

CDM Inspection report for the fdzformat database

Package Version: 1.2.3

Date: Wed Jan 17 16:40:18 2024

Authors: mfinster

# Table of content

# 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 | fdzformat |
| Database acronym | OMOP\_format1 |
| Contact Person |  |
| Email |  |
| SME |  |
| Contact Person |  |
| Email SME |  |

## Database Description

ETL from Format 1 FDZ to OMOP Reported year 2010

## SME Role

Describe the involvement of the SME in the ETL Development process

# ETL Development General

This section describes 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 | N\_PERSONS |
| --- | --- | --- |
| cost | 48,048,000 | NA |
| observation | 21,334,696 | 6,863,000 |
| visit\_occurrence | 13,728,000 | 6,863,000 |
| condition\_occurrence | 12,420,958 | 6,818,401 |
| payer\_plan\_period | 6,864,000 | 6,863,000 |
| person | 6,863,000 | 6,863,000 |
| observation\_period | 6,863,000 | 6,863,000 |
| procedure\_occurrence | 3,456,112 | 3,437,748 |
| drug\_exposure | 3,433,000 | 3,432,000 |
| death | 3,432,000 | 3,431,500 |
| location | 17,280 | NA |
| measurement | 9,680 | 5,368 |
| provider | 0 | NA |
| specimen | 0 | 0 |
| dose\_era | 0 | 0 |
| device\_exposure | 0 | 0 |
| visit\_detail | 0 | 0 |
| drug\_era | 0 | 0 |
| condition\_era | 0 | 0 |
| care\_site | 0 | NA |
| note | 0 | 0 |

Query executed in 72.82 secs

## Data density plots

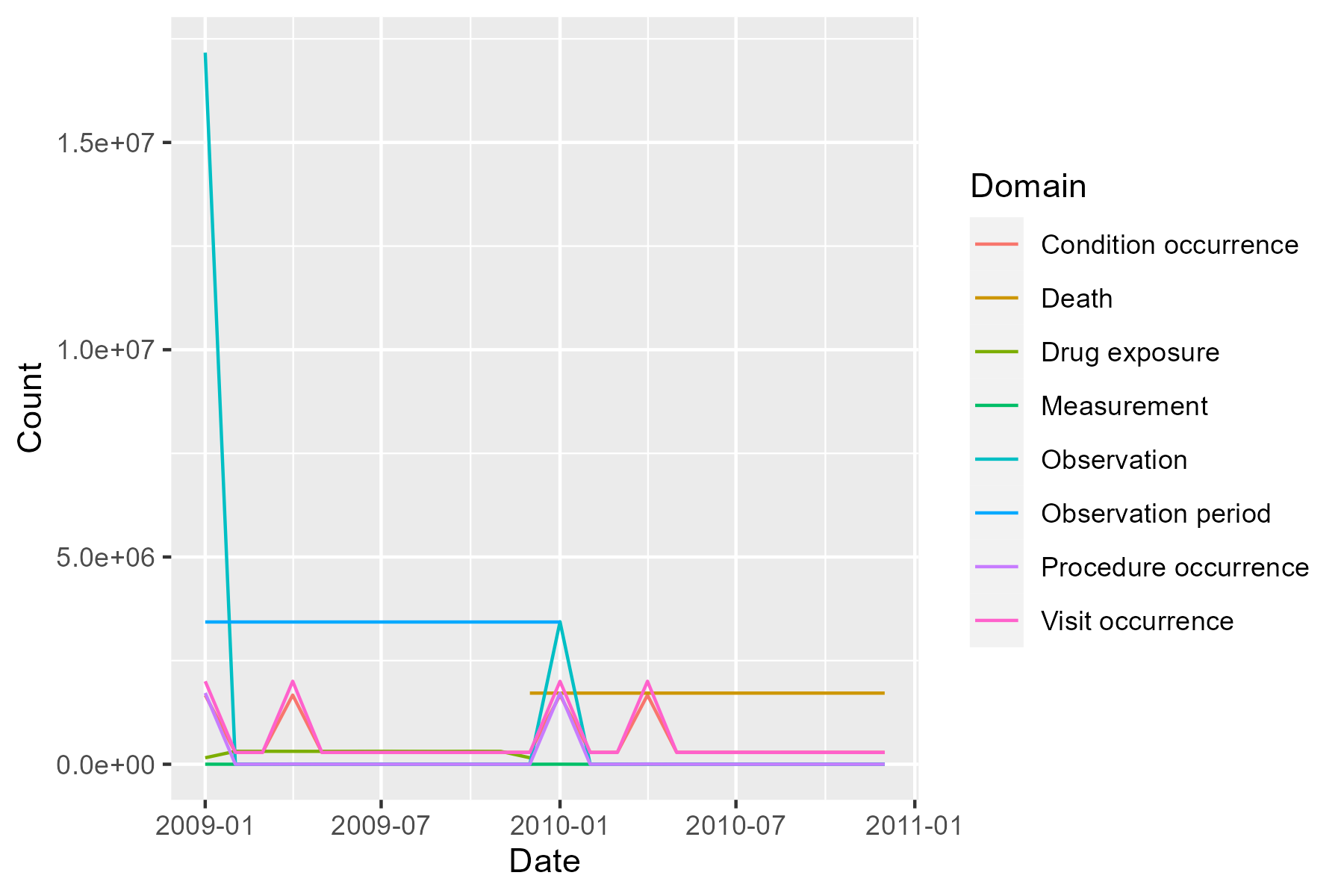


Figure 1. Total record count over time per data domain

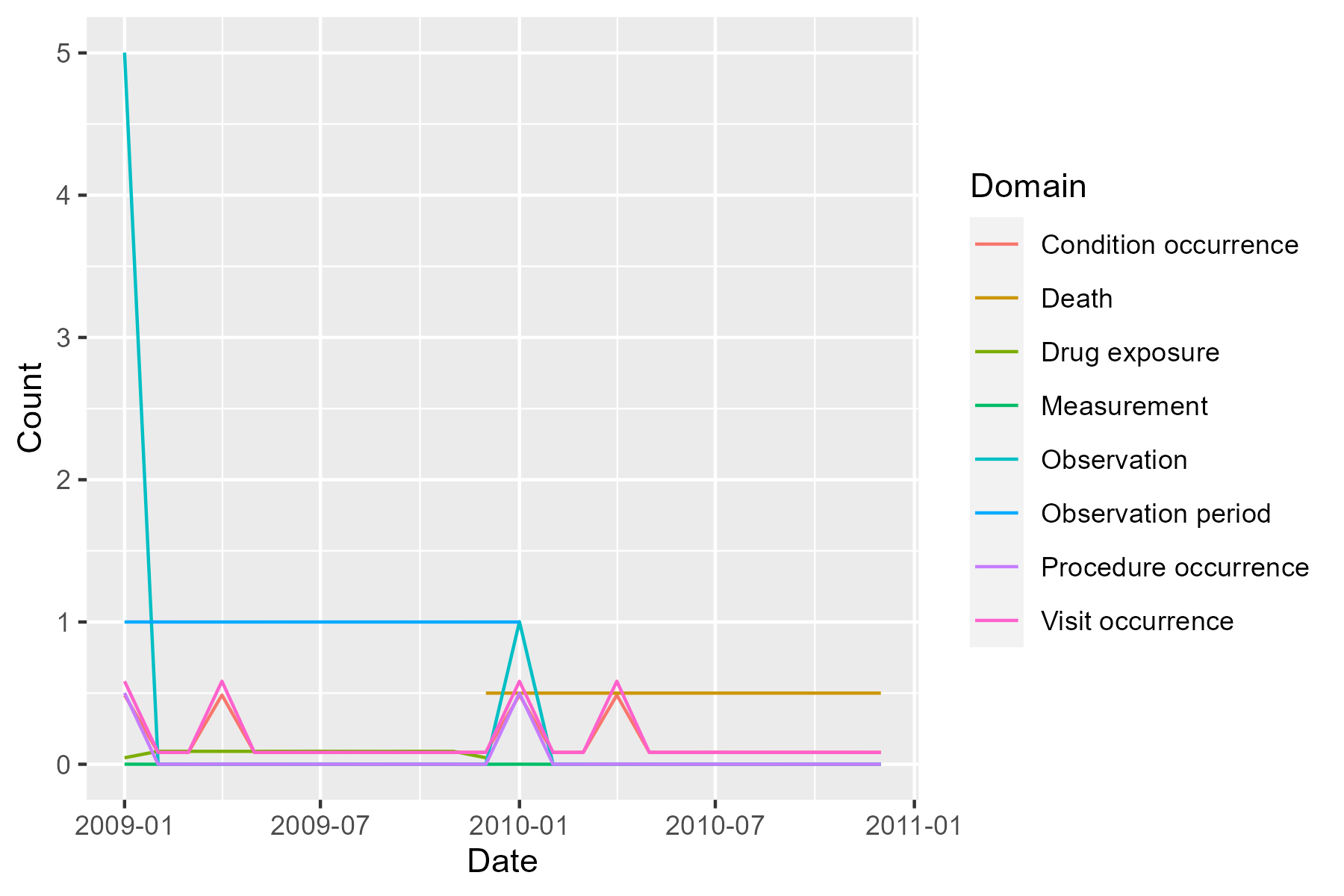


Figure 2. Number of records per person over time per data domain

## Distinct concepts per person

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

| Domain | Min | P10 | P25 | MEDIAN | P75 | P90 | Max |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Condition occurrence | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| Measurement | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| Drug exposure | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Procedure occurrence | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| Observation | 1 | 1 | 1 | 1 | 1 | 1 | 3 |

# 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. Ideally these lists are sorted descending by source code frequency.

## Vocabularies

Vocabulary version: v5.0 23-JAN-23

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) | 2018-06-26 ABMS | 85 | 0 | 13 |
| ATC | WHO Anatomic Therapeutic Chemical Classification | RxNorm 20210907 | 0 | 6,509 | 231 |
| CDM | OMOP Common DataModel | CDM v6.0.0 | 1,045 | 0 | 0 |
| CMS Place of Service | Place of Service Codes for Professional Claims (CMS) | 2009-01-11 | 51 | 0 | 9 |
| 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 |
| CPT4 | Current Procedural Terminology version 4 (AMA) | 2022 Release | 11,909 | 3,559 | 1,454 |
| Currency | International Currency Symbol (ISO 4217) | 2008 | 180 | 0 | 0 |
| 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 |
| EBM | German Uniform Assessment | 2000001037 | 0 | 0 | 3,614 |
| Episode | OMOP Episode | Episode 20201014 | 14 | 0 | 4 |
| Ethnicity | OMOP Ethnicity | NA | 2 | 0 | 0 |
| Gender | OMOP Gender | NA | 2 | 0 | 3 |
| HCPCS | Healthcare Common Procedure Coding System (CMS) | 20221001 Alpha Numeric HCPCS File | 8,461 | 0 | 2,808 |
| ICD10CM | International Classification of Diseases, Tenth Revision, Clinical Modification (NCHS) | ICD10CM FY2023 code descriptions | 0 | 0 | 98,583 |
| ICD10GM | International Classification of Diseases, Tenth Revision, German Edition | ICD10GM 2022 | 0 | 0 | 17,213 |
| ICD9CM | International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 1 and 2 (NCHS) | ICD9CM v32 master descriptions | 0 | 0 | 17,564 |
| ICD9Proc | International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 3 (NCHS) | ICD9CM v32 master descriptions | 2,223 | 0 | 2,434 |
| Insured days | Forschungsdatenzentrum (DE) Format1 Insured days | 2023-03-21 | 0 | 0 | 6 |
| Korean Revenue Code | Korean Revenue Code (KNHIS) | NA | 7 | 0 | 0 |
| Language | OMOP Language | Language 20221030 | 1 | 0 | 0 |
| LOINC | Logical Observation Identifiers Names and Codes (Regenstrief Institute) | LOINC 2.73 | 113,893 | 49,168 | 102,015 |
| Meas Type | OMOP Measurement Type | NA | 0 | 0 | 12 |
| Medicare Specialty | Medicare provider/supplier specialty codes (CMS) | 2018-06-26 Specialty | 112 | 0 | 8 |
| Metadata | OMOP Metadata | NA | 1 | 0 | 1 |
| NDC | National Drug Code (FDA and manufacturers) | NDC 20230122 | 11,403 | 0 | 1,126,729 |
| None | OMOP Standardized Vocabularies | v5.0 23-JAN-23 | 0 | 0 | 1 |
| Note Type | OMOP Note Type | NA | 0 | 0 | 10 |
| NUCC | National Uniform Claim Committee Health Care Provider Taxonomy Code Set (NUCC) | 2018-06-26 NUCC | 674 | 0 | 181 |
| Observation Type | OMOP Observation Type | NA | 0 | 0 | 29 |
| Obs Period Type | OMOP Observation Period Type | NA | 0 | 0 | 6 |
| OMOP Extension | OMOP Extension (OHDSI) | OMOP Extension 20230110 | 1,187 | 0 | 53 |
| OPS | Operations and Procedures Classification (OPS) | OPS Version 2022 | 0 | 0 | 42,959 |
| OSM | OpenStreetMap (OSMF) | OSM Release 2019-02-21 | 203,339 | 0 | 0 |
| Plan | OMOP Health Plan | NA | 11 | 0 | 0 |
| Plan Stop Reason | OMOP Plan Stop Reason | NA | 13 | 0 | 0 |
| Procedure Type | OMOP Procedure Occurrence Type | NA | 0 | 0 | 97 |
| Professional groups | Forschungsdatenzentrum (DE) Format1 Insured days | 2023-05-03 | 0 | 0 | 1 |
| Race | Race and Ethnicity Code Set (USBC) | Version 1.0 | 50 | 0 | 3 |
| Relationship | OMOP Relationship | NA | 14 | 0 | 698 |
| Revenue Code | UB04/CMS1450 Revenue Codes (CMS) | 2010 Release | 538 | 0 | 0 |
| RxNorm | RxNorm (NLM) | RxNorm 20230103 | 149,993 | 35,341 | 119,532 |
| RxNorm Extension | OMOP RxNorm Extension | RxNorm Extension 2023-01-16 | 1,834,465 | 0 | 275,964 |
| 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 |
| SPL | Structured Product Labeling (FDA) | NDC 20230122 | 0 | 626,507 | 14,948 |
| Sponsor | OMOP Sponsor | 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 | Census regions of the United States (USCB) | 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 | 143 |

Query executed in 7.70 secs

## Table counts

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

| TABLENAME | COUNT |
| --- | --- |
| concept\_class | 418 |
| relationship | 690 |
| concept | 5,945,608 |
| concept\_synonym | 2,101,511 |
| vocabulary | 64 |
| domain | 50 |
| concept\_ancestor | 70,537,715 |
| drug\_strength | 2,936,738 |
| concept\_relationship | 46,843,676 |

Query executed in 25.93 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 | 154,986 | 23,084 | 14.9% | 12,420,958 | 1,974,500 | 15.9% |
| Condition status | 6 | 5 | 83.3% | 12,420,958 | 11,040,194 | 88.9% |
| Death cause | 0 | NA | NA | NA | NA | NA |
| Device | 0 | NA | NA | NA | NA | NA |
| Drug | 1,717,000 | 0 | 0% | 3,433,000 | 0 | 0% |
| Measurement | 122 | 122 | 100% | 9,680 | 9,680 | 100% |
| Measurement unit | 0 | NA | NA | NA | NA | NA |
| Measurement value | 0 | NA | NA | NA | NA | NA |
| Observation | 31,888 | 31,887 | 100.0% | 21,334,696 | 738,696 | 3.5% |
| Observation unit | 0 | NA | NA | NA | NA | NA |
| Observation value | 0 | NA | NA | NA | NA | NA |
| Procedure | 273 | 273 | 100% | 3,456,112 | 3,456,112 | 100% |
| Provider Specialty | 0 | NA | NA | NA | NA | NA |
| Specimen | 0 | NA | NA | NA | NA | NA |
| Visit | 3 | 2 | 66.7% | 13,728,000 | 10,296,000 | 75.0% |

Query executed in 68.25 secs

## Drug Mappings

Table 6. The level of the drug mappings

| Class | #Records | #Patients | #Codes |
| --- | --- | --- | --- |
|  |  |  |  |

Query executed in 11.96 secs

## Unmapped Codes

Table 7 omitted because no unmapped drugs were found.

Query executed in 26.28 secs

Table 8. Top 25 of unmapped conditions. Counts are rounded up to the nearest hundred.

| # | Source Value | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | G02.5 | 100 | 100 |
| 2 | D52.09\* | 100 | 100 |
| 3 | D521- | 100 | 100 |
| 4 | D52.3 | 100 | 100 |
| 5 | D52.3\* | 100 | 100 |
| 6 | D52.5 | 100 | 100 |
| 7 | D52.5\* | 100 | 100 |
| 8 | D52.5+ | 100 | 100 |
| 9 | D52.5- | 100 | 100 |
| 10 | D525! | 100 | 100 |
| 11 | D52.6 | 100 | 100 |
| 12 | D526- | 100 | 100 |
| 13 | D53.01# | 100 | 100 |
| 14 | D53.01\* | 100 | 100 |
| 15 | D53.01+ | 100 | 100 |
| 16 | D53.02 | 100 | 100 |
| 17 | D53.02! | 100 | 100 |
| 18 | D53.02# | 100 | 100 |
| 19 | D53.02\* | 100 | 100 |
| 20 | D53.02+ | 100 | 100 |
| 21 | D53.02- | 100 | 100 |
| 22 | D53.03! | 100 | 100 |
| 23 | D53.04! | 100 | 100 |
| 24 | D53.04\* | 100 | 100 |
| 25 | D53.04+ | 100 | 100 |

Query executed in 73.32 secs

Table 9 omitted because no unmapped measurements were found.

Query executed in 0.68 secs

Table 10. All 1 unmapped observations. Counts are rounded up to the nearest hundred.

| # | Source Value | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | NA | 20,596,000 | 6,863,000 |

Query executed in 35.28 secs

Table 11 omitted because no unmapped procedures were found.

Query executed in 1.32 secs

Table 12 omitted because no unmapped devices were found.

Query executed in 0.82 secs

Table 13. All 1 unmapped visits. Counts are rounded up to the nearest hundred.

| # | Source Value | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | 0 | 3,432,000 | 3,431,500 |

Query executed in 7.41 secs

## Mapped Codes

Table 14 omitted because no mapped drugs were found.

Query executed in 1.28 secs

Table 15. Top 25 of mapped conditions. Counts are rounded up to the nearest hundred.

| # | Concept Name | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | Finding related to pregnancy | 36,000 | 19,900 |
| 2 | Disorder of upper arm | 8,300 | 4,500 |
| 3 | Infection due to resistant bacteria | 5,000 | 2,900 |
| 4 | Poisoning by drug AND/OR medicinal substance | 4,900 | 2,700 |
| 5 | Abnormal finding on evaluation procedure | 4,600 | 2,500 |
| 6 | Fetal condition affecting obstetrical care of mother | 4,200 | 2,300 |
| 7 | Dislocations/sprains/strains | 4,000 | 2,200 |
| 8 | Cerebrovascular disease | 3,900 | 2,100 |
| 9 | Disorder of joint of ankle and/or foot | 3,800 | 2,200 |
| 10 | Disorder of lower leg | 3,700 | 2,100 |
| 11 | Arthropathy associated with another disorder | 3,500 | 2,000 |
| 12 | Disorder of free lower limb | 3,500 | 1,900 |
| 13 | Staphylococcal arthritis | 3,500 | 1,900 |
| 14 | Disorder of sacrum | 3,400 | 1,900 |
| 15 | Muscle and tendon injury | 3,400 | 1,900 |
| 16 | Juvenile osteochondrosis of spine | 3,400 | 1,900 |
| 17 | Disorder of forearm | 3,400 | 1,900 |
| 18 | Gouty arthropathy | 3,300 | 1,900 |
| 19 | Pregnancy with abortive outcome | 3,300 | 1,900 |
| 20 | Purulent myositis | 3,300 | 1,900 |
| 21 | Disorder of lumbar spine | 3,300 | 1,800 |
| 22 | Postmeningococcal arthritis | 3,200 | 1,800 |
| 23 | Gonococcal bursitis | 3,200 | 1,800 |
| 24 | Ankylosing spondylitis | 3,200 | 1,800 |
| 25 | Disorder of cervical spine | 3,200 | 1,700 |

Query executed in 18.00 secs

Table 16. Top 25 of mapped measurements. Counts are rounded up to the nearest hundred.

| # | Concept Name | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | Screening for disorder | 1,200 | 700 |
| 2 | HIV CDC stage finding | 900 | 600 |
| 3 | Antenatal screening | 800 | 500 |
| 4 | Barthel index | 800 | 400 |
| 5 | Measurement of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) | 500 | 300 |
| 6 | Screening for malignant neoplasm of prostate | 400 | 300 |
| 7 | Albumin measurement | 400 | 300 |
| 8 | Hematology screening test | 400 | 300 |
| 9 | Alpha-1-Fetoprotein measurement, amniotic fluid | 400 | 300 |
| 10 | In vivo test of hypersensitivity | 400 | 300 |
| 11 | Fetal chromosomal abnormality screening | 400 | 300 |
| 12 | Screening for malignant neoplasm of breast | 400 | 300 |
| 13 | Antenatal screening for isoimmunization | 400 | 300 |
| 14 | Screening for malignant neoplasm of stomach | 400 | 300 |
| 15 | Screening for cancer | 400 | 200 |
| 16 | Antibody titer measurement | 400 | 300 |
| 17 | Blood pressure taking | 300 | 200 |
| 18 | Serum/plasma protein test | 300 | 200 |
| 19 | Blood test | 300 | 200 |
| 20 | Immunology laboratory test | 300 | 200 |
| 21 | Antenatal screening for malformation | 300 | 200 |
| 22 | Enzyme measurement | 300 | 200 |
| 23 | Screening for cardiovascular system disease | 100 | 100 |
| 24 | Screening for malignant neoplasm of intestinal tract | 100 | 100 |
| 25 | Globulin measurement | 100 | 100 |

Query executed in 1.10 secs

Table 17. Top 25 of mapped observations. Counts are rounded up to the nearest hundred.

| # | Concept Name | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | Disorder excluded | 686,600 | 686,400 |
| 2 | History of event | 8,700 | 5,000 |
| 3 | Follow-up encounter | 2,600 | 1,500 |
| 4 | Family history of clinical finding | 1,700 | 1,000 |
| 5 | Aftercare | 1,400 | 800 |
| 6 | Requires vaccination | 1,400 | 800 |
| 7 | Allergy to substance | 1,400 | 800 |
| 8 | Body mass index 40+ - severely obese | 1,000 | 600 |
| 9 | Patient encounter procedure | 1,000 | 600 |
| 10 | On waiting list for organ transplant | 900 | 500 |
| 11 | Exposure to communicable disease | 800 | 400 |
| 12 | Need for personal care assistance | 800 | 400 |
| 13 | Family history of cancer | 800 | 400 |
| 14 | Requires polio vaccination | 700 | 400 |
| 15 | Donor for medical or surgical procedure | 700 | 400 |
| 16 | Tobacco user | 700 | 400 |
| 17 | Accident due to cold of man-made origin | 600 | 300 |
| 18 | Bone tissue donor | 500 | 300 |
| 19 | Family history of malignant neoplasm of thoracic cavity structure | 500 | 300 |
| 20 | Blood donor | 500 | 400 |
| 21 | Adverse reaction to drug | 500 | 300 |
| 22 | Requires rabies vaccination course | 500 | 300 |
| 23 | Speech therapy | 500 | 300 |
| 24 | Eye donor for corneal transplant | 500 | 300 |
| 25 | Suicide attempt | 500 | 300 |

Query executed in 11.39 secs

Table 18. Top 25 of mapped procedures. Counts are rounded up to the nearest hundred.

| # | Concept Name | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | Dialysis procedure | 3,432,100 | 3,431,600 |
| 2 | Preventive procedure | 1,600 | 900 |
| 3 | Fitting procedure | 1,500 | 800 |
| 4 | Adjustment of implant | 1,500 | 800 |
| 5 | Attention to device, anastomosis AND/OR stoma | 1,300 | 800 |
| 6 | Surgical procedure | 1,200 | 700 |
| 7 | Medical examination for suspected condition | 1,100 | 600 |
| 8 | Analysis of electronic device used in cardiovascular system | 900 | 500 |
| 9 | Reproductive care management | 800 | 500 |
| 10 | Specialized medical examination | 700 | 400 |
| 11 | General examination of patient | 700 | 400 |
| 12 | Contraception care management | 600 | 400 |
| 13 | Examination for suspected mental disorder | 600 | 300 |
| 14 | Fitting of prosthesis | 500 | 300 |
| 15 | Postnatal maternal examination | 500 | 300 |
| 16 | In vitro fertilization | 500 | 300 |
| 17 | Prophylactic immunotherapy | 500 | 300 |
| 18 | Rehabilitation therapy | 500 | 300 |
| 19 | Ophthalmic examination and evaluation | 500 | 300 |
| 20 | Neurological mental status determination | 500 | 300 |
| 21 | Examination for suspected tuberculosis | 500 | 300 |
| 22 | Reversal of sterilization | 500 | 300 |
| 23 | Hearing examination | 500 | 300 |
| 24 | Genetic counseling | 500 | 300 |
| 25 | Attention to gastrostomy | 400 | 300 |

Query executed in 14.60 secs

Table 19 omitted because no mapped devices were found.

Query executed in 0.83 secs

Table 20. All 2 mapped visits. Counts are rounded up to the nearest hundred.

| # | Concept Name | #Records | #Subjects |
| --- | --- | --- | --- |
| 1 | Outpatient Hospital | 6,864,000 | 6,863,000 |
| 2 | Outpatient Visit | 3,432,000 | 3,431,500 |

Query executed in 30.59 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 21. Source to concept map breakdown

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

Query executed in 0.82 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 22. cdm\_source table content

| field | 1 | 2 | 3 | 4 |
| --- | --- | --- | --- | --- |
| CDM\_SOURCE\_NAME | FDZ\_format 1 | FDZ\_format 1 | format1 | format1 |
| CDM\_SOURCE\_ABBREVIATION | format1 | format1 | MEVIS\_format1 | MEVIS\_format1 |
| CDM\_HOLDER | MEVIS\_mfinster | MEVIS\_mfinster | MEVIS | MEVIS |
| SOURCE\_DESCRIPTION |  |  | FDZ\_format1 | FDZ\_format1 |
| SOURCE\_DOCUMENTATION\_REFERENCE |  |  | Datensatzb. | Datensatzb. |
| CDM\_ETL\_REFERENCE |  |  | https://github.com/mfinster1/FDZFormat\_to\_OMOP | https://github.com/mfinster1/FDZFormat\_to\_OMOP |
| SOURCE\_RELEASE\_DATE | 2023-03-23 | 2023-03-23 | 2010-12-31 | 2010-12-31 |
| CDM\_RELEASE\_DATE | 2023-03-23 | 2023-03-23 | 2024-01-17 | 2024-01-17 |
| CDM\_VERSION | 5.4 | 5.4 | v5.4 | v5.4 |
| CDM\_VERSION\_CONCEPT\_ID | 756265 | 756265 | 756265 | 756265 |
| VOCABULARY\_VERSION | v5.0 23-JAN-23 | v5.0 23-JAN-23 | v5.0 23-JAN-23 | v5.0 23-JAN-23 |

## HADES packages

Table 23. Versions of all installed HADES R packages

| Package | Version |
| --- | --- |
| DatabaseConnector | 6.1.0 |
| ParallelLogger | 3.3.0 |
| ROhdsiWebApi | 1.3.3 |
| SqlRender | 1.15.0 |

Missing HADES packages: CohortMethod, SelfControlledCaseSeries, SelfControlledCohort, EvidenceSynthesis, PatientLevelPrediction, EnsemblePatientLevelPrediction, Capr, CirceR, CohortGenerator, PhenotypeLibrary, EmpiricalCalibration, MethodEvaluation, CohortDiagnostics, Andromeda, BigKnn, Cyclops, Eunomia, FeatureExtraction, Hydra, OhdsiSharing

## System Information

Installed R version: R version 4.2.3 (2023-03-15 ucrt)

System CPU vendor: GenuineIntel

System CPU model: 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz

System CPU number of cores: 8

System RAM: 34.07 GB

DBMS: postgresql

WebAPI version: unknown

## Vocabulary Query Performance

The number of 'Maps To' relations is equal to 4372361. This query was executed in 78.06 secs

## Achilles Query Performance

Table 24. Execution time of queries of the Achilles R-Package

| ID | NAME | DURATION |
| --- | --- | --- |
| 0 | Source name | 6.87 |
| 1 | Number of persons | 2.61 |
| 2 | Number of persons by gender | 7.3 |
| 3 | Number of persons by year of birth | 7.28 |
| 4 | Number of persons by race | 7.29 |
| 5 | Number of persons by ethnicity | 7.4 |
| 7 | Number of persons with invalid provider\_id | 0.45 |
| 8 | Number of persons with invalid location\_id | 0.98 |
| 9 | Number of persons with invalid care\_site\_id | 0.44 |
| 10 | Number of all persons by year of birth by gender | 7.93 |
| 11 | Number of non-deceased persons by year of birth by gender | 8.03 |
| 12 | Number of persons by race and ethnicity | 8.14 |
| 101 | Number of persons by age, with age at first observation period | 36.9 |
| 102 | Number of persons by gender by age, with age at first observation period | 36.41 |
| 108 | Number of persons by length of observation period, in 30d increments | 24.04 |
| 109 | Number of persons with continuous observation in each year | 47.35 |
| 110 | Number of persons with continuous observation in each month | 19.66 |
| 111 | Number of persons by observation period start month | 13.34 |
| 112 | Number of persons by observation period end month | 13.56 |
| 113 | Number of persons by number of observation periods | 15.14 |
| 114 | Number of persons with observation period before year-of-birth | 24.59 |
| 115 | Number of persons with observation period end < observation period start | 0.41 |
| 116 | Number of persons with at least one day of observation in each year by gender and age decile | 41.64 |
| 117 | Number of persons with at least one day of observation in each month | 5.12 |
| 118 | Number of observation periods with invalid person\_id | 11.77 |
| 119 | Number of observation period records by period\_type\_concept\_id | 0.77 |
| 200 | Number of persons with at least one visit occurrence, by visit\_concept\_id | 36.58 |
| 201 | Number of visit occurrence records, by visit\_concept\_id | 23.01 |
| 202 | Number of persons by visit occurrence start month, by visit\_concept\_id | 41.79 |
| 204 | Number of persons with at least one visit occurrence, by visit\_concept\_id by calendar year by gender by age decile | 59.96 |
| 207 | Number of visit records with invalid person\_id | 12.25 |
| 209 | Number of visit records with invalid care\_site\_id | 0.85 |
| 210 | Number of visit\_occurrence records outside a valid observation period | 9.86 |
| 211 | Number of visit records with end date < start date | 0.71 |
| 212 | Number of persons with at least one visit occurrence, by calendar year by gender by age decile | 40.4 |
| 220 | Number of visit occurrence records by visit occurrence start month | 26.15 |
| 221 | Number of persons by visit start year | 36.2 |
| 225 | Number of visit\_occurrence records by visit\_source\_concept\_id | 21.71 |
| 226 | Number of records by domain by visit\_concept\_id | 2.34 |
| 230 | Number of visit\_occurrence records inside valid observation period | 9.96 |
| 231 | Proportion of people with at least one visit\_occurrence record outside a valid observation period | 17.8 |
| 232 | Proportion of visit\_occurrence records outside a valid observation period | 10.57 |
| 300 | Number of providers | 0.03 |
| 301 | Number of providers by specialty concept\_id | 0.03 |
| 303 | Number of providers records by specialty\_concept\_id and visit\_concept\_id | 0.13 |
| 325 | Number of provider records by specialty\_source\_concept\_id | 0.03 |
| 400 | Number of persons with at least one condition occurrence, by condition\_concept\_id | 31.32 |
| 401 | Number of condition occurrence records, by condition\_concept\_id | 21.67 |
| 402 | Number of persons by condition occurrence start month, by condition\_concept\_id | 38.85 |
| 404 | Number of persons with at least one condition occurrence, by condition\_concept\_id by calendar year by gender by age decile | 58.6 |
| 405 | Number of condition occurrence records, by condition\_concept\_id by condition\_type\_concept\_id | 22.38 |
| 409 | Number of condition occurrence records with invalid person\_id | 12.05 |
| 410 | Number of condition occurrence records outside valid observation period | 9.74 |
| 411 | Number of condition occurrence records with end date < start date | 0.65 |
| 412 | Number of condition occurrence records with invalid provider\_id | 0.78 |
| 413 | Number of condition occurrence records with invalid visit\_id | 23.1 |
| 414 | Number of condition occurrence records, by condition\_status\_concept\_id | 21.8 |
| 415 | Number of condition occurrence records, by condition\_type\_concept\_id | 21.58 |
| 416 | Number of condition occurrence records, by condition\_status\_concept\_id, condition\_type\_concept\_id | 22.01 |
| 420 | Number of condition occurrence records by condition occurrence start month | 15.09 |
| 425 | Number of condition\_occurrence records by condition\_source\_concept\_id | 21.83 |
| 430 | Number of condition\_occurrence records inside a valid observation period | 11.76 |
| 431 | Proportion of people with at least one condition\_occurrence record outside a valid observation period | 17.77 |
| 432 | Proportion of condition\_occurrence records outside a valid observation period | 10.12 |
| 500 | Number of persons with death, by cause\_concept\_id | 8.87 |
| 501 | Number of records of death, by cause\_concept\_id | 6.25 |
| 502 | Number of persons by death month | 10.27 |
| 504 | Number of persons with a death, by calendar year by gender by age decile | 23.29 |
| 505 | Number of death records, by death\_type\_concept\_id | 6.35 |
| 509 | Number of death records with invalid person\_id | 10.18 |
| 510 | Number of death records outside valid observation period | 7.12 |
| 525 | Number of death records by cause\_source\_concept\_id | 0.41 |
| 530 | Number of death records inside a valid observation period | 6.21 |
| 531 | Proportion of people with at least one death record outside a valid observation period | 8.58 |
| 532 | Proportion of death records that occur outside a valid observation period | 7.21 |
| 600 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id | 9.11 |
| 601 | Number of procedure occurrence records, by procedure\_concept\_id | 6.31 |
| 602 | Number of persons by procedure occurrence start month, by procedure\_concept\_id | 10.41 |
| 604 | Number of persons with at least one procedure occurrence, by procedure\_concept\_id by calendar year by gender by age decile | 24.13 |
| 605 | Number of procedure occurrence records, by procedure\_concept\_id by procedure\_type\_concept\_id | 6.59 |
| 609 | Number of procedure occurrence records with invalid person\_id | 10.05 |
| 610 | Number of procedure occurrence records outside valid observation period | 7.32 |
| 612 | Number of procedure occurrence records with invalid provider\_id | 0.26 |
| 613 | Number of procedure occurrence records with invalid visit\_id | 19.6 |
| 620 | Number of procedure occurrence records by procedure occurrence start month | 7.06 |
| 625 | Number of procedure\_occurrence records by procedure\_source\_concept\_id | 6.36 |
| 630 | Number of procedure\_occurrence records inside a valid observation period | 6.1 |
| 631 | Proportion of people with at least one procedure\_occurrence record outside a valid observation period | 8.51 |
| 632 | Proportion of procedure\_occurrence records outside a valid observation period | 7.48 |
| 691 | Percentage of total persons that have at least x procedures | 12.26 |
| 700 | Number of persons with at least one drug exposure, by drug\_concept\_id | 8.9 |
| 701 | Number of drug exposure records, by drug\_concept\_id | 6.45 |
| 702 | Number of persons by drug exposure start month, by drug\_concept\_id | 10.99 |
| 704 | Number of persons with at least one drug exposure, by drug\_concept\_id by calendar year by gender by age decile | 23.67 |
| 705 | Number of drug exposure records, by drug\_concept\_id by drug\_type\_concept\_id | 6.73 |
| 709 | Number of drug exposure records with invalid person\_id | 10 |
| 710 | Number of drug exposure records outside valid observation period | 7.68 |
| 711 | Number of drug exposure records with end date < start date | 0.27 |
| 712 | Number of drug exposure records with invalid provider\_id | 0.27 |
| 713 | Number of drug exposure records with invalid visit\_id | 14.78 |
| 720 | Number of drug exposure records by drug exposure start month | 7.44 |
| 725 | Number of drug\_exposure records by drug\_source\_concept\_id | 6.56 |
| 730 | Number of drug\_exposure records inside a valid observation period | 6.33 |
| 731 | Proportion of people with at least one drug\_exposure record outside a valid observation period | 8.4 |
| 732 | Proportion of drug\_exposure records outside a valid observation period | 7.45 |
| 791 | Percentage of total persons that have at least x drug exposures | 12.07 |
| 800 | Number of persons with at least one observation occurrence, by observation\_concept\_id | 42.22 |
| 801 | Number of observation occurrence records, by observation\_concept\_id | 24.58 |
| 802 | Number of persons by observation occurrence start month, by observation\_concept\_id | 52.89 |
| 804 | Number of persons with at least one observation occurrence, by observation\_concept\_id by calendar year by gender by age decile | 1.16 |
| 805 | Number of observation occurrence records, by observation\_concept\_id by observation\_type\_concept\_id | 25.65 |
| 807 | Number of observation occurrence records, by observation\_concept\_id and unit\_concept\_id | 25.59 |
| 809 | Number of observation records with invalid person\_id | 13.99 |
| 810 | Number of observation records outside valid observation period | 11.79 |
| 812 | Number of observation records with invalid provider\_id | 1.3 |
| 813 | Number of observation records with invalid visit\_id | 20.93 |
| 814 | Number of observation records with no value (numeric, string, or concept) | 1.28 |
| 820 | Number of observation records by observation start month | 30.08 |
| 822 | Number of observation records, by observation\_concept\_id and value\_as\_concept\_id | 25.33 |
| 823 | Number of observation records, by observation\_concept\_id and qualifier\_concept\_id | 25.21 |
| 825 | Number of observation records by observation\_source\_concept\_id | 25.49 |
| 826 | Number of observation records by value\_as\_concept\_id | 25.42 |
| 827 | Number of observation records by unit\_concept\_id | 24.99 |
| 830 | Number of observation records inside a valid observation period | 23.99 |
| 831 | Proportion of people with at least one observation record outside a valid observation period | 25.79 |
| 832 | Proportion of observation records outside a valid observation period | 12.13 |
| 891 | Percentage of total persons that have at least x observations | 40.33 |
| 900 | Number of persons with at least one drug era, by drug\_concept\_id | 1.27 |
| 901 | Number of drug era records, by drug\_concept\_id | 1.27 |
| 902 | Number of persons by drug era start month, by drug\_concept\_id | 1.27 |
| 904 | Number of persons with at least one drug era, by drug\_concept\_id by calendar year by gender by age decile | 0.11 |
| 908 | Number of drug eras without valid person | 0.04 |
| 910 | Number of drug\_era records outside valid observation period | 0.08 |
| 911 | Number of drug eras with end date < start date | 0.06 |
| 920 | Number of drug era records by drug era start month | 1.33 |
| 930 | Number of drug\_era records inside a valid observation period | 1.2 |
| 931 | Proportion of people with at least one drug\_era record outside a valid observation period | 0.06 |
| 932 | Proportion of drug\_era records outside a valid observation period | 0.07 |
| 1,000 | Number of persons with at least one condition era, by condition\_concept\_id | 1.24 |
| 1,001 | Number of condition era records, by condition\_concept\_id | 1.27 |
| 1,002 | Number of persons by condition era start month, by condition\_concept\_id | 1.3 |
| 1,004 | Number of persons with at least one condition era, by condition\_concept\_id by calendar year by gender by age decile | 0.13 |
| 1,008 | Number of condition eras without valid person | 0.05 |
| 1,010 | Number of condition\_era records outside a valid observation period | 0.05 |
| 1,011 | Number of condition eras with end date < start date | 0.04 |
| 1,020 | Number of condition era records by condition era start month | 1.41 |
| 1,030 | Number of condition\_era records inside a valid observation period | 1.22 |
| 1,031 | Proportion of people with at least one condition\_era record outside a valid observation period | 0.06 |
| 1,032 | Proportion of condition\_era records outside a valid observation period | 0.06 |
| 1,100 | Number of persons by location 3-digit zip | 11.76 |
| 1,101 | Number of persons by location state | 11.72 |
| 1,102 | Number of care sites by location 3-digit zip | 0.05 |
| 1,103 | Number of care sites by location state | 0.04 |
| 1,200 | Number of persons by place of service | 0.42 |
| 1,201 | Number of visits by place of service | 0.82 |
| 1,202 | Number of care sites by place of service | 0.03 |
| 1,203 | Number of visits by place of service discharge type | 0.85 |
| 1,300 | Number of persons with at least one visit detail, by visit\_detail\_concept\_id | 0.38 |
| 1,301 | Number of visit detail records, by visit\_detail\_concept\_id | 0.13 |
| 1,302 | Number of persons by visit detail start month, by visit\_detail\_concept\_id | 0.12 |
| 1,304 | Number of persons with at least one visit detail, by visit\_detail\_concept\_id by calendar year by gender by age decile | 0.12 |
| 1,307 | Number of visit records with invalid person\_id | 0.04 |
| 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 | 0.05 |
| 1,311 | Number of visit\_detail records with end date < start date | 0.04 |
| 1,312 | Number of persons with at least one visit detail, by calendar year by gender by age decile | 0.14 |
| 1,320 | Number of visit detail records by visit detail start month | 0.13 |
| 1,321 | Number of persons by visit start year | 0.17 |
| 1,325 | Number of visit\_detail records by visit\_detail\_source\_concept\_id | 0.12 |
| 1,326 | Number of records by domain by visit\_detail\_concept\_id | 6.1 |
| 1,330 | Number of visit\_detail records inside a valid observation period | 0.14 |
| 1,331 | Proportion of people with at least one visit\_detail record outside a valid observation period | 0.06 |
| 1,332 | Proportion of visit\_detail records outside a valid observation period | 0.08 |
| 1,408 | Number of persons by length of payer plan period, in 30d increments | 34.29 |
| 1,409 | Number of persons with continuous payer plan in each year | 27.78 |
| 1,410 | Number of persons with continuous payer plan in each month | 49.31 |
| 1,411 | Number of persons by payer plan period start month | 21.86 |
| 1,412 | Number of persons by payer plan period end month | 21.65 |
| 1,413 | Number of persons by number of payer plan periods | 28.35 |
| 1,414 | Number of persons with payer plan period before year-of-birth | 24.71 |
| 1,415 | Number of persons with payer plan period end < payer plan period start | 0.43 |
| 1,425 | Number of payer\_plan\_period records by payer\_source\_concept\_id | 0.76 |
| 1,800 | Number of persons with at least one measurement occurrence, by measurement\_concept\_id | 0.66 |
| 1,801 | Number of measurement occurrence records, by measurement\_concept\_id | 0.68 |
| 1,802 | Number of persons by measurement occurrence start month, by measurement\_concept\_id | 0.66 |
| 1,804 | Number of persons with at least one mesurement occurrence, by measurement\_concept\_id by calendar year by gender by age decile | 0.55 |
| 1,805 | Number of measurement occurrence records, by measurement\_concept\_id by measurement\_type\_concept\_id | 0.74 |
| 1,807 | Number of measurement occurrence records, by measurement\_concept\_id and unit\_concept\_id | 0.75 |
| 1,809 | Number of measurement records with invalid person\_id | 0.08 |
| 1,810 | Number of measurement records outside valid observation period | 4.68 |
| 1,811 | Number of measurement records with a valid value (with a mapped, non-null value\_as\_number) | 0.35 |
| 1,812 | Number of measurement records with invalid provider\_id | 0.04 |
| 1,813 | Number of measurement records with invalid visit\_id | 0.16 |
| 1,814 | Number of measurement records with no value (numeric, string, or concept) | 0.44 |
| 1,818 | Number of measurement records below/within/above normal range, by measurement\_concept\_id and unit\_concept\_id | 0.52 |
| 1,819 | Number of measurement records, by measurement\_concept\_id, with a valid value (with a mapped, non-null value\_as\_number) | 0.38 |
| 1,820 | Number of measurement records by measurement start month | 0.69 |
| 1,821 | Number of measurement records with no numeric value | 0.04 |
| 1,822 | Number of measurement records, by measurement\_concept\_id and value\_as\_concept\_id | 0.73 |
| 1,823 | Number of measurement records, by measurement\_concept\_id and operator\_concept\_id | 0.71 |
| 1,825 | Number of measurement records by measurement\_source\_concept\_id | 0.74 |
| 1,826 | Number of measurement records by value\_as\_concept\_id | 0.74 |
| 1,827 | Number of measurement records by unit\_concept\_id | 0.69 |
| 1,830 | Number of visit\_detail records inside a valid observation period | 0.6 |
| 1,831 | Proportion of people with at least one measurement record outside a valid observation period | 4.73 |
| 1,832 | Proportion of measurement records outside a valid observation period | 4.69 |
| 1,833 | Proportion of measurement records inside a valid observation period and without a value | 0.72 |
| 1,891 | Percentage of total persons that have at least x measurements | 0.73 |
| 1,900 | Source values mapped to concept\_id 0 by table, by column, by source\_value | 1.03 |
| 2,000 | Number of patients with at least 1 Dx and 1 Rx | 35.36 |
| 2,001 | Number of patients with at least 1 Dx and 1 Proc | 34.97 |
| 2,002 | Number of patients with at least 1 Meas, 1 Dx and 1 Rx | 31.76 |
| 2,003 | Number of patients with at least 1 Visit | 31.21 |
| 2,004 | Number of distinct patients that overlap between specific domains | 19.23 |
| 2,100 | Number of persons with at least one device exposure, by device\_concept\_id | 0.15 |
| 2,101 | Number of device exposure records, by device\_concept\_id | 0.11 |
| 2,102 | Number of persons by device records start month, by device\_concept\_id | 0.13 |
| 2,104 | Number of persons with at least one device exposure, by device\_concept\_id by calendar year by gender by age decile | 0.13 |
| 2,105 | Number of device exposure records, by device\_concept\_id by device\_type\_concept\_id | 0.12 |
| 2,110 | Number of device\_exposure records outside valid observation period | 0.06 |
| 2,125 | Number of device\_exposure records by device\_source\_concept\_id | 0.12 |
| 2,130 | Number of device\_exposure records inside a valid observation period | 0.12 |
| 2,131 | Proportion of people with at least one device\_exposure record outside a valid observation period | 0.06 |
| 2,132 | Proportion of device\_exposure records outside a valid observation period | 0.08 |
| 2,191 | Percentage of total persons that have at least x device exposures | 0.12 |
| 2,200 | Number of persons with at least one note by note\_type\_concept\_id | 0.04 |
| 2,201 | Number of note records, by note\_type\_concept\_id | 0.04 |

Query executed in 1.26 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: