**.** ****

**OMOP Common Data Model (CDM V5.0)**

**Swedish Health Registry Mapping Specification**

**20 July 2016**

**Last Update: 18 July 2017**

**Am**

**Am**

**Am**

Jolanda Strubel, Maxim Moinat

The Hyve, Utrecht, The Netherlands

Contents

[1.0 Abbreviations used: 3](#_Toc488143442)

[2.0 Introduction 3](#_Toc488143444)

[3.0 Source Data Mapping Approach 3](#_Toc488143445)

[3.1 Vocabulary mapping 4](#_Toc488143447)

[4.0 Source Data and ETL 5](#_Toc488143448)

[4.1 Table: person 6](#_Toc488143450)

[4.2 Table: observation\_period 7](#_Toc488143451)

[4.3 Table: visit\_occurrence 7](#_Toc488143452)

[4.4 Table: drug\_exposure 8](#_Toc488143453)

[4.5 Table: condition\_occurrence 9](#_Toc488143475)

[4.6 Table: procedure\_occurrence 10](#_Toc488143550)

[4.7 Table: death 11](#_Toc488143551)

[4.8 Table: measurement 12](#_Toc488143552)

[4.9 Table: observation 13](#_Toc488143553)

[4.10 Table: provider 19](#_Toc488143554)

[4.11 Table: care\_site 19](#_Toc488143556)

[4.12 Table: location 19](#_Toc488143557)

[4.13 Table: drug\_era 20](#_Toc488143558)

[4.14 Table: condition\_era 20](#_Toc488143559)

# Abbreviations used:

CDM Common Data Model

ETL Export Transform Load

# Introduction

This document reflects the requirements, assumptions, business rules and transformations for the implementation of the Common Data Model Version 5.0 (CDM) as implemented by The Hyve.

The purpose of this document is to describe the ETL mapping of the data from Swedish registries to the OMOP Common Data Model. It is based on the OMOP ETL Specifications. General information that is covered by the OMOP ETL Specification will not be covered in this document, but a detailed discussion of this project specific aspects of mapping and converting data to the standard CDM is provided.

The document is composed of a section per target OMOP CDM table. In each section, the tables and their mapping are individually reviewed along with any source specific rules and exceptions.

# Source Data Mapping Approach

This section covers the high-level assumptions and approach to extraction, transformation and loading (ETL) of raw source data into the Common Data Model (CDM).

In the table below an overview is given of which source tables were mapped to which OMOP table:

|  |  |  |
| --- | --- | --- |
| **Source table** | **Table description** | **CDM table** |
| patient\_sluten  patient\_oppen  patient\_dag\_kiru | Patient registers for “hospital care” (1977-2013), “out patients” (2001-2013) and “day care” (1997-2000) | Observation period  Visit\_occurence  Condition occurence  Device exposure  Procedure occurence  Measurement  Observation |
| lisa2013 | Socio economic register at 01.-01-2013 for the data cut | Person  Measurement  Observation  Provider  Care site  location |
| lmed2005-lmed2015 | Drug register 2005 to 2015 | Drug exposure  Provider  Measurement |
| dod\_6113 | Death register 2005-2014 | Death |
| dod-2015 | Death register 2015 only; only a few of the variables | Death |

In addition, lists of all providers, hospitals and counties were retrieved from the White Rabbit reports to fill the provider, care\_site and location tables respectively.

General na

## Vocabulary mapping

We mapped four source vocabularies to standard vocabularies in OMOP. An overview of the mapping results is given in the table below. In addition to these vocabularies, some small code lists were also mapped to standard OMOP concepts (e.g. gender and visit type).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Type** | **Source Vocabulary** | **Target Vocabulary** | **Target Class** | **Mapping type** | **Number of concepts mapped** | **Coverage** |
| Condition, Ekod, Morsak | ICD10-SE | SNOMED | Clinical Finding | Automatic | 33,485 |  |
| Procedure | NOMESCO | SNOMED | Procedure | Manual | 90 |  |
| Procedure | KVA | SNOMED | Procedure | Manual | 13 |  |
| Drug | NPL /  Varunummer | RxNorm | Ingredient, Clinical Comp/Form/Drug | Automatic | 6865 |  |
| Provider Specialty | SPKOD | Specialty | Specialty | Manual | 162 | 100% |

The ICD10-SE, NOMESCO, drug and SPKOD mappings are stored in separate csv files. The small code lists are hard coded in the ETL scripts.

#### ICD10-SE

A complete list of Swedish ICD10 codes (ICD10-SE) was retrieved from Social Styrelsen. It consists of diagnostic codes, including extension codes and the ‘External causes of disease and death’ chapter. These codes were mapped to SNOMED using the international ICD10 codes:

1. Mapping to the OMOP ICD10 vocabulary by choosing the best matching concept code. e.g. S5250 to S52.5 *‘Fracture of lower end of radius’.* The mapping is done at best on all characters and at worst on the three first characters. As can be seen in the table below, the last character was often truncated. Thus, most mappings were done on a more granular level (e.g. R029A ‘*Dermatitis gangrenosa infantum’* to R02 ‘*Gangrene, not elsewhere classified’*).
2. Mapping to SNOMED vocabulary with the existing concept relation ‘Maps to’. e.g. S52.5 to 4134322 *‘Fracture of distal end of radius’.*
3. Manual curation of the 20 most frequent unmapped concepts using Usagi.

|  |  |
| --- | --- |
| **n trailing characters truncated** | **Number of concepts mapped to SNOMED** |
| 0 | 9,645 |
| 1 | 8,643 |
| 2 | 15,197 |
| Not mapped | 362 |

#### Drugs

All Swedish drug concepts are encoded by a varunummer (article code). Here, we mapped these varunummer to corresponding RxNorm concepts using their ATC code, dose form and strength.

#### Procedures (NOMESCO and KVA)

A small subset of the NOMESCO (surgical) and KVA (non-surgical) vocabularies were mapped manually to SNOMED procedure codes. For this, the Swedish concept name was translated to English and given as input to Usagi. With this tool, an automatic suggestion was made, which was curated by hand.

For example, the code FNG05 “Perkutan transluminal koronarangioplastik (PTCA) med inläggande av stent (i)” was translated and mapped to SNOMED concept 4283892 (Placement of stent in coronary artery).

#### Provider Speciality

Concept mapping for spkod encoding. The mapping was done on the first 2 digits of the spkod code. Codes are either 2, 4 or 5 digits. 4 and 5 digit codes truncated to first 2 digits.

For some code description, a perfect match in the OMOP vocabulary for “specialties” could not be found. The following suboptimal mappings were made:

* Mapped “Allmänmedicin” (family medicine) to “General practice”
* No direct equivalent for “Företagshälsovård” (company doctor). Mapped to “Occupational Therapy”
* Mapped “Medicinska njursjukdomar” (renal diseases) to “urology”.
* “Occupational and environmental medicine” not present in vocabularies. Mapped to “Occupational Therapy”
* “Oncology” (“Onkologi”) maps to either “medical oncology”, “radiation oncology” or “surgical oncology”. Chosen for “medical oncology”.
* “skin and venereal disease” not one specialism in the vocabulary. Mapped to “Infectious Disease”
* “Bacteriology” is a subdivision of “microbiology”. For the first no concept in OMOP. Mapped to “Pathology - Medical Microbiology”
* “Clinical cytology” mapped as “cytopathology”
* “Social medicine” mapped to “sociologist” (no entry coming close regarding epidemiology)
* “Disease of teeth” mapped to “Dental Providers, Dentist”.

# Source Data and ETL

*This section describes the mapping process and ETL conversions of the data into Common Data Model.*

The following prefixes were used to indicate in which data set a certain source field is present:

PATREG: patient\_sluten, patient\_oppen, patient\_dag\_kiru

LISA: lisa2013

DRUG: lmed2005-lmed2015

DEATH: dod\_6113, dod-2015

#### Preprocessing and filtereing

The first step in the ETL is to transform the patient and death registries from a wide to a long format. For the patient registries, the diagnoses and procedures of one visit are all in one row. After transformation each diagnosis and procedure is split on separate rows and given a visit id. For the death registry the same is done for multiple causes of death.

In addition, the patient and drug registers are filtered before the ETL procedure; all rows without a specific date in admission date (indatuma) or discharge date (utdatuma) or without an age are removed.

## Table: person

All information of each person is first extracted from the patient, drug and death registers. This data is aggregated per person and logic is applied to get one value for gender, year of birth, immigration and emigration date per person. The resulting intermediate table is called PERSON\_AGG and the table below contains a description of each of the fields.

Rounding for the gender (kon) field is needed because the recorded gender is not always the same between the registries. The age can change during the year, depending on the birthday. The lowest year of birth found is considered the correct year of birth.

|  |  |  |  |
| --- | --- | --- | --- |
| **Field name** | **Logic** | **Comment** | **Aggregated from** |
| PERSON\_AGG.kon | ROUND( AVG(*kon*) ) | Kon is an integer (1 or 2). | PATREG  DRUG |
| PERSON\_AGG.year\_of\_birth | FLOOR( AVG(*ar* – *alder*) ) | Depending on the birthday, the yob can be 1 year too high. | PATREG  DRUG  DEATH |
| PERSON\_AGG.seninv | MAX(*seninv*) | If sentuv is an integer, else Null. | PATREG |
| PERSON\_AGG.senutv | MAX(*senutv*) |  | PATREG |

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| person\_id | PERSON\_AGG.lpnr  LISA.lpnr |  | Some patients not in DRUG nor in PATREG. Use LISA to fill this field. |
| gender\_source\_value | PERSON\_AGG.kon |  |  |
| gender\_concept\_id | PERSON\_AGG.kon | *See above* |  |
| year\_of\_birth | PERSON\_AGG.year\_of\_birth | *See above* | Persons without year of birth are discarded. |
| month\_of\_birth |  |  |  |
| day\_of\_birth |  |  |  |
| time\_of\_birth |  |  |  |
| race\_concept\_id |  | 8552 | unknown race |
| ethnicity\_concept\_id |  | 0 |  |
| location\_id | LISA.lan |  |  |
| provider\_id |  |  |  |
| care\_site\_id |  |  |  |
| person\_source\_value | LISA.lpnr |  |  |
| gender\_source\_concept\_id |  |  |  |
| race\_source\_value |  |  |  |
| race\_source\_concept\_id |  |  |  |
| ethnicity\_source\_value |  |  |  |
| ethnicity\_source\_concept\_id |  |  |  |

Concept mapping for source field PERSON\_AGG.kon (gender):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 1 | Male | Male | 8507 |  |
| 2 | Female | Female | 8532 |  |
| (empty) | (empty) | UNKNOWN | 8551 | If no gender found in PATREG or DRUG |

If lpnr from LISA not found in PATREG or DRUG, then it is not included in the person table. Main reason: no date of birth.

## Table: observation\_period

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| observation\_period\_id |  | *Automatically Generated* |  |
| person\_id | LISA.lpnr |  |  |
| observation\_period\_start\_date |  | Start date of the data cut: 1997-01-01 | Configurable |
| observation\_period\_end\_date |  | End date of the data cut: 2015-08-01 | Configurable |
| period\_type\_concept\_id |  | 44814724 – Period covering healthcare encounters |  |

## Table: visit\_occurrence

Visit\_occurence\_id generated

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id | PATREG.lpnr |  |  |
| visit\_occurrence\_id |  | *Automatically generated per line in PATREG.* |  |
| visit\_concept\_id |  | Derived from PATREG source file (see below) |  |
| visit\_start\_date | PATREG.indatuma |  | If date empty, use 01-01-1900. |
| visit\_start\_time |  |  |  |
| visit\_end\_date | PATREG.utdatuma |  | If date empty, use 01-01-1900. |
| visit\_end\_time |  |  |  |
| visit\_type\_concept\_id |  | 44818518 – Visit derived from EHR record |  |
| provider\_id |  |  |  |
| care\_site\_id | PATREG.sjukhus |  | See care site table for (concept) mapping |
| visit\_source\_value |  | ‘sluten’, ‘oppen’ or ‘dag kiru’ |  |
| visit\_source\_concept\_id |  |  |  |

Concept mapping for visit\_concept\_id. Determined by the source file the patient register is in (either patient\_sluten, oppen or dag-kiru).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| Patient\_sluten | Inhospital | Inpatient Visit | 9201 | Visit domain |
| Patient\_oppen | Outpatient care | Outpatient Visit | 9202 | Visit domain |
| Dag\_kiru | Day Surgery | Outpatient Visit | 9202 | Or 45878057 (“Same day/Ambulatory Surgery”) from ‘Maes Value’ domain. |

## Table: drug\_exposure

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| drug\_exposure\_id |  | *Automatically Generated* |  |
| person\_id | DRUG.lpnr |  |  |
| drug\_concept\_id | DRUG.varunr | Mapped to RxNorm |  |
| drug\_type\_concept\_id |  | 43542356 Physician administered drug (identified from EHR problem list) |  |
| drug\_source\_value | DRUG.varunr  DRUG.lnamn |  | Field allows just 50 characters. Lnamn is trimmed. |
| quantity | DRUG.forpstl  DRUG.antal | Integer(forpstl) \* antal.  Forpstl is a text field. First integer is taken. Quantity is null if forpstl contains reference to volume or weight instead of number of pills (‘mi’, ‘,’ or ‘gr’). | If antal is negative, the record is omitted. |
| drug\_exposure\_start\_date | DRUG.edatum |  |  |
| dose\_unit\_concept\_id | DRUG.styrka\_enh | Mapped to OMOP vocabulary |  |
| dose\_unit\_source\_value | DRUG.styrka\_enh |  |  |
| provider\_id | DRUG.spkod1 | Mapped to provider table.  If not mappable, set to an unknown provider (id=99999) |  |
| days\_supply | DRUG.forpstl  DRUG.antal | Same as quantity. | Assumes that the daily prescription is 1 |
| sig | DRUG.doser |  | Raw text with prescription directions. |
| effective\_drug\_dose | DRUG.styrknum |  | Numeric strength, unmodified |
| route\_source\_value | DRUG.lformgrupp |  |  |

Concept mapping for source field DRUG.styrka\_enh: (SNOMED)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| mg | Milligram | Milligram | 8576 |  |
| mikrog | Microgram | Microgram | 9655 |  |
| g | Gram | Gram | 8504 |  |
| (All other) |  | Other | 9177 |  |

## Table: condition\_occurrence

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id | PATREG.lpnr |  |  |
| condition\_start\_date | PATREG.indatuma |  |  |
| condition\_occurrence\_id |  | *Automatically generated* |  |
| condition\_concept\_id | PATREG.hdia  PATREG.bdia# | Trailing ‘-‘, ‘X’, ‘P’ and ‘T’ are removed first, then mapped from ICD10-SE to SNOMED.  0 if condition could not be mapped |  |
| condition\_source\_concept\_id | PATREG.hdia  PATREG.bdia# | Mapped from ICD10-SE to ICD10 concept (international). | ICD10 concept\_id. Used as intermediate between ICD10-SE and SNOMED. |
| condition\_type\_concept\_id |  | *See table below* | Distinguishes between hdia, bdia1 and the other bdia# |
| condition\_end\_date |  |  |  |
| condition\_source\_value | PATREG.hdia  PATREG.bdia# |  |  |
| stop\_reason |  |  |  |
| provider\_id |  |  |  |
| visit\_occurrence\_id |  | *Automatically generated* |  |

Based on the column, the condition codes were given a condition\_type\_concept\_id. The following table was used for the mapping.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source variable name** | **Source Description** | **OMOP Condition type description** | **OMOP Condition type Concept ID** | **Remarks** |
| hdia | main diagnosis | Primary Condition | 44786627 |  |
| bdia1 | secondary diagnosis | First Position Condition | 44786628 | bdia at the first position. To distinquish bdia1 from the rest. |
| bdia# | secondary diagnosis | Secondary Condition | 44786629 | All bdia2 and higher numbers become secondary condition types |



## Table: procedure\_occurrence

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id | PATREG.lpnr |  |  |
| procedure\_occurrence\_id |  |  |  |
| procedure\_date | PATREG.utdatuma |  | Precise date of procedure is unknown, but must be in between indatum and utdatum. As a convention, we use utdatum. |
| procedure\_source\_concept\_id |  |  |  |
| procedure\_concept\_id | PATREG.op# | Mapped from NOMESCO to SNOMED concept. |  |
| procedure\_type\_concept\_id |  | ‘op1’ = 44786630  ‘op2’ = 44786631  ‘op#’ = 44786631 | Distinction between primary and secondary procedure. |
| visit\_occurrence\_id |  | *Automatically generated* |  |
| modifier\_concept\_id |  |  |  |
| quantity |  |  |  |
| provider\_id |  |  |  |
| procedure\_source\_value | PATREG.op# |  |  |
| qualifier\_source\_value |  |  |  |







































## Table: death

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id | DEATH.lpnr |  |  |
| death\_date | DEATH.dodsdat | Trailing zeros are mapped to middle of time period  20071200 => 2007-12-15  20070000 => 2007-06-01 | Unknown (empty) death date not allowed. |
| death\_type\_concept\_id |  | 38003569 (EHR record patient status "Deceased") |  |
| cause\_concept\_id | DEATH.ulorsak | Trailing ‘-‘, ‘X’, ‘P’ and ‘T’ are removed first, then mapped from ICD10-SE to SNOMED. | Both missing and unmappable concepts are given concept\_id 0. |
| cause\_source\_value | DEATH.ulorsak |  |  |
| cause\_source\_concept\_id |  |  |  |

An addendum table is created as addition to the OMOP CDM to store the circumstances of the death.

#### Death addendum table

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Type | Source Field | Comment |
| person\_id | integer | DEATH.lpnr | Links to death table |
| alcohol | boolean | DEATH.alkohol |  |
| narcotic | boolean | DEATH.narkotik |  |
| work | boolean | DEATH. Aolycka |  |
| abroad | boolean | DEATH.dodutl |  |
| surgical\_procedure | boolean | DEATH.opererad |  |
| place\_of\_death\_concept\_id | integer | DEATH.dodspl |  |



Mapping for DEATH.dodspl (place of death). Concepts are from the place of service domain.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 1 | hospital | Inpatient Hospital | 8717 |  |
| 2 | old peoples home, rehabilitation clinic etc | Other Place of Service | 8844 |  |
| 3 | private home | Home | 8536 |  |
| 4 | other/unknown | Other | 44814649 | Class - ‘undefined’  Domain - ‘Observation’ |
| *(empty)* |  | Other | 44814649 | Class - ‘undefined’  Domain - ‘Observation’ |

**Secondary death causes**

All additional causes of death (Morsak**)** are saved as observation with observation\_concept\_id 4083743 (Cause of Death; SNOMED Clinical Finding). The SNOMED concept\_id for the death cause is saved as value\_as\_concept\_id. See section 4.9 Observation.

## Table: measurement



#### Income

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| measurement\_id |  | *Automatically generated* |  |
| person\_id | LISA.lpnr |  |  |
| measurement\_concept\_id |  | 4073460 (Income) |  |
| measurement\_date |  | First day of the year of the lisa file. |  |
| measurement\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| value\_as\_number | LISA.dispinkpersf04 |  |  |
| unit\_concept\_id |  | 44818647 (Currency: Swedish krona/kronor) |  |
| measurement\_source\_value |  | ‘dispinkpersf04’ |  |
| value\_source\_value | LISA.dispinkpersf04 |  |  |

Mappings made for the measurement\_concept\_id and unit\_concept\_id.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source Variable name** | **Description** | **OMOP DESCRIPTION** | **OMOP CONCEPT\_ID** | **Remarks** |
| LISA.dispinkpersf04 | Income | Individual income | 4073460 |  |
| LISA.dispinkpersf04 | Unit | Swedish krona/kronor | 44818647 |  |

## Table: observation

Nine different data fields are loaded into the observation table. From the Lisa registry, the fields ethnic background, education and work status are loaded. From the patient registry, the fields with hospitalization status, discharge location, planned visit and marital status. From the death registry, the additional causes of death are recorded. Below are the syntactic mappings for each data field. After the syntactic mapping, the semantic mappings are listed in several tables.

#### Ethnic Background

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | LISA.lpnr |  |  |
| observation\_concept\_id | LISA.utlsvbakgalt | 11,12 = Foreign background  21,22,23 = Swedish background |  |
| observation\_date |  | First day of the year of the lisa file. |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | LISA.utlsvbakgalt |  |  |
| qualifier\_source\_value |  | ‘utlsvbakgalt’ | The column name as string of the observation column. |



Concept mapping for source field LISA.utlsvbakgalt:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 11 | born abroad by two parents who also were born abroad | Immigrant | 4058588 | Codes 11 and 12 are counted as foreign background |
| 12 | born in Sweden by two parents who were born abroad | Immigrant | 4058588 |  |
| 21 | Born abroad, at least one of the parents was born in Sweden | Native | 43021808 | codes 21,22,23 are counted as Swedish background |
| 22 | Born in Sweden, one parent born in Sweden, the other born abroad | Native | 43021808 |  |
| 23 | Born in Sweden, both parents also born in Sweden | Native | 43021808 |  |

#### Education

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | LISA.lpnr |  |  |
| observation\_concept\_id | LISA.sun2000niva | See mapping below |  |
| observation\_date |  | First day of the year of the lisa file. |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | LISA.sun2000niva |  |  |
| value\_as\_concept\_id |  | 4188539 (Yes to suggestive statement) |  |
| qualifier\_source\_value |  | ‘sun2000niva’ |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

#### Work Status

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | LISA.lpnr |  |  |
| observation\_concept\_id | LISA.syssstat11 | See mapping below |  |
| observation\_date |  | First day of the year of the lisa file. |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | LISA.syssstat11 |  |  |
| value\_as\_concept\_id |  | 4188539 (Yes to suggestive statement) |  |
| qualifier\_source\_value |  | ‘syssstat11’ |  |

Concept mapping for source field LISA.syssstat11

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 1 | Working | | Employed | 4076340 | SNOMED, Clinical Finding |
| 5 | Not Working, with tax previous year | Unemployed | | 4251171 | SNOMED, Clinical Finding |
| 6 | Not Working, without tax previous year | | Unemployed | 4251171 | SNOMED, Clinical Finding |

#### Insatt

From patient\_sluten

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | PATREG.lpnr |  |  |
| observation\_concept\_id | PATREG.insatt | See mapping below |  |
| observation\_date | PATREG.indatuma |  |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | PATREG.insatt |  |  |
| value\_as\_concept\_id |  | 4188539 (Yes to suggestive statement) |  |
| qualifier\_source\_value |  | ‘insatt’ |  |
| visit\_occurrence\_id |  | *Automatically generated* |  |

Concept mapping for source field PATREG.INSATT: (SNOMED)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 1 | Other hospital or clinic | Hospital admission, transfer from other hospital or health care facility. | 4164916 |  |
| 2 | Home for disabled or elderly people  (särskilt boende) | Hospital Admission | 8715 |  |
| 3 | Own home | Hospital Admission | 8715 |  |

#### Utsatt

From patient\_sluten

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | PATREG.lpnr |  |  |
| observation\_concept\_id | PATREG. utsatt | See mapping below |  |
| observation\_date | PATREG.utdatuma |  |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | PATREG.utsatt |  |  |
| value\_as\_concept\_id |  | 4188539 (Yes to suggestive statement) |  |
| qualifier\_source\_value |  | ‘utsatt’ |  |
| visit\_occurrence\_id |  | *Automatically generated* |  |

Concept mapping for source field PATREG.UTSATT: (SNOMED)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 1 | to other hospital or clinic | Discharge to hospital | 4142018 |  |
| 2 | to home for elderly or disabled | Discharge to nursing home | 4143443 |  |
| 3 | ordinary living (ordinärt boende) | Discharge to home | 4140634 |  |
| 4 | died during hospital stay | Patient died in hospital | 4081608 |  |

#### Pvard

From patient\_sluten and patient\_oppen

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | PATREG.lpnr |  |  |
| observation\_concept\_id | PATREG. pvard | See mapping below |  |
| observation\_date | PATREG.indatuma |  |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | PATREG.pvard |  |  |
| value\_as\_concept\_id |  | 4188539 (Yes to suggestive statement) |  |
| qualifier\_source\_value |  | ‘pvard’ |  |
| visit\_occurrence\_id |  | *Automatically generated* |  |

Concept mapping for source field PATREG.pvard (visit planned): SNOMED (available OMOP visit type concepts not applicable)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| 1 | Yes, planned | Planned admission | 4228491 |  |
| 2 | No, not planned | Unplanned local admission | 44803024 |  |
| (empty) |  |  | 0 |  |

#### Ekod (accident)

From patient\_sluten and patient\_oppen

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | PATREG.lpnr |  |  |
| observation\_concept\_id | PATREG. ekod1-ekod5 | Trailing ‘-‘, ‘X’, ‘P’ and ‘T’ are removed first, then mapped to SNOMED |  |
| observation\_date | PATREG.indatuma |  | Actual date is not known, use intake date. |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | PATREG.ekod1-ekod5 |  |  |
| observation\_source\_concept\_id | PATREG. ekod1-ekod5 | Mapped from ICD10-SE to SNOMED |  |
| qualifier\_source\_value |  | ‘ekod’ |  |
| visit\_occurrence\_id |  | *Automatically generated* |  |

#### Marriage status

From patient\_sluten and patient\_oppen

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | PATREG.lpnr |  |  |
| observation\_concept\_id | PATREG.civil | See mapping below |  |
| observation\_date | PATREG.indatuma |  |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| observation\_source\_value | PATREG.civil |  |  |
| value\_as\_concept\_id |  | 4188539 (Yes to suggestive statement) |  |
| qualifier\_source\_value |  | ‘civil’ |  |
| visit\_occurrence\_id |  | *Automatically generated* |  |

Concept mapping for source field PATREG.civil:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | **Source Description** | **OMOP description** | **OMOP Concept ID** | **Remarks** |
| G | married | Married | 4338692 |  |
| OG | not married | Single person | 4053842 |  |
| O | not married |  | 0 |  |
| S | divorced | Divorced | 4069297 |  |
| Ä | widow/widower | Widowed | 4143188 |  |
| RP | registered partnership | Domestic partnership | 4325710 |  |
| SP | divorced partnership | Divorced | 4069297 |  |
| EP | widowed partnership | Widowed | 4143188 |  |

#### Additional death causes

From patient\_sluten and patient\_oppen

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| observation\_id |  | *Automatically generated* |  |
| person\_id | DEATH.lpnr |  |  |
| observation\_concept\_id |  | 4083743 (Cause of Death) |  |
| observation\_date | DEATH.dodsdat | Trailing zeroes converted to middle of time period. (examples see Death table section) |  |
| observation\_type\_concept\_id |  | 38000280 (Observation recorded from EHR) |  |
| value\_as\_string | DEATH.morsak1-51 |  |  |
| value\_as\_concept\_id | DEATH.morsak1-51 | Trailing ‘-‘, ‘X’, ‘P’ and ‘T’ are removed first, then mapped from ICD10-SE to SNOMED |  |
| observation\_source\_value |  | ‘morsak<#>’ |  |

## Table: provider































List of providers extracted from drug files.

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| provider\_id | Spkod1 |  |  |
| provider\_name | source\_name |  |  |
| npi |  |  |  |
| dea |  |  |  |
| specialty\_concept\_id | Spkod1 | Mapped to specialty concept |  |
| care\_site\_id |  |  |  |
| year\_of\_birth |  |  |  |
| gender\_concept\_id |  |  |  |
| provider\_source\_value |  |  |  |
| specialty\_source\_value | source\_name |  |  |
| specialty\_source\_concept\_id |  |  |  |
| gender\_source\_value |  |  |  |
| gender\_source\_concept\_id |  |  |  |

## Table: care\_site

Reading from *sjukhus\_care\_site.csv*

List of hospitals (Sjukhus)

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| care\_site\_id |  | *Automatically generated* |  |
| care\_site\_name | Sjukhus\_name |  |  |
| place\_of\_service\_concept\_id |  |  |  |
| location\_id | lan\_kom | Mapped to location table |  |
| care\_site\_source\_value | Sjukhus\_id |  |  |
| place\_of\_service\_source\_value |  |  |  |

## Table: location

Reading from *lan\_locations.csv*

List of counties (lan)

| Destination Field | Source Field | Logic | Comment |
| --- | --- | --- | --- |
| location\_id | lan\_kom |  |  |
| address\_1 |  |  |  |
| address\_2 |  |  |  |
| city |  |  |  |
| state |  |  |  |
| zip |  |  |  |
| county | county\_name |  |  |
| location\_source\_value | lan\_kom |  |  |

## Table: drug\_era

The default, non-stockpiling script from the OHDSI community was used, shared by Chris Knoll on Github (<https://gist.github.com/chrisknoll/a18c8e15ff66f26fac84)>.

For this mapping, the drug eras are determined in these steps:

1. All drug exposure concepts are mapped to their respective ingredients.
2. The end of a drug exposure is set to the start date plus the number of days supply.
3. Combine the exposures of an ingredient to one era if the end and start are within 30 days of each other.

## Table: condition\_era

Standard script from OHDSI community to build condition\_era