGenerator | Not installed |
| DARWIN EU® | deckR | Not installed |
| OHDSI HADES | Achilles | 1.7.2 |
| OHDSI HADES | Andromeda | 0.6.3 |
| OHDSI HADES | CirceR | 1.3.1 |
| OHDSI HADES | CohortDiagnostics | 3.2.4 |
| OHDSI HADES | CohortGenerator | 0.8.1 |
| OHDSI HADES | DataQualityDashboard | 2.5.0 |
| OHDSI HADES | DatabaseConnector | 6.3.2 |
| OHDSI HADES | Eunomia | 1.0.2 |
| OHDSI HADES | FeatureExtraction | 3.3.2 |
| OHDSI HADES | OhdsiSharing | 0.2.2 |
| OHDSI HADES | OhdsiShinyModules | 2.0.0 |
| OHDSI HADES | ParallelLogger | 3.3.0 |
| OHDSI HADES | ROhdsiWebApi | 1.3.3 |
| OHDSI HADES | ResultModelManager | 0.5.3 |
| OHDSI HADES | SqlRender | 1.16.1 |
| OHDSI HADES | BigKnn | Not installed |
| OHDSI HADES | BrokenAdaptiveRidge | Not installed |
| OHDSI HADES | Capr | Not installed |
| OHDSI HADES | Characterization | Not installed |
| OHDSI HADES | CohortExplorer | Not installed |
| OHDSI HADES | CohortMethod | Not installed |
| OHDSI HADES | Cyclops | Not installed |
| OHDSI HADES | DeepPatientLevelPrediction | Not installed |
| OHDSI HADES | EmpiricalCalibration | Not installed |
| OHDSI HADES | EnsemblePatientLevelPrediction | Not installed |
| OHDSI HADES | EvidenceSynthesis | Not installed |
| OHDSI HADES | Hydra | Not installed |
| OHDSI HADES | IterativeHardThresholding | Not installed |
| OHDSI HADES | MethodEvaluation | Not installed |
| OHDSI HADES | PatientLevelPrediction | Not installed |
| OHDSI HADES | PheValuator | Not installed |
| OHDSI HADES | PhenotypeLibrary | Not installed |
| OHDSI HADES | SelfControlledCaseSeries | Not installed |
| OHDSI HADES | SelfControlledCohort | Not installed |
| OHDSI HADES | ShinyAppBuilder | Not installed |

* 1. System Information

Installed R version: R version 4.3.2 (2023-10-31)

System CPU vendor:

System CPU model: Apple M1 Pro

System CPU number of cores: 10

System RAM: 17.18 GB

DBMS: PostgreSQL 14.6 on aarch64-apple-darwin20.6.0, compiled by Apple clang version 12.0.5 (clang-1205.0.22.9), 64-bit

WebAPI version: 2.12.1

* 1. Vocabulary Query Performance

The number of 'Maps To' relations is equal to 1,117,767 and queried in 3.00 seconds (2.68227e-06 s/#).

* 1. Applied indexes

1. The indexes applied on the OMOP CDM tables Ⓢ

| TABLENAME | INDEXNAMES |
| --- | --- |
| care\_site | idx\_care\_site\_id\_1,xpk\_care\_site |
| condition\_era | idx\_condition\_era\_concept\_id\_1,idx\_condition\_era\_person\_id\_1,xpk\_condition\_era |
| condition\_occurrence | idx\_condition\_concept\_id\_1,idx\_condition\_person\_id\_1,idx\_condition\_visit\_id\_1,xpk\_condition\_occurrence |
| cost | idx\_cost\_event\_id,xpk\_cost |
| death | idx\_death\_person\_id\_1 |
| device\_exposure | idx\_device\_concept\_id\_1,idx\_device\_person\_id\_1,idx\_device\_visit\_id\_1,xpk\_device\_exposure |
| dose\_era | idx\_dose\_era\_concept\_id\_1,idx\_dose\_era\_person\_id\_1,xpk\_dose\_era |
| drug\_era | idx\_drug\_era\_concept\_id\_1,idx\_drug\_era\_person\_id\_1,xpk\_drug\_era |
| drug\_exposure | idx\_drug\_concept\_id\_1,idx\_drug\_person\_id\_1,idx\_drug\_visit\_id\_1,xpk\_drug\_exposure |
| episode | xpk\_episode |
| fact\_relationship | idx\_fact\_relationship\_id1,idx\_fact\_relationship\_id2,idx\_fact\_relationship\_id3 |
| location | idx\_location\_id\_1,xpk\_location |
| measurement | idx\_measurement\_concept\_id\_1,idx\_measurement\_person\_id\_1,idx\_measurement\_visit\_id\_1,xpk\_measurement |
| metadata | idx\_metadata\_concept\_id\_1,xpk\_metadata |
| note | idx\_note\_concept\_id\_1,idx\_note\_person\_id\_1,idx\_note\_visit\_id\_1,xpk\_note |
| note\_nlp | idx\_note\_nlp\_concept\_id\_1,idx\_note\_nlp\_note\_id\_1,xpk\_note\_nlp |
| observation | idx\_observation\_concept\_id\_1,idx\_observation\_person\_id\_1,idx\_observation\_visit\_id\_1,xpk\_observation |
| observation\_period | idx\_observation\_period\_id\_1,xpk\_observation\_period |
| payer\_plan\_period | idx\_period\_person\_id\_1,xpk\_payer\_plan\_period |
| person | idx\_gender,idx\_person\_id,xpk\_person |
| procedure\_occurrence | idx\_procedure\_concept\_id\_1,idx\_procedure\_person\_id\_1,idx\_procedure\_visit\_id\_1,xpk\_procedure\_occurrence |
| provider | idx\_provider\_id\_1,xpk\_provider |
| specimen | idx\_specimen\_concept\_id\_1,idx\_specimen\_person\_id\_1,xpk\_specimen |
| visit\_detail | idx\_visit\_det\_concept\_id\_1,idx\_visit\_det\_occ\_id,idx\_visit\_det\_person\_id\_1,xpk\_visit\_detail |
| visit\_occurrence | idx\_visit\_concept\_id\_1,idx\_visit\_person\_id\_1,xpk\_visit\_occurrence |

Query executed in 0.05 seconds

The following expected indexes are missing:

xpk\_concept, xpk\_vocabulary, xpk\_domain, xpk\_concept\_class, xpk\_relationship, idx\_concept\_concept\_id, idx\_concept\_code, idx\_concept\_vocabluary\_id, idx\_concept\_domain\_id, idx\_concept\_class\_id, idx\_vocabulary\_vocabulary\_id, idx\_domain\_domain\_id, idx\_concept\_class\_class\_id, idx\_concept\_relationship\_id\_1, idx\_concept\_relationship\_id\_2, idx\_concept\_relationship\_id\_3, idx\_relationship\_rel\_id, idx\_concept\_synonym\_id, idx\_concept\_ancestor\_id\_1, idx\_concept\_ancestor\_id\_2, idx\_source\_to\_concept\_map\_3, idx\_source\_to\_concept\_map\_1, idx\_source\_to\_concept\_map\_2, idx\_source\_to\_concept\_map\_c, idx\_drug\_strength\_id\_1, idx\_drug\_strength\_id\_2, idx\_episode\_person\_id\_1, idx\_episode\_concept\_id\_1, idx\_episode\_event\_id\_1, idx\_ee\_field\_concept\_id\_1

* 1. Achilles Query Performance

WARNING: Achilles v1.7 was used. The run time is NOT standardised to one unit. Here, we assume they are all in seconds. This might not be accurate.

1. Execution time of Achilles analyses. Total: 2m 35s. Median: 60ms. Longest duration: 22.1s (analysis 117). Ⓐ

| ID | NAME | DURATION |
| --- | --- | --- |
| 0 | Source name | 0.04 |
| 1 | Number of persons | 0.01 |
| 2 | Number of persons by gender | 0.01 |
| 3 | Number of persons by year of birth | 0.02 |
| 4 | Number of persons by race | 0.01 |
| 5 | Number of persons by ethnicity | 0.01 |
| 7 | Number of persons with invalid provider\_id | 0.01 |
| 8 | Number of persons with invalid location\_id | 0.01 |
| 9 | Number of persons with invalid care\_site\_id | 0.01 |
| 10 | Number of all persons by year of birth by gender | 0.02 |
| 11 | Number of non-deceased persons by year of birth by gender | 0.03 |
| 12 | Number of persons by race and ethnicity | 0.01 |
| 101 | Number of persons by age, with age at first observation period | 0.03 |
| 102 | Number of persons by gender by age, with age at first observation period | 0.03 |
| 108 | Number of persons by length of observation period, in 30d increments | 0.05 |
| 109 | Number of persons with continuous observation in each year | 1.61 |
| 110 | Number of persons with continuous observation in each month | 1.58 |
| 111 | Number of persons by observation period start month | 0.02 |
| 112 | Number of persons by observation period end month | 0.02 |
| 113 | Number of persons by number of observation periods | 0.01 |
| 114 | Number of persons with observation period before year-of-birth | 0.02 |
| 115 | Number of persons with observation period end < observation period start | 0.01 |
| 116 | Number of persons with at least one day of observation in each year by gender and age decile | 1.1 |
| 117 | Number of persons with at least one day of observation in each month | 22.06 |
| 118 | Number of observation periods with invalid person\_id | 0.02 |
| 119 | Number of observation period records by period\_type\_concept\_id | 0.01 |
| 200 | Number of persons with at least one visit occurrence, by visit\_concept\_id | 0.72 |
| 201 | Number of visit occurrence records, by visit\_concept\_id | 0.1 |
| 202 | Number of persons by visit occurrence start month, by visit\_concept\_id | 0.69 |
| 204 | Number of persons with at least one visit occurrence, by visit\_concept\_id by calendar year by gender by age decile | 1.01 |
| 207 | Number of visit records with invalid person\_id | 0.06 |
| 209 | Number of visit records with invalid care\_site\_id | 0.05 |
| 210 | Number of visit\_occurrence records outside a valid observation period | 0.07 |
| 211 | Number of visit records with end date < start date | 0.04 |
| 212 | Number of persons with at least one visit occurrence, by calendar year by gender by age decile | 0.86 |
| 220 | Number of visit occurrence records by visit occurrence start month | 0.21 |
| 221 | Number of persons by visit start year | 0.49 |
| 225 | Number of visit\_occurrence records by visit\_source\_concept\_id | 0.11 |
| 226 | Number of records by domain by visit\_concept\_id | 3.43 |
| 230 | Number of visit\_occurrence records inside valid observation period | 0.08 |
| 231 | Proportion of people with at least one visit\_occurrence record outside a valid observation period | 0.3 |
| 232 | Proportion of visit\_occurrence records outside a valid observation period | 0.1 |
| 300 | Number of providers | 0.01 |
| 301 | Number of providers by specialty concept\_id | 0.01 |
| 303 | Number of providers records by specialty\_concept\_id and visit\_concept\_id | 0.09 |
| 325 | Number of provider records by specialty\_source\_concept\_id | 0.01 |
| 400 | Number of persons with at least one condition occurrence, by condition\_concept\_id | 0.2 |
| 401 | Number of condition occurrence records, by condition\_concept\_id | 0.04 |
| 402 | Number of persons by condition occurrence start month, by condition\_concept\_id | 0.39 |
| 404 | Number of persons with at least one condition occurrence, by condition\_concept\_id by calendar year by gender by age decile | 0.47 |
| 405 | Number of condition occurrence records, by condition\_concept\_id by condition\_type\_concept\_id | 0.06 |
| 409 | Number of condition occurrence records with invalid person\_id | 0.03 |
| 410 | Number of condition occurrence records outside valid observation period | 0.03 |
| 411 | Number of condition occurrence records with end date < start date | 0.02 |
| 412 | Number of condition occurrence records with invalid provider\_id | 0.03 |
| 413 | Number of condition occurrence records with invalid visit\_id | 0.14 |
| 414 | Number of condition occurrence records, by condition\_status\_concept\_id | 0.04 |
| 415 | Number of condition occurrence records, by condition\_type\_concept\_id | 0.04 |
| 416 | Number of condition occurrence records, by condition\_status\_concept\_id, condition\_type\_concept\_id | 0.05 |
| 420 | Number of condition occurrence records by condition occurrence start month | 0.09 |
| 425 | Number of condition\_occurrence records by condition\_source\_concept\_id | 0.05 |
| 430 | Number of condition\_occurrence records inside a valid observation period | 0.03 |
| 431 | Proportion of people with at least one condition\_occurrence record outside a valid observation period | 0.1 |
| 432 | Proportion of condition\_occurrence records outside a valid observation period | 0.04 |
| 500 | Number of persons with death, by cause\_concept\_id | 0.01 |
| 501 | Number of records of death, by cause\_concept\_id | 0.01 |
| 502 | Number of persons by death month | 0.01 |
| 504 | Number of persons with a death, by calendar year by gender by age decile | 0.01 |
| 505 | Number of death records, by death\_type\_concept\_id | 0.01 |
| 509 | Number of death records with invalid person\_id | 0.01 |
| 510 | Number of death records outside valid observation period | 0.01 |
| 525 | Number of death records by cause\_source\_concept\_id | 0.01 |
| 530 | Number of death records inside a valid observation period | 0.01 |
| 531 | Proportion of people with at least one death record outside a valid observation period | 0.01 |
| 532 | Proportion of death records that occur outside a valid observation period | 0.01 |
| 600 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id | 0.61 |
| 601 | Number of procedure occurrence records, by procedure\_concept\_id | 0.1 |
| 602 | Number of persons by procedure occurrence start month, by procedure\_concept\_id | 0.82 |
| 604 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id by calendar year by gender by age decile | 1.09 |
| 605 | Number of procedure occurrence records, by procedure\_concept\_id by procedure\_type\_concept\_id | 0.12 |
| 609 | Number of procedure occurrence records with invalid person\_id | 0.06 |
| 610 | Number of procedure occurrence records outside valid observation period | 0.07 |
| 612 | Number of procedure occurrence records with invalid provider\_id | 0.06 |
| 613 | Number of procedure occurrence records with invalid visit\_id | 0.21 |
| 620 | Number of procedure occurrence records by procedure occurrence start month | 0.21 |
| 625 | Number of procedure\_occurrence records by procedure\_source\_concept\_id | 0.11 |
| 630 | Number of procedure\_occurrence records inside a valid observation period | 0.09 |
| 631 | Proportion of people with at least one procedure\_occurrence record outside a valid observation period | 0.29 |
| 632 | Proportion of procedure\_occurrence records outside a valid observation period | 0.09 |
| 691 | Percentage of total persons that have at least x procedures | 0.24 |
| 700 | Number of persons with at least one drug exposure, by drug\_concept\_id | 0.76 |
| 701 | Number of drug exposure records, by drug\_concept\_id | 0.12 |
| 702 | Number of persons by drug exposure start month, by drug\_concept\_id | 0.95 |
| 704 | Number of persons with at least one drug exposure, by drug\_concept\_id by calendar year by gender by age decile | 1.27 |
| 705 | Number of drug exposure records, by drug\_concept\_id by drug\_type\_concept\_id | 0.14 |
| 709 | Number of drug exposure records with invalid person\_id | 0.06 |
| 710 | Number of drug exposure records outside valid observation period | 0.08 |
| 711 | Number of drug exposure records with end date < start date | 0.06 |
| 712 | Number of drug exposure records with invalid provider\_id | 0.08 |
| 713 | Number of drug exposure records with invalid visit\_id | 0.24 |
| 720 | Number of drug exposure records by drug exposure start month | 0.24 |
| 725 | Number of drug\_exposure records by drug\_source\_concept\_id | 0.15 |
| 730 | Number of drug\_exposure records inside a valid observation period | 0.09 |
| 731 | Proportion of people with at least one drug\_exposure record outside a valid observation period | 0.33 |
| 732 | Proportion of drug\_exposure records outside a valid observation period | 0.11 |
| 791 | Percentage of total persons that have at least x drug exposures | 0.24 |
| 800 | Number of persons with at least one observation occurrence, by observation\_concept\_id | 0.17 |
| 801 | Number of observation occurrence records, by observation\_concept\_id | 0.04 |
| 802 | Number of persons by observation occurrence start month, by observation\_concept\_id | 0.38 |
| 804 | Number of persons with at least one observation occurrence, by observation\_concept\_id by calendar year by gender by age decile | 0.44 |
| 805 | Number of observation occurrence records, by observation\_concept\_id by observation\_type\_concept\_id | 0.05 |
| 807 | Number of observation occurrence records, by observation\_concept\_id and unit\_concept\_id | 0.06 |
| 809 | Number of observation records with invalid person\_id | 0.03 |
| 810 | Number of observation records outside valid observation period | 0.03 |
| 812 | Number of observation records with invalid provider\_id | 0.03 |
| 813 | Number of observation records with invalid visit\_id | 0.16 |
| 814 | Number of observation records with no value (numeric, string, or concept) | 0.02 |
| 820 | Number of observation records by observation start month | 0.09 |
| 822 | Number of observation records, by observation\_concept\_id and value\_as\_concept\_id | 0.05 |
| 823 | Number of observation records, by observation\_concept\_id and qualifier\_concept\_id | 0.05 |
| 825 | Number of observation records by observation\_source\_concept\_id | 0.06 |
| 826 | Number of observation records by value\_as\_concept\_id | 0.05 |
| 827 | Number of observation records by unit\_concept\_id | 0.05 |
| 830 | Number of observation records inside a valid observation period | 0.04 |
| 831 | Proportion of people with at least one observation record outside a valid observation period | 0.11 |
| 832 | Proportion of observation records outside a valid observation period | 0.04 |
| 891 | Percentage of total persons that have at least x observations | 0.1 |
| 900 | Number of persons with at least one drug era, by drug\_concept\_id | 0.08 |
| 901 | Number of drug era records, by drug\_concept\_id | 0.04 |
| 902 | Number of persons by drug era start month, by drug\_concept\_id | 0.34 |
| 904 | Number of persons with at least one drug era, by drug\_concept\_id by calendar year by gender by age decile | 0.37 |
| 908 | Number of drug eras without valid person | 0.02 |
| 910 | Number of drug\_era records outside valid observation period | 0.02 |
| 911 | Number of drug eras with end date < start date | 0.01 |
| 920 | Number of drug era records by drug era start month | 0.1 |
| 930 | Number of drug\_era records inside a valid observation period | 0.02 |
| 931 | Proportion of people with at least one drug\_era record outside a valid observation period | 0.05 |
| 932 | Proportion of drug\_era records outside a valid observation period | 0.03 |
| 1,000 | Number of persons with at least one condition era, by condition\_concept\_id | 0.15 |
| 1,001 | Number of condition era records, by condition\_concept\_id | 0.05 |
| 1,002 | Number of persons by condition era start month, by condition\_concept\_id | 0.42 |
| 1,004 | Number of persons with at least one condition era, by condition\_concept\_id by calendar year by gender by age decile | 0.59 |
| 1,008 | Number of condition eras without valid person | 0.02 |
| 1,010 | Number of condition\_era records outside a valid observation period | 0.03 |
| 1,011 | Number of condition eras with end date < start date | 0.02 |
| 1,020 | Number of condition era records by condition era start month | 0.11 |
| 1,030 | Number of condition\_era records inside a valid observation period | 0.04 |
| 1,031 | Proportion of people with at least one condition\_era record outside a valid observation period | 0.09 |
| 1,032 | Proportion of condition\_era records outside a valid observation period | 0.05 |
| 1,100 | Number of persons by location 3-digit zip | 0.01 |
| 1,101 | Number of persons by location state | 0.01 |
| 1,102 | Number of care sites by location 3-digit zip | 0.01 |
| 1,103 | Number of care sites by location state | 0.01 |
| 1,200 | Number of persons by place of service | 0.01 |
| 1,201 | Number of visits by place of service | 0.04 |
| 1,202 | Number of care sites by place of service | 0.01 |
| 1,203 | Number of visits by place of service discharge type | 0.05 |
| 1,300 | Number of persons with at least one visit detail, by visit\_detail\_concept\_id | 0.76 |
| 1,301 | Number of visit detail records, by visit\_detail\_concept\_id | 0.1 |
| 1,302 | Number of persons by visit detail start month, by visit\_detail\_concept\_id | 0.69 |
| 1,304 | Number of persons with at least one visit detail, by visit\_detail\_concept\_id by calendar year by gender by age decile | 1.01 |
| 1,307 | Number of visit records with invalid person\_id | 0.06 |
| 1,309 | Number of visit\_detail records with invalid care\_site\_id | 0.04 |
| 1,310 | Number of visit\_detail records outside a valid observation period | 0.07 |
| 1,311 | Number of visit\_detail records with end date < start date | 0.05 |
| 1,312 | Number of persons with at least one visit detail, by calendar year by gender by age decile | 0.89 |
| 1,320 | Number of visit detail records by visit detail start month | 0.21 |
| 1,321 | Number of persons by visit start year | 0.5 |
| 1,325 | Number of visit\_detail records by visit\_detail\_source\_concept\_id | 0.11 |
| 1,326 | Number of records by domain by visit\_detail\_concept\_id | 2.78 |
| 1,330 | Number of visit\_detail records inside a valid observation period | 0.09 |
| 1,331 | Proportion of people with at least one visit\_detail record outside a valid observation period | 0.31 |
| 1,332 | Proportion of visit\_detail records outside a valid observation period | 0.1 |
| 1,408 | Number of persons by length of payer plan period, in 30d increments | 0.01 |
| 1,409 | Number of persons with continuous payer plan in each year | 0.01 |
| 1,410 | Number of persons with continuous payer plan in each month | 0.02 |
| 1,411 | Number of persons by payer plan period start month | 0.01 |
| 1,412 | Number of persons by payer plan period end month | 0.01 |
| 1,413 | Number of persons by number of payer plan periods | 0.01 |
| 1,414 | Number of persons with payer plan period before year-of-birth | 0.01 |
| 1,415 | Number of persons with payer plan period end < payer plan period start | 0.01 |
| 1,425 | Number of payer\_plan\_period records by payer\_source\_concept\_id | 0.01 |
| 1,800 | Number of persons with at least one measurement occurrence, by measurement\_concept\_id | 11.41 |
| 1,801 | Number of measurement occurrence records, by measurement\_concept\_id | 6.21 |
| 1,802 | Number of persons by measurement occurrence start month, by measurement\_concept\_id | 12.99 |
| 1,804 | Number of persons with at least one mesurement occurrence, by measurement\_concept\_id by calendar year by gender by age decile | 16.92 |
| 1,805 | Number of measurement occurrence records, by measurement\_concept\_id by measurement\_type\_concept\_id | 3.78 |
| 1,807 | Number of measurement occurrence records, by measurement\_concept\_id and unit\_concept\_id | 3.45 |
| 1,809 | Number of measurement records with invalid person\_id | 0.6 |
| 1,810 | Number of measurement records outside valid observation period | 2.7 |
| 1,811 | Number of measurement records with a valid value (with a mapped, non-null value\_as\_number) | 2.61 |
| 1,812 | Number of measurement records with invalid provider\_id | 2.32 |
| 1,813 | Number of measurement records with invalid visit\_id | 3.47 |
| 1,814 | Number of measurement records with no value (numeric, string, or concept) | 2.02 |
| 1,818 | Number of measurement records below/within/above normal range, by measurement\_concept\_id and unit\_concept\_id | 0.48 |
| 1,819 | Number of measurement records, by measurement\_concept\_id, with a valid value (with a mapped, non-null value\_as\_number) | 0.85 |
| 1,820 | Number of measurement records by measurement start month | 2.12 |
| 1,821 | Number of measurement records with no numeric value | 0.47 |
| 1,822 | Number of measurement records, by measurement\_concept\_id and value\_as\_concept\_id | 1.21 |
| 1,823 | Number of measurement records, by measurement\_concept\_id and operator\_concept\_id | 1.13 |
| 1,825 | Number of measurement records by measurement\_source\_concept\_id | 1.18 |
| 1,826 | Number of measurement records by value\_as\_concept\_id | 1.07 |
| 1,827 | Number of measurement records by unit\_concept\_id | 1.1 |
| 1,830 | Number of visit\_detail records inside a valid observation period | 0.76 |
| 1,831 | Proportion of people with at least one measurement record outside a valid observation period | 3.86 |
| 1,832 | Proportion of measurement records outside a valid observation period | 0.82 |
| 1,833 | Proportion of measurement records inside a valid observation period and without a value | 1.17 |
| 1,891 | Percentage of total persons that have at least x measurements | 2.12 |
| 2,100 | Number of persons with at least one device exposure, by device\_concept\_id | 0.07 |
| 2,101 | Number of device exposure records, by device\_concept\_id | 0.01 |
| 1,900 | Source values mapped to concept\_id 0 by table, by column, by source\_value | 2.86 |
| 2,000 | Number of patients with at least 1 Dx and 1 Rx | 0.54 |
| 2,001 | Number of patients with at least 1 Dx and 1 Proc | 0.51 |
| 2,002 | Number of patients with at least 1 Meas, 1 Dx and 1 Rx | 2.79 |
| 2,003 | Number of patients with at least 1 Visit | 0.37 |
| 2,004 | Number of distinct patients that overlap between specific domains | 1.98 |
| 2,100 | Number of persons with at least one device exposure, by device\_concept\_id | 0.01 |
| 2,101 | Number of device exposure records, by device\_concept\_id | 0.01 |
| 2,102 | Number of persons by device records start month, by device\_concept\_id | 0.01 |
| 2,104 | Number of persons with at least one device exposure, by device\_concept\_id by calendar year by gender by age decile | 0.01 |
| 2,105 | Number of device exposure records, by device\_concept\_id by device\_type\_concept\_id | 0.01 |
| 2,110 | Number of device\_exposure records outside valid observation period | 0.01 |
| 2,125 | Number of device\_exposure records by device\_source\_concept\_id | 0.01 |
| 2,130 | Number of device\_exposure records inside a valid observation period | 0.01 |
| 2,131 | Proportion of people with at least one device\_exposure record outside a valid observation period | 0.01 |
| 2,132 | Proportion of device\_exposure records outside a valid observation period | 0.01 |
| 2,191 | Percentage of total persons that have at least x device exposures | 0.01 |
| 2,200 | Number of persons with at least one note by note\_type\_concept\_id | 0.01 |
| 2,201 | Number of note records, by note\_type\_concept\_id | 0.01 |

Query executed in 0.06 seconds

1. Appendix
   1. Vocabulary table counts
2. The number of records in all vocabulary tables. Ⓞ

| TABLENAME | COUNT |
| --- | --- |
| concept\_relationship | 19,269,350 |
| concept\_ancestor | 10,909,302 |
| concept | 1,575,149 |
| concept\_synonym | 1,419,014 |
| drug\_strength | 200,666 |
| relationship | 672 |
| concept\_class | 416 |
| domain | 50 |
| vocabulary | 44 |

Query executed in 1.36 seconds

* 1. Vocabulary concept counts

Vocabulary version: v5.0 31-OCT-22

1. 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 | - |
| --- | --- | --- | --- | --- | --- |
| ATC | WHO Anatomic Therapeutic Chemical Classification | RxNorm 20210907 | 0 | 6,509 | 231 |
| CDM | OMOP Common DataModel | CDM v6.0.0 | 1,045 | 0 | 0 |
| Cohort Type | OMOP Cohort Type | NA | 0 | 0 | 1 |
| Concept Class | OMOP Concept Class | NA | 0 | 0 | 416 |
| Condition Status | OMOP Condition Status | NA | 22 | 0 | 0 |
| Condition Type | OMOP Condition Occurrence Type | NA | 0 | 0 | 118 |
| Cost | OMOP Cost | NA | 51 | 0 | 0 |
| Cost Type | OMOP Cost Type | NA | 0 | 0 | 8 |
| Death Type | OMOP Death Type | NA | 0 | 0 | 14 |
| Device Type | OMOP Device Type | NA | 0 | 0 | 4 |
| Domain | OMOP Domain | NA | 0 | 0 | 65 |
| Drug Type | OMOP Drug Exposure Type | NA | 0 | 0 | 16 |
| Episode | OMOP Episode | Episode 20201014 | 14 | 0 | 4 |
| Ethnicity | OMOP Ethnicity | NA | 2 | 0 | 0 |
| Gender | OMOP Gender | NA | 2 | 0 | 3 |
| Korean Revenue Code | Korean Revenue Code | NA | 7 | 0 | 0 |
| Language | Language (OMOP) | Language 20221030 | 1 | 0 | 0 |
| Meas Type | OMOP Measurement Type | NA | 0 | 0 | 12 |
| Metadata | Metadata | NA | 1 | 0 | 1 |
| None | OMOP Standardized Vocabularies | v5.0 31-OCT-22 | 0 | 0 | 1 |
| Note Type | OMOP Note Type | NA | 0 | 0 | 10 |
| OMOP Extension | OMOP Extension (OHDSI) | OMOP Extension 20221026 | 1,158 | 0 | 53 |
| OSM | OpenStreetMap | OSM Release 2019-02-21 | 203,339 | 0 | 0 |
| Obs Period Type | OMOP Observation Period Type | NA | 0 | 0 | 6 |
| Observation Type | OMOP Observation Type | NA | 0 | 0 | 29 |
| Plan | Health Plan - contract to administer healthcare transactions by the payer, facilitated by the sponsor | NA | 11 | 0 | 0 |
| Plan Stop Reason | Plan Stop Reason - Reason for termination of the Health Plan | NA | 13 | 0 | 0 |
| Procedure Type | OMOP Procedure Occurrence Type | NA | 0 | 0 | 97 |
| Race | Race and Ethnicity Code Set (USBC) | Version 1.0 | 50 | 0 | 3 |
| Relationship | OMOP Relationship | NA | 14 | 0 | 680 |
| RxNorm | RxNorm (NLM) | RxNorm 20221003 | 149,688 | 35,316 | 119,254 |
| SNOMED | Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO) | 2021-07-31 SNOMED CT International Edition; 2021-09-01 SNOMED CT US Edition; 2021-11-24 SNOMED CT UK Edition | 538,283 | 0 | 516,652 |
| SOPT | Source of Payment Typology (PHDSC) | SOPT Version 9.2 | 162 | 0 | 6 |
| Sponsor | Sponsor - institution or individual financing healthcare transactions | NA | 6 | 0 | 0 |
| Type Concept | OMOP Type Concept | Type Concept 20221030 | 80 | 0 | 0 |
| UB04 Point of Origin | UB04 Claim Source Inpatient Admission Code (CMS) | NA | 0 | 0 | 23 |
| UB04 Pri Typ of Adm | UB04 Claim Inpatient Admission Type Code (CMS) | NA | 6 | 0 | 0 |
| UB04 Pt dis status | UB04 Patient Discharge Status Code (CMS) | NA | 0 | 0 | 55 |
| UB04 Typ bill | UB04 Type of Bill - Institutional (USHIK) | NA | 4 | 0 | 294 |
| UCUM | Unified Code for Units of Measure (Regenstrief Institute) | Version 1.8.2 | 1,029 | 0 | 89 |
| US Census | United States Census Bureau | US Census 2017 Release | 13 | 0 | 0 |
| Visit | OMOP Visit | Visit 20211216 | 19 | 0 | 0 |
| Visit Type | OMOP Visit Type | NA | 0 | 0 | 18 |
| Vocabulary | OMOP Vocabulary | NA | 0 | 0 | 142 |

Query executed in 0.43 seconds