ETL SPECIFICATION – HUS OMOP

Juha Koski

Draft 26.4.2021

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Author | Date | Content |
| 0.1 | Juha Koski | 22.12.2020 | First version |
| 0.2 | Juha Koski | 15.4.2021 | Version with HUS AI views |
| 0.21 | Juha Koski | 19.4.2021 | Updated after meeting with HUS/TAYS/TYKS |
| 0.22 | Pasi Rikala | 20.4.2021 | Revised |
| 0.3 | Juha Koski | 26.4.2021 | Updated after Pasi’s comments |

**Content**

[ETL SPECIFICATION – HUS OMOP 1-1](#_Toc70365498)

[1 General 1-4](#_Toc70365499)

[1.1 OMOP CDM 1-4](#_Toc70365500)

[1.2 Source systems 1-4](#_Toc70365501)

[1.3 Forecoming systems 1-4](#_Toc70365502)

[1.4 Tools 1-4](#_Toc70365503)

[1.5 Co-operation 1-5](#_Toc70365504)

[2 Source Data Mapping Approach to CDMV5.3.1 2-6](#_Toc70365505)

[AI TEAM VIEWS 2-6](#_Toc70365506)

[2.1 Table name: person 2-7](#_Toc70365507)

[2.1.1 Reading from asiakas 2-7](#_Toc70365508)

[2.2 Table name: death 2-9](#_Toc70365509)

[2.2.1 Reading from asiakas 2-9](#_Toc70365510)

[2.3 Table name: visit\_occurrence 2-10](#_Toc70365511)

[2.3.1 Reading from kaynti 2-10](#_Toc70365512)

[2.3.2 Reading from osastohoito 2-12](#_Toc70365513)

[2.4 Table name: observation 2-14](#_Toc70365514)

[2.4.1 Reading from diagnoosi 2-14](#_Toc70365515)

[2.5 Table name: observation\_period 2-16](#_Toc70365516)

[2.6 Table name: condition\_occurrence 2-17](#_Toc70365517)

[2.6.1 Reading from diagnoosi 2-17](#_Toc70365518)

[2.7 Table name: measurement 2-19](#_Toc70365519)

[2.7.1 Reading from diagnoosi 2-19](#_Toc70365520)

[2.7.2 Reading from multilab\_tulos 2-21](#_Toc70365521)

[2.7.3 Reading from multilab\_mikrobitulos 2-24](#_Toc70365522)

[2.7.4 Reading from multilab\_lisatiedot 2-27](#_Toc70365523)

[2.7.5 Reading from hoitotaulukko\_ods 2-30](#_Toc70365524)

[2.8 Table name: drug\_exposure 2-32](#_Toc70365525)

[2.8.1 Reading from cressida\_laakitys 2-32](#_Toc70365526)

[2.8.2 Reading from potilaan\_hoitokuurit\_ods 2-34](#_Toc70365527)

[2.8.3 Reading from potilaan\_syopalaakkeet\_ods 2-37](#_Toc70365528)

[2.9 Table name: procedure\_occurrence 2-40](#_Toc70365529)

[2.9.1 Reading from toimenpide 2-40](#_Toc70365530)

[2.9.2 Reading from diagnoosi 2-41](#_Toc70365531)

[2.10 Table name: location 2-43](#_Toc70365532)

[2.11 Table name: device\_exposure 2-44](#_Toc70365533)

[2.12 Table name: fact\_relationship 2-45](#_Toc70365534)

[2.13 Table name: note 2-46](#_Toc70365535)

[2.13.1 Reading from potilaskertomus 2-46](#_Toc70365536)

[2.14 Table name: note\_nlp 2-48](#_Toc70365537)

[2.15 Table name: specimen 2-49](#_Toc70365538)

[2.15.1 Reading from qpati\_sampledata 2-49](#_Toc70365539)

[2.16 Table name: visit\_detail 2-51](#_Toc70365540)

[2.17 Table name: cohort 2-52](#_Toc70365541)

[2.18 Table name: cohort\_attribute 2-53](#_Toc70365542)

[2.19 Table name: condition\_era 2-54](#_Toc70365543)

[2.20 Table name: dose\_era 2-55](#_Toc70365544)

[2.21 Table name: drug\_era 2-56](#_Toc70365545)

[2.22 Table name: cost 2-57](#_Toc70365546)

[2.23 Table name: payer\_plan\_period 2-58](#_Toc70365547)

[2.24 Table name: care\_site 2-59](#_Toc70365548)

[2.24.1 Reading from kaynti 2-59](#_Toc70365549)

[2.24.2 Reading from osastohoito 2-59](#_Toc70365550)

[2.25 Table name: provider 2-61](#_Toc70365551)

[2.26 Table name: cdm\_source 2-62](#_Toc70365552)

[2.27 Table name: metadata 2-63](#_Toc70365553)

[2.28 Table name: attribute\_definition 2-64](#_Toc70365554)

[2.29 Table name: cohort\_definition 2-65](#_Toc70365555)

[3 Appendix: Data Dictionary 3-66](#_Toc70365556)

[3.1 Source tables 3-66](#_Toc70365557)

[3.2 Source fields 3-67](#_Toc70365558)

# General

This document describes the ETL specification for HUS level OMOP CDM. Data includes all the patients who have been treated at the Hospital District of Helsinki and Uusimaa. The dataset is going to be continually updated. The project belongs to Treatment of acute leukemia work package which is the part of the CleverHealth Network’s eCare for Me project.

## OMOP CDM

The target version in OMOP Common Data Model (CDM) is 5.3.1. The OMOP database is Azure Synapse Analytics in the Helsinki University Hospital’s Microsoft Azure Environment.

## Source systems

Used source systems at HUS OMOP:

* Uranus (patient data system 🡪 use has ended 10/2020)
* Multilab (laboratory system)
* QPati (pathology system) 🡪 On hold

The source systems itself are not directly used in the mapping. All the source systems data is first loaded into the HUS Datalake. On the top of raw data layer in the HUS Data lake the HUS AI team constructs views for further use (for example information services, HUS data warehouse and HUS OMOP CDM). These HUS AI views are used to create ”the one truth” the data needers can rely on. Source tables from HUS Data lake are used only if there is no HUS AI views for the dataset.

## Forecoming systems

* Apotti (new patient data system)
* Genome warehouse

## Tools

* White Rabbit
* Rabbit in a Hat (v.0.10.2)

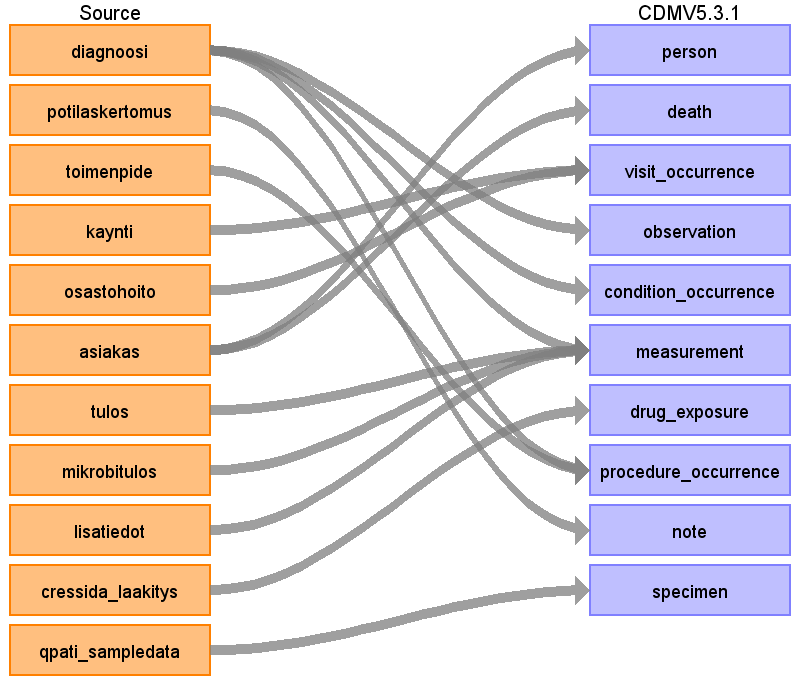
The data profiling with White Rabbit -tool is executed directly from HUS Data Lake and provided to the use of the project. ETL specification is generated with Rabbit in a Hat -tool in cooperation with medical and AI experts of HUS as stated in OHDSI best practices.

## Co-operation

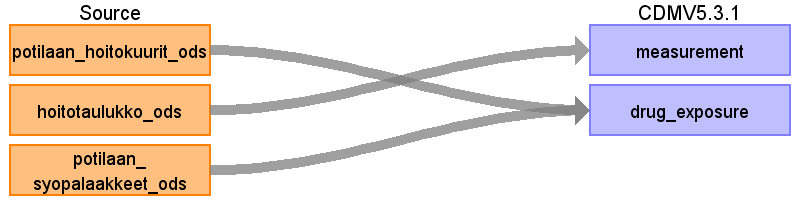
In Finland, several university hospitals (for example Tampere University Hospital and Turku University Hospital) are also planning or having on going OMOP projects. We are co-operating with other University hospitals and national organizations (like THL, Finngen and FIMM) in Finland to ensure we are having a solid base in Finnish OMOP CDM mappings. We are pursuing to create nationally coherent outcome.

# Source Data Mapping Approach to CDMV5.3.1

## AI TEAM VIEWS

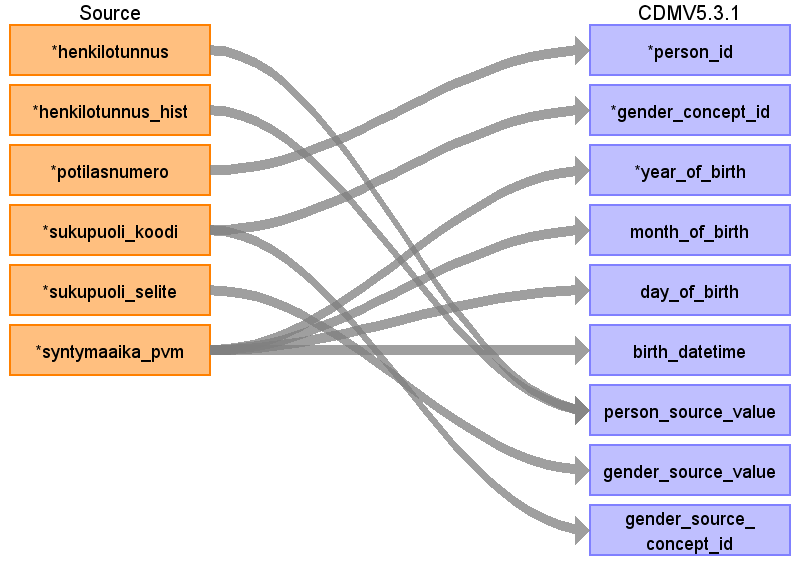


CRESSIDA\_ODS



## Table name: person

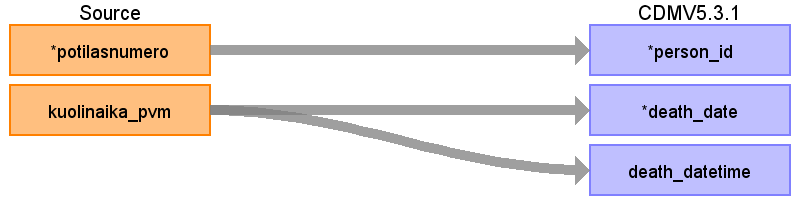
### Reading from asiakas



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id | Potilasnumero |  | Generic ID, Retrieved from lookup table |
| gender\_concept\_id | sukupuoli\_koodi |  | 1 --> 8507 Male  2 --> 8532 Female  For others use 0 |
| year\_of\_birth | syntymaaika\_pvm |  | Extracted year from time of birth |
| month\_of\_birth | syntymaaika\_pvm |  | Extracted month from time of birth |
| day\_of\_birth | syntymaaika\_pvm |  | Extracted day from time of birth |
| birth\_datetime | syntymaaika\_pvm |  |  |
| race\_concept\_id |  |  | Not available in source data. Use 0. |
| ethnicity\_concept\_id |  |  | Not available in source data. Use 0. |
| location\_id |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| care\_site\_id |  |  | Not available in source data. Use NULL. |
| person\_source\_value | henkilotunnus |  | Changes in social security numbers has to be taken into account. AI-team view from Uranus data always gives the newest patient identification. With Uranus no longer problem because it is not updated anymore. Must rethink when Apotti data is available. |
| gender\_source\_value | sukupuoli\_selite |  |  |
| gender\_source\_concept\_id |  |  | NULL |
| race\_source\_value |  |  | Not available in source data. Use NULL. |
| race\_source\_concept\_id |  |  | Not available in source data. Use NULL. |
| ethnicity\_source\_value |  |  | Not available in source data. Use NULL. |
| ethnicity\_source\_concept\_id |  |  | Not available in source data. Use NULL. |

## Table name: death

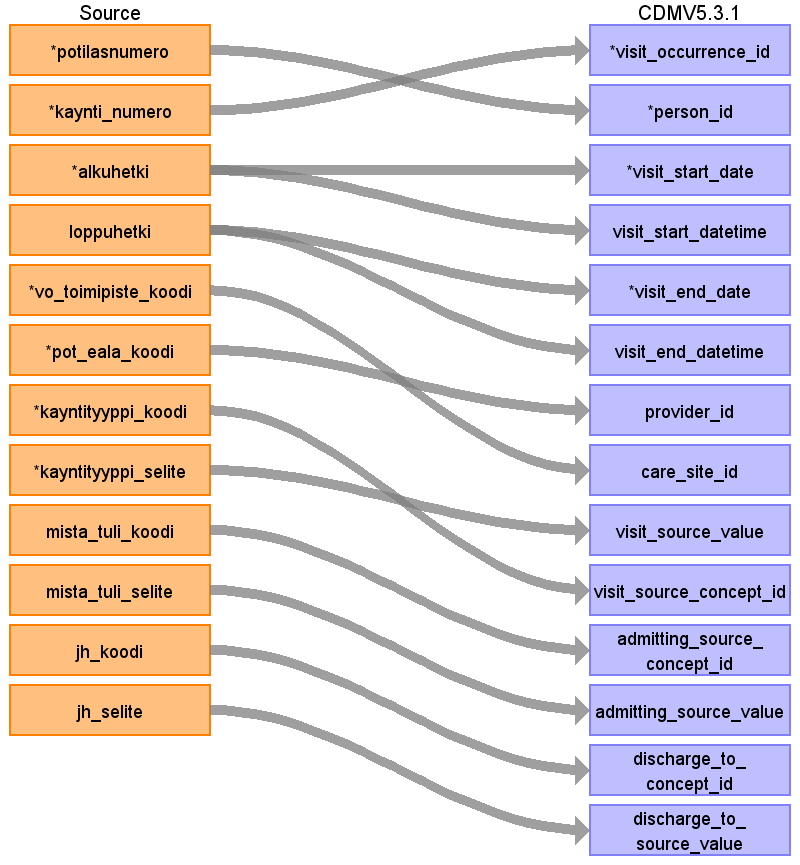
### Reading from asiakas



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| person\_id | Potilasnumero |  | Generic ID, Retrieved from lookup table |
| death\_date | kuolinaika\_pvm |  | Remember: data may be delayed for 1 month because the time of death is updated to the source system only once per month. With Uranus no longer problem because it is not updated anymore. Must rethink when Apotti data is available.  Another possibility is to use discharge source value (jh\_koodi = 3) which means death during the inpatient visit |
| death\_datetime | kuolinaika\_pvm |  |  |
| death\_type\_concept\_id |  |  | Use standard: 32817 EHR |
| cause\_concept\_id |  |  | Not available in source data. Use 0. |
| cause\_source\_value |  |  | Not available in source data. Use NULL. |
| cause\_source\_concept\_id |  |  | Not available in source data. Use NULL. |

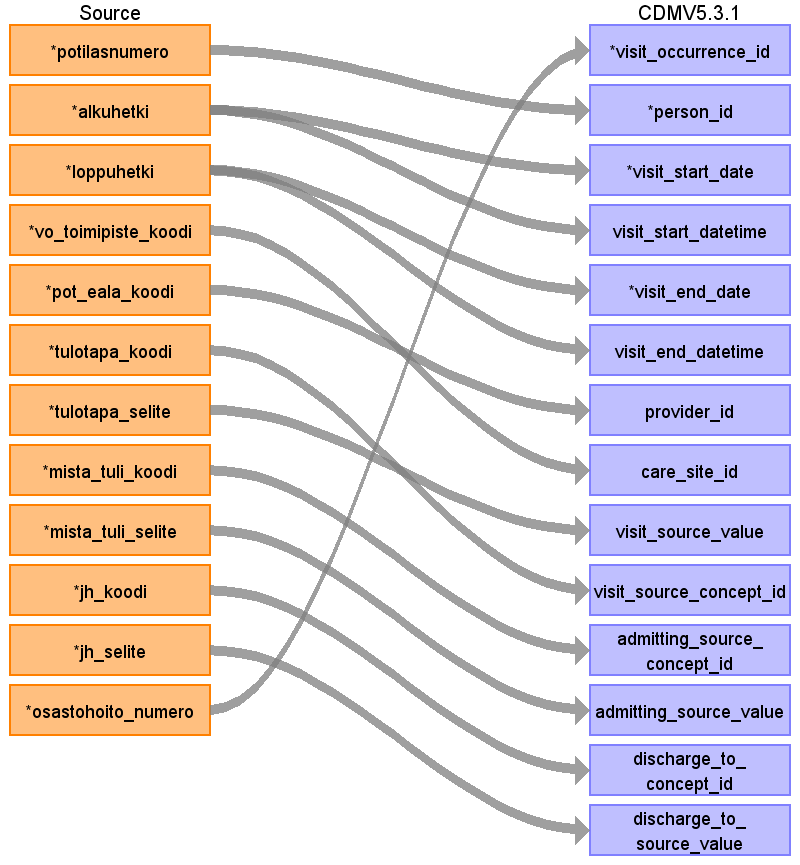
## Table name: visit\_occurrence

### Reading from kaynti



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| visit\_occurrence\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| visit\_concept\_id |  |  | 9202 Outpatient visit for Kaynti |
| visit\_start\_date | alkuhetki |  |  |
| visit\_start\_datetime | alkuhetki |  |  |
| visit\_end\_date | loppuhetki |  |  |
| visit\_end\_datetime | loppuhetki |  |  |
| visit\_type\_concept\_id |  |  | Use standard: 32817 EHR |
| provider\_id | pot\_eala\_koodi |  | We don't want to use doctor information. Use specialty instead. Use lookup to get the provider\_id that matches pot\_eala\_koodi. Not yet implemented. |
| care\_site\_id | vo\_toimipiste\_koodi |  | Use lookup to get the care\_site\_id that matches vo\_toimipiste\_koodi. |
| visit\_source\_value | kayntityyppi\_selite |  |  |
| visit\_source\_concept\_id | ~~kayntityyppi\_koodi~~ |  | NULL, Not OMOP supported concepts available? |
| admitting\_source\_concept\_id | ~~mista\_tuli\_koodi~~ |  | NULL, Not OMOP supported concepts available? |
| admitting\_source\_value | mista\_tuli\_selite |  |  |
| discharge\_to\_concept\_id | jh\_koodi |  | Not yet mapped |
| discharge\_to\_source\_value | jh\_selite |  |  |
| preceding\_visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |

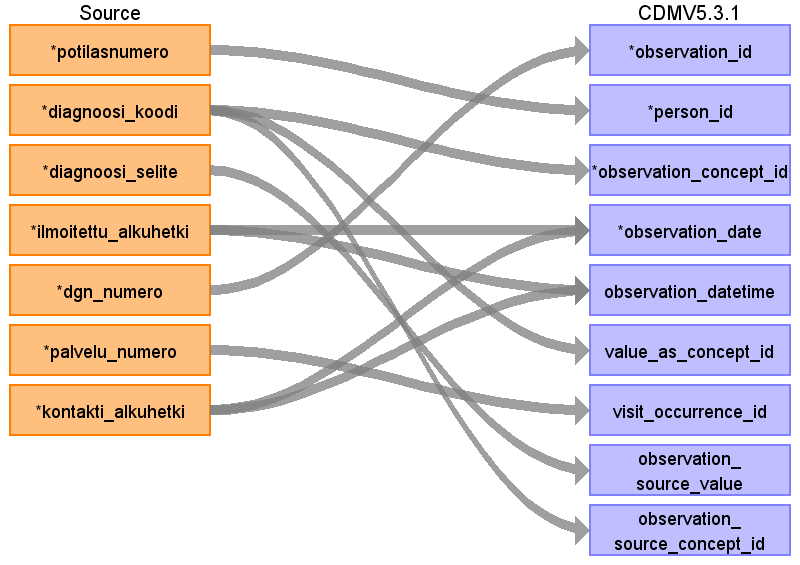
### Reading from osastohoito



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| visit\_occurrence\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| visit\_concept\_id |  |  | 9201 Inpatient visit for Osastohoito |
| visit\_start\_date | alkuhetki |  |  |
| visit\_start\_datetime | alkuhetki |  |  |
| visit\_end\_date | loppuhetki |  |  |
| visit\_end\_datetime | loppuhetki |  |  |
| visit\_type\_concept\_id |  |  | Use standard: 32817 EHR |
| provider\_id | pot\_eala\_koodi |  | We don't want to use doctor information. Use specialty instead. Use lookup to get the provider\_id that matches pot\_eala\_koodi. Not yet implemented. |
| care\_site\_id | vo\_toimipiste\_koodi |  | Use lookup to get the care\_site\_id that matches vo\_toimipiste\_koodi. |
| visit\_source\_value | tulotapa\_selite |  |  |
| visit\_source\_concept\_id | tulotapa\_koodi |  | NULL, Not OMOP supported concepts available? |
| admitting\_source\_concept\_id | mista\_tuli\_koodi |  | NULL, Not OMOP supported concepts available? |
| admitting\_source\_value | mista\_tuli\_selite |  |  |
| discharge\_to\_concept\_id | jh\_koodi |  | Not yet mapped |
| discharge\_to\_source\_value | jh\_selite |  |  |
| preceding\_visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |

## Table name: observation

### Reading from diagnoosi



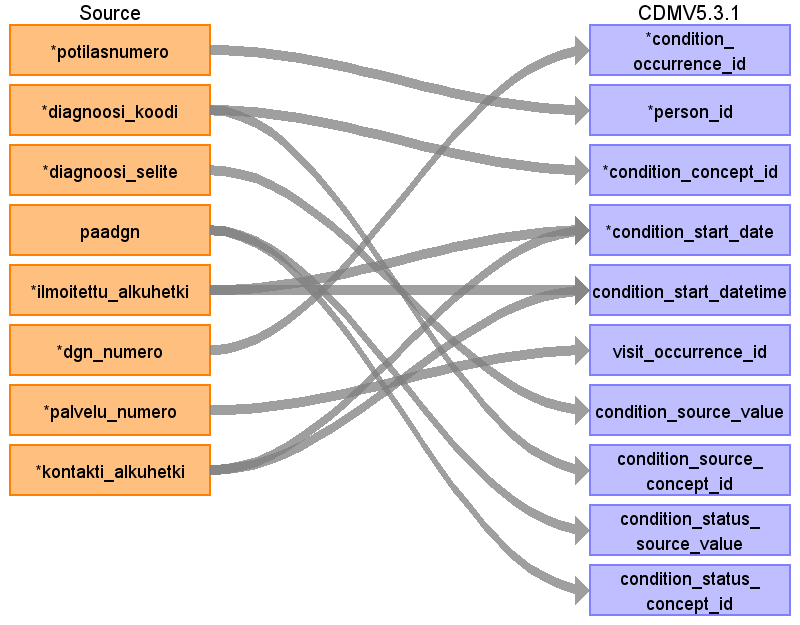
|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| observation\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| observation\_concept\_id | diagnoosi\_koodi |  |  |
| observation\_date | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  | Take the oldest datetime from connected fields |
| observation\_datetime | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  | Take the oldest datetime from connected fields |
| observation\_type\_concept\_id |  |  | 32817 EHR |
| value\_as\_number |  |  | Not available in source data. Use NULL. |
| value\_as\_string |  |  | Not available in source data. Use NULL. |
| value\_as\_concept\_id | diagnoosi\_koodi |  | If code is a concept in ICD10 vocabulary, look up using relationship\_id = 'Maps to value' . |
| qualifier\_concept\_id |  |  | Not available in source data. Use NULL. |
| unit\_concept\_id |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id | palvelu\_numero |  | Look up using visit\_occurrence data. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| observation\_source\_value | diagnoosi\_selite |  | 50 char: dgn\_vaihe + '/' + cast(paadgn as varchar) + '/' + laji + '/' + dg\_koodi |
| observation\_source\_concept\_id | diagnoosi\_koodi |  |  |
| unit\_source\_value |  |  | Not available in source data. Use NULL. |
| qualifier\_source\_value |  |  | Not available in source data. Use NULL. |

## Table name: observation\_period

This is not mapped but is implemented. From all the tables used in OMOP is taken the patients earliest occurrence date (observation\_period\_start\_date) and latest occurrence date (observation\_period\_end\_date).

## Table name: condition\_occurrence

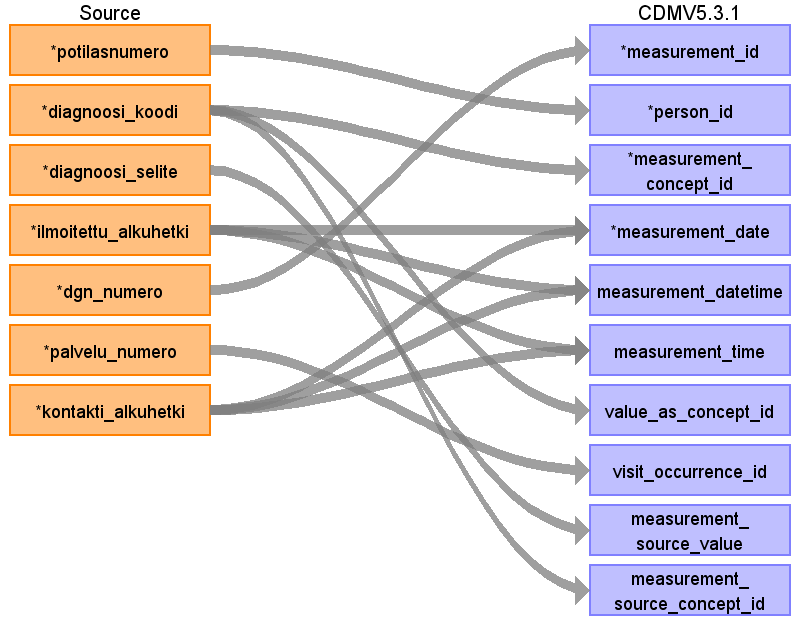
### Reading from diagnoosi



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| condition\_occurrence\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| condition\_concept\_id | diagnoosi\_koodi |  | Uses 3 mappings in order: HUS, TAYS, ICD10 |
| condition\_start\_date | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  | Take the oldest datetime from connected fields |
| condition\_start\_datetime | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  | Take the oldest datetime from connected fields |
| condition\_end\_date |  |  | Not reliable. Not used. |
| condition\_end\_datetime |  |  | Not reliable. Not used. |
| condition\_type\_concept\_id |  |  | Use 32817 HER. AI team view automatically includes only actualized (“toteutuneet”) diagnoses. |
| stop\_reason |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Specialty.  Not available in source data. Use NULL. |
| visit\_occurrence\_id | palvelu\_numero |  | Look up using visit occurrence data. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| condition\_source\_value | diagnoosi\_selite |  |  |
| condition\_source\_concept\_id | diagnoosi\_koodi |  |  |
| condition\_status\_source\_value | paadgn |  |  |
| condition\_status\_concept\_id | paadgn |  | 32902 Primary diagnosis  32908 Secondary diagnosis  Long term diagnoses are planned to be taken to the AI team view. Do we want to take them into the OMOP? |

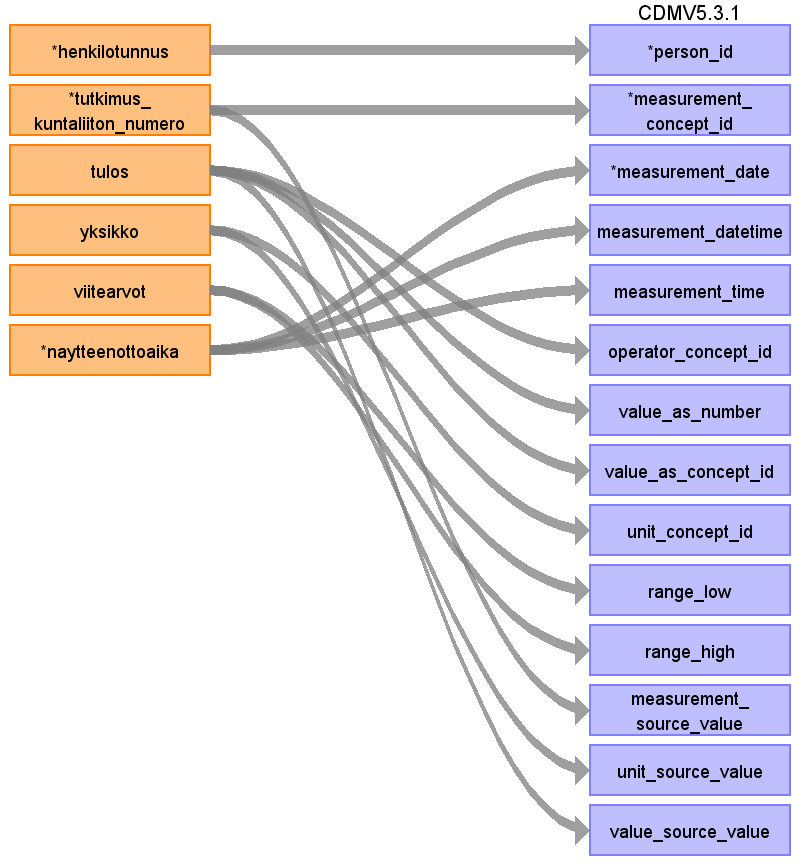
## Table name: measurement

### Reading from diagnoosi



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| measurement\_concept\_id | diagnoosi\_koodi |  | ICD10 vocabulary |
| measurement\_date | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  | Take the oldest datetime from connected fields |
| measurement\_datetime | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  |  |
| measurement\_time | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  | Take the oldest datetime from previous fields |
| measurement\_type\_concept\_id |  |  | Cressida: 32817 EHR |
| operator\_concept\_id |  |  | Not available in source data. Use NULL. |
| value\_as\_number |  |  | Not available in source data. Use NULL. |
| value\_as\_concept\_id | diagnoosi\_koodi |  | ”Maps to value” as boundary condition for relationship view |
| unit\_concept\_id |  |  | Not available in source data. Use NULL. |
| range\_low |  |  | Not available in source data. Use NULL. |
| range\_high |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id | palvelu\_numero |  |  |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| measurement\_source\_value | diagnoosi\_selite |  |  |
| measurement\_source\_concept\_id | diagnoosi\_koodi |  |  |
| unit\_source\_value |  |  | Not available in source data. Use NULL. |
| value\_source\_value |  |  | Not available in source data. Use NULL. |

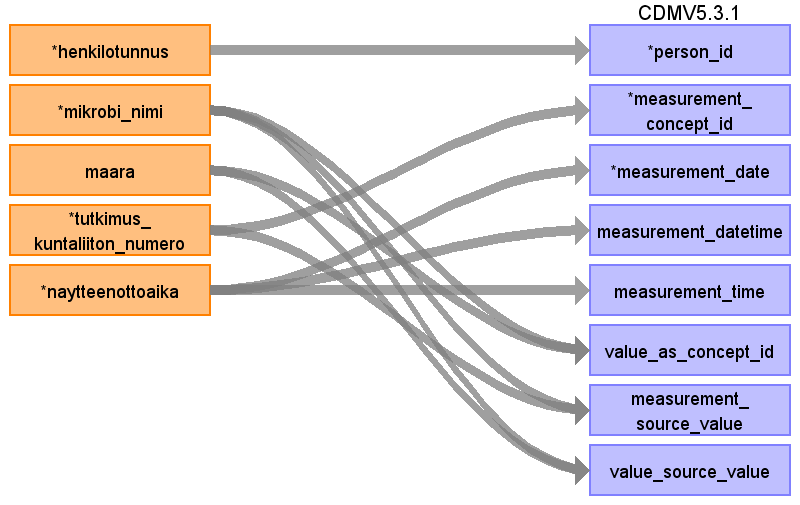
### Reading from multilab\_tulos



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Generic ID |
| person\_id | henkilotunnus |  | Retrieved from lookup table |
| measurement\_concept\_id | tutkimus\_kuntaliiton\_numero |  | LOINC |
| measurement\_date | naytteenottoaika |  | Multilab: Näytteenottoaika exact enough |
| measurement\_datetime | naytteenottoaika |  |  |
| measurement\_time | naytteenottoaika |  |  |
| measurement\_type\_concept\_id |  |  | Multilab: 32856 lab |
| operator\_concept\_id | tulos |  | <, >, <=, >=  If available |
| value\_as\_number | tulos |  | Only numerical information, otherwise null  Multilab: has to be determined from the text field |
| value\_as\_concept\_id | tulos |  | esim. +, ++, +++ merkitykselle pitäisi löytää konsepti. Koskee myös muita ei-numeerisia arvoja, jotka halutaan saada tuloksina kiinni  Tehty, NEGAT, POSIT, esim.positiiviselle on 3 erilaista standardikonseptia  EI TEHDA, EI TEHTY, EI ONNISTU tyyppiset jätetään pois  Not mapped yet |
| unit\_concept\_id | yksikko |  |  |
| range\_low | viitearvot |  | Parse with regular expression |
| range\_high | viitearvot |  | Parse with regular expression |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| measurement\_source\_value | tutkimus\_kuntaliiton\_numero |  | Kuntaliiton numero |
| measurement\_source\_concept\_id |  |  | Not available in source data. Use NULL. |
| unit\_source\_value | yksikko |  |  |
| value\_source\_value | tulos |  |  |

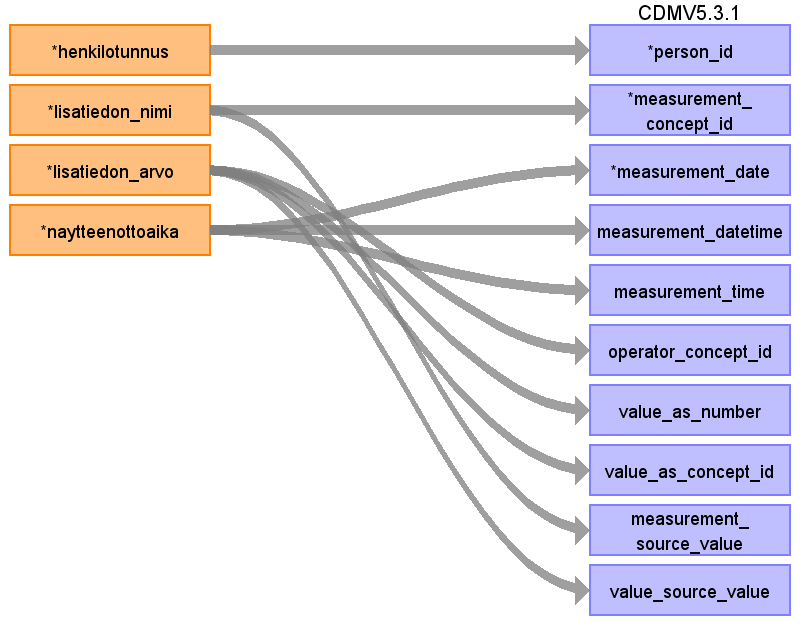
### Reading from multilab\_mikrobitulos

Not yet implemented



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Generic ID |
| person\_id | henkilotunnus |  | Retrieved from lookup table |
| measurement\_concept\_id | tutkimus\_kuntaliiton\_numero |  | Not yet mapped. Definition is not ready yet!  Mikrobitulokset: Kuntaliiton numero ratkaisee konseptin, tulos on löytyy/ei löydy |
| measurement\_date | naytteenottoaika |  | Multilab: Näytteenottoaika exact enough |
| measurement\_datetime | naytteenottoaika |  |  |
| measurement\_time | naytteenottoaika |  |  |
| measurement\_type\_concept\_id |  |  | Multilab: 32856 Lab |
| operator\_concept\_id |  |  | <, >, <=, >=  If available |
| value\_as\_number |  |  | Only numerical information, otherwise null  Multilab: has to be determined from the text field |
| value\_as\_concept\_id | maara  mikrobi\_nimi |  | Not yet mapped  esim. +, ++, +++ merkitykselle pitäisi löytää konsepti. Koskee myös muita ei-numeerisia arvoja, jotka halutaan saada tuloksina kiinni  Tehty, NEGAT, POSIT, esim positiiviselle on 3 erilaista standardikonseptia  EI TEHDA, EI TEHTY, EI ONNISTU tyyppiset jätetään pois |
| unit\_concept\_id |  |  | Not available in source data. Use NULL. |
| range\_low |  |  | Not available in source data. Use NULL. |
| range\_high |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| measurement\_source\_value | tutkimus\_kuntaliiton\_numero  mikrobi\_nimi |  | Huslab number |
| measurement\_source\_concept\_id |  |  |  |
| unit\_source\_value |  |  | Not available in source data. Use NULL. |
| value\_source\_value | mikrobi\_nimi  maara |  |  |

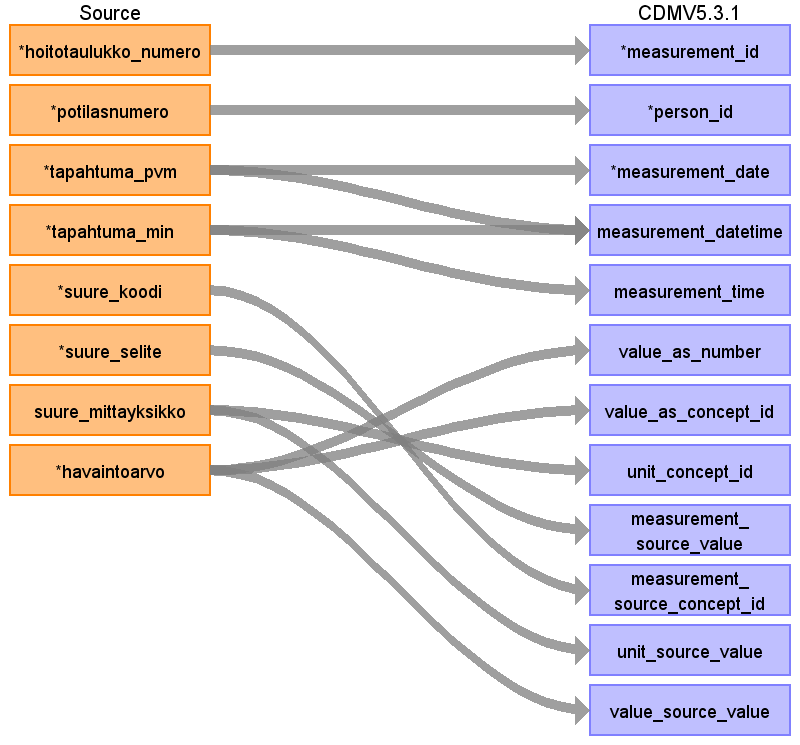
### Reading from multilab\_lisatiedot



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Generic ID |
| person\_id | henkilotunnus |  | Retrieved from lookup table |
| measurement\_concept\_id | lisatiedon\_nimi |  | \*TARKISTETTAVA tuleeko lisäksi kuntaliiton\_numero |
| measurement\_date | naytteenottoaika |  | Multilab: Näytteenottoaika exact enough |
| measurement\_datetime | naytteenottoaika |  |  |
| measurement\_time | naytteenottoaika |  |  |
| measurement\_type\_concept\_id |  |  | Multilab: 32856 Lab |
| operator\_concept\_id | lisatiedon\_arvo |  | <, >, <=, >=  If available |
| value\_as\_number | lisatiedon\_arvo |  | Only numerical information, otherwise null |
| value\_as\_concept\_id | lisatiedon\_arvo |  | Not yet mapped  esim. +, ++, +++ merkitykselle pitäisi löytää konsepti. Koskee myös muita ei-numeerisia arvoja, jotka halutaan saada tuloksina kiinni  Tehty, NEGAT, POSIT, esim positiiviselle on 3 erilaista standardikonseptia  EI TEHDA, EI TEHTY, EI ONNISTU tyyppiset jätetään pois |
| unit\_concept\_id |  |  | Not available in source data. Use NULL. |
| range\_low |  |  | Not available in source data. Use NULL. |
| range\_high |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| measurement\_source\_value | lisatiedon\_nimi |  |  |
| measurement\_source\_concept\_id |  |  | Not available in source data. Use NULL. |
| unit\_source\_value |  |  | Not available in source data. Use NULL. |
| value\_source\_value | lisatiedon\_arvo |  |  |

### Reading from hoitotaulukko\_ods

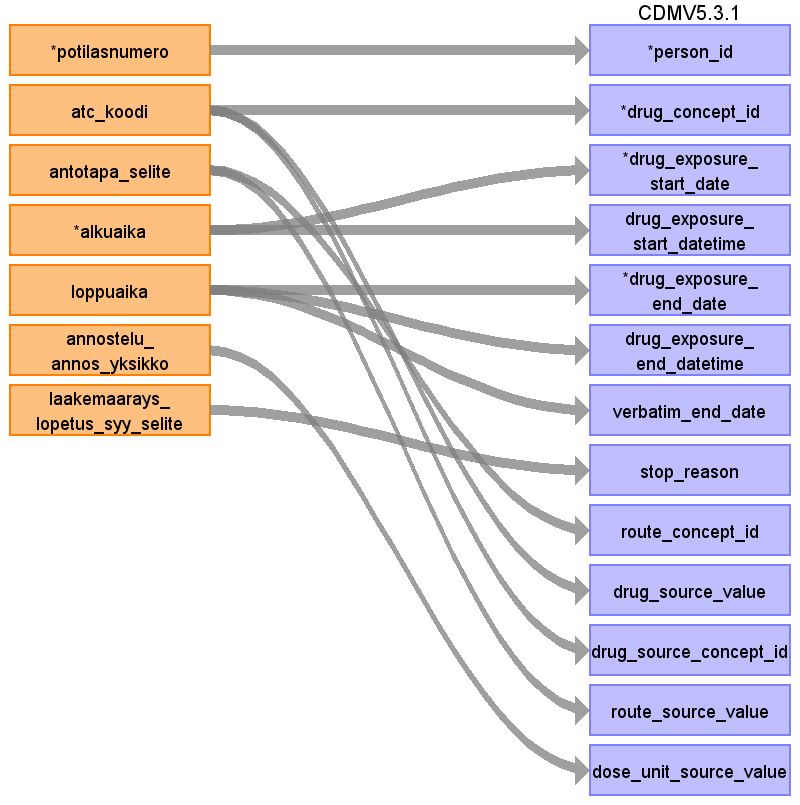
Not yet implemented



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| measurement\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| measurement\_concept\_id |  |  | Hoitotaulukossa suoraan LOINC-koodina saatavilla? |
| measurement\_date | tapahtuma\_pvm |  |  |
| measurement\_datetime | tapahtuma\_pvm  tapahtuma\_min |  |  |
| measurement\_time | tapahtuma\_min |  |  |
| measurement\_type\_concept\_id |  |  | voidaanko olla varmoja, että kaikki hoitotaulukon tiedot ovat näitä? Hoitotaulukossa paljon erilaisia arvoja. Mistä voidaan tietää onko potilaan raportoima, labratulos tms? |
| operator\_concept\_id |  |  |  |
| value\_as\_number | havaintoarvo |  |  |
| value\_as\_concept\_id | havaintoarvo |  |  |
| unit\_concept\_id | suure\_mittayksikko |  |  |
| range\_low |  |  | Not available in source data. Use NULL. |
| range\_high |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Hoitotaulukosta ei suoraa linkkiä käyntiin tai hoitojaksoon. Haettava potilaan, päivämäärän ja toimipisteen perusteella. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| measurement\_source\_value | suure\_selite |  |  |
| measurement\_source\_concept\_id | suure\_koodi |  |  |
| unit\_source\_value | suure\_mittayksikko |  |  |
| value\_source\_value | havaintoarvo |  |  |

## Table name: drug\_exposure

### Reading from cressida\_laakitys



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| drug\_exposure\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| drug\_concept\_id | atc\_koodi |  |  |
| drug\_exposure\_start\_date | alkuaika |  |  |
| drug\_exposure\_start\_datetime | alkuaika |  |  |
| drug\_exposure\_end\_date | loppuaika |  |  |
| drug\_exposure\_end\_datetime | loppuaika |  |  |
| verbatim\_end\_date | loppuaika |  |  |
| drug\_type\_concept\_id |  |  | 32830 EHR medication list |
| stop\_reason | laakemaarays\_lopetus\_syy\_selite |  |  |
| refills |  |  | Not available in source data. Use NULL. |
| quantity |  |  | Not available from Uranus, later from Apotti? |
| days\_supply |  |  | Not available in source data. Use NULL. |
| sig |  |  | Not available in source data. Use NULL. |
| route\_concept\_id | antotapa\_selite |  |  |
| lot\_number |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| drug\_source\_value | atc\_koodi |  |  |
| drug\_source\_concept\_id | atc\_koodi |  |  |
| route\_source\_value | antotapa\_selite |  |  |
| dose\_unit\_source\_value | annostelu\_annos\_yksikko |  |  |

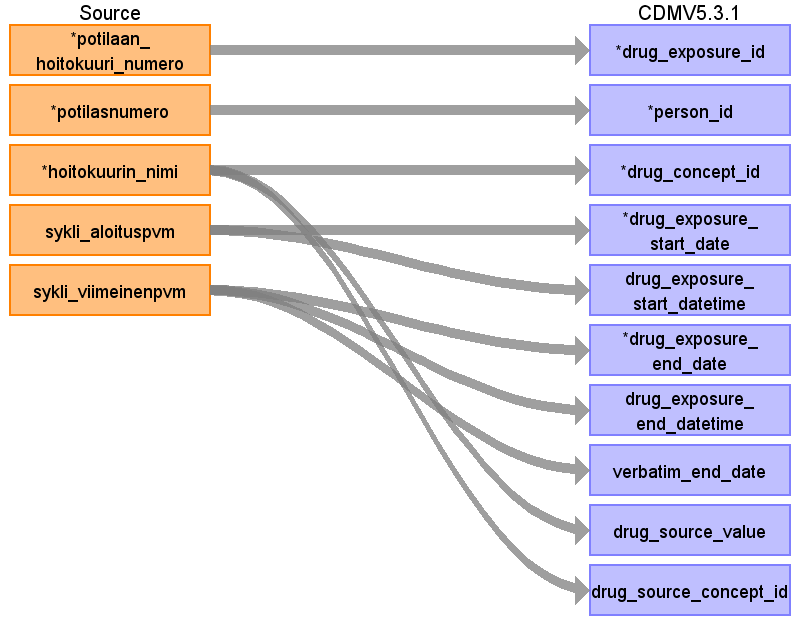
### Reading from potilaan\_hoitokuurit\_ods

Kun tuodaan erikseen hoitokuuri ja lääkkeen annot, niin tulee sekä hoitokuurista rivejä että lääkkeenantorivejä eli kahden rajapinnan joinaus, mistä tiedot saadaan

Hoitokuurien osalta pohdittava otetaanko mukaan nyt vai myöhemmin. Versiossa 6.1 tulossa onkologian laajennus, jossa mukana EPISODE + EPISODE\_EVENT -taulut tähän käyttötarkoitukseen. Pohdittava tehdäänkö 5.3.1:een väliaikainen ratkaisu vai odotetaanko versiota 6.1 ennen kuin hoitokuurit otetaan matkaan. EPISODE -taulun myötä esim. syöpien osalta voidaan myös hallita taudin vaiheita paikallinen/levinnyt jne. kun diagnoosi ei muutu levinneisyyden mukaan vaan on aina löytöpaikan mukaan.

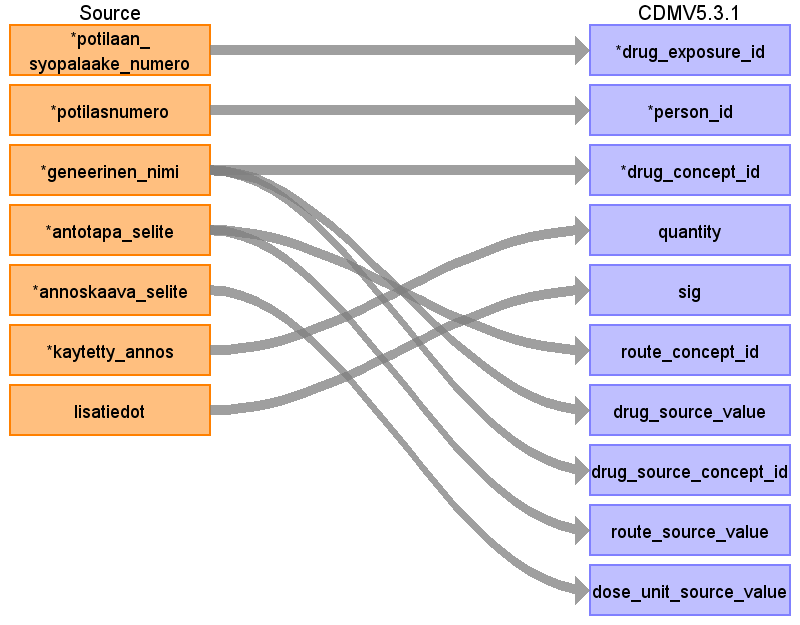
Syöpälääkkeet ja antokirjaukset AI-tiimin näkymältä eli tiedot yhdistetty keskenään. Syöpälääkkeen numero tunnisteeksi, aikaleimat tuodaan antokirjauksilta

Hoitokuurilta otetaan syöpälääketietojen yhteydessä tietoja kuten päivä, potilas ja syklin alku- ja loppuaika



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| drug\_exposure\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| drug\_concept\_id |  |  | Retrieved from potilaan\_syopalaakkeet\_ods |
| drug\_exposure\_start\_date | sykli\_aloituspvm |  |  |
| drug\_exposure\_start\_datetime | sykli\_aloituspvm |  |  |
| drug\_exposure\_end\_date | sykli\_viimeinenpvm |  | Saadaanko hoitokuurin kesto eli hoitokuurin syklin loppupäivä? Tutkitaan dataa. |
| drug\_exposure\_end\_datetime | sykli\_viimeinenpvm |  |  |
| verbatim\_end\_date | sykli\_viimeinenpvm |  |  |
| drug\_type\_concept\_id |  |  | 32817 EHR |
| stop\_reason |  |  | Not available in source data. Use NULL. |
| refills |  |  | Not available in source data. Use NULL. |
| quantity |  |  | Not available in source data. Use NULL. |
| days\_supply |  |  | Not available in source data. Use NULL. |
| sig |  |  | Not available in source data. Use NULL. |
| route\_concept\_id |  |  | Not available in source data. Use NULL. |
| lot\_number |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| drug\_source\_value |  |  | geneerinen\_nimi retrieved from potilaan\_syopalaakkeet\_ods |
| drug\_source\_concept\_id |  |  | Not available in source data. Use NULL. |
| route\_source\_value |  |  | Not available in source data. Use NULL. |
| dose\_unit\_source\_value |  |  | Not available in source data. Use NULL. |

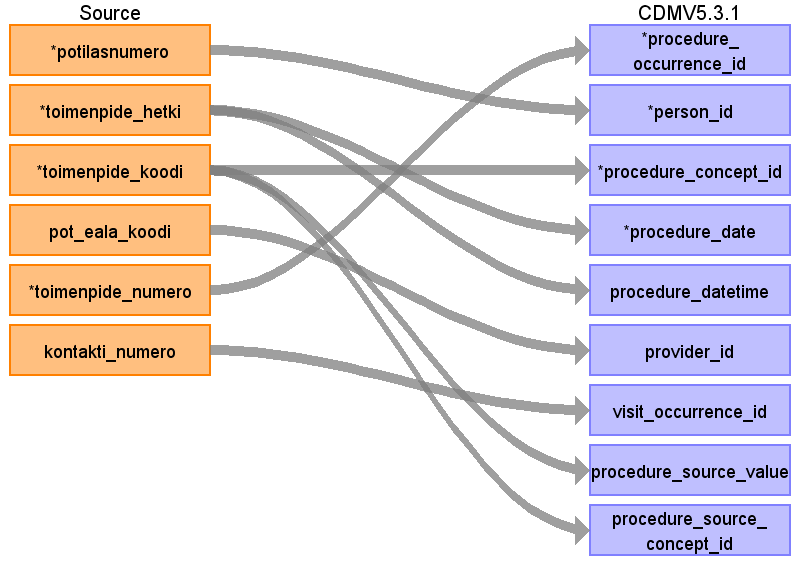
### Reading from potilaan\_syopalaakkeet\_ods



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| drug\_exposure\_id | potilaan\_syopalaake\_numero |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| drug\_concept\_id | geneerinen\_nimi |  |  |
| drug\_exposure\_start\_date |  |  | Retrieved from potilaan\_hoitokuurit\_ods |
| drug\_exposure\_start\_datetime |  |  | Retrieved from potilaan\_hoitokuurit\_ods |
| drug\_exposure\_end\_date |  |  | Retrieved from potilaan\_hoitokuurit\_ods |
| drug\_exposure\_end\_datetime |  |  | Retrieved from potilaan\_hoitokuurit\_ods |
| verbatim\_end\_date |  |  | Retrieved from potilaan\_hoitokuurit\_ods |
| drug\_type\_concept\_id |  |  | Retrieved from potilaan\_hoitokuurit\_ods |
| stop\_reason |  |  | Not available in source data. Use NULL. |
| refills |  |  | Not available in source data. Use NULL. |
| quantity |  |  | Not available in source data. Use NULL. |
| days\_supply |  |  | Not available in source data. Use NULL. |
| sig | lisatiedot |  | Tekstimuotoinen ohje lääkkeen käytöstä mikä on resepteihin kirjoitettu (kuinka monta tablettia, monta kertaa päivässä, kuinka pitkään jne)  Pohdittava syöpälääkkeiden osalta mitä tuodaan vai tuodaanko mitään |
| route\_concept\_id | antotapa\_selite |  | For example oral:  suu = 4132161  Huom! ks. kirjoitusasu ja mitä kaikkia vaihtoehtoja taulussa oikeasti on |
| lot\_number |  |  | Not available in source data. Use NULL. |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id |  |  | Not available in source data. Use NULL. |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| drug\_source\_value | geneerinen\_nimi |  | lääkeainetermi RX-norm  Suomessa myynnissä oleva kauppanimi, vaikuttavan aineen mukaan RX-normista lopullinen concept  Vaikuttavan aineen merkitys suurempi kuin kauppanimen |
| drug\_source\_concept\_id | ~~geneerinen\_nimi~~ |  | NULL, Not OMOP supported concepts available? |
| route\_source\_value | antotapa\_selite |  |  |
| dose\_unit\_source\_value | annoskaava\_selite |  |  |

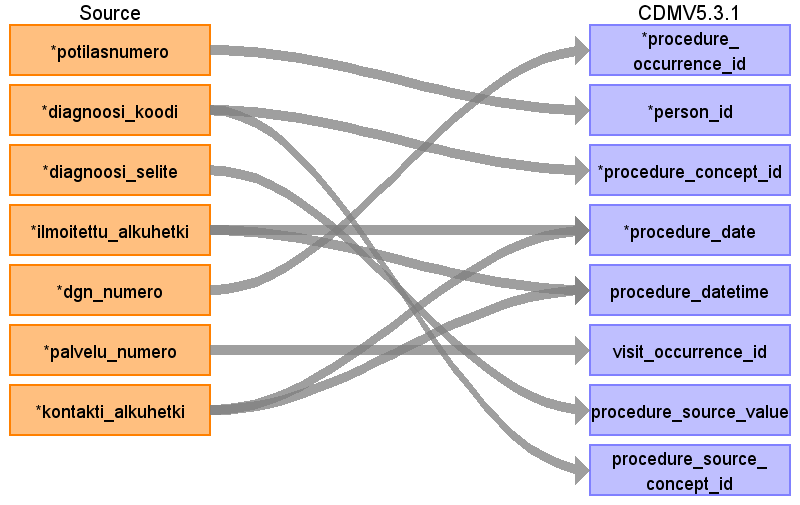
## Table name: procedure\_occurrence

### Reading from toimenpide



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| procedure\_occurrence\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| procedure\_concept\_id | toimenpide\_koodi |  | TAYS mapping at use |
| procedure\_date | toimenpide\_hetki |  |  |
| procedure\_datetime | toimenpide\_hetki |  |  |
| procedure\_type\_concept\_id |  |  | 32817 EHR |
| modifier\_concept\_id |  |  | Not available in source data. Use NULL. |
| quantity |  |  | use quantity 1 as constant value |
| provider\_id | pot\_eala\_koodi |  |  |
| visit\_occurrence\_id | kontakti\_numero |  |  |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| procedure\_source\_value | toimenpide\_koodi |  |  |
| procedure\_source\_concept\_id | ~~toimenpide\_koodi~~ |  | NULL, Not OMOP supported concepts available? |
| modifier\_source\_value |  |  | Not available in source data. Use NULL. |

### Reading from diagnoosi



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| procedure\_occurrence\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| procedure\_concept\_id | diagnoosi\_koodi |  | ICD10 vocabulary |
| procedure\_date | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  |  |
| procedure\_datetime | ilmoitettu\_alkuhetki  kontakti\_alkuhetki |  |  |
| procedure\_type\_concept\_id |  |  | 32817 EHR |
| modifier\_concept\_id |  |  | Not available in source data. Use NULL. |
| quantity |  |  | use quantity 1 as constant value |
| provider\_id |  |  | Not available in source data. Use NULL. |
| visit\_occurrence\_id | palvelu\_numero |  |  |
| visit\_detail\_id |  |  | Not available in source data. Use NULL. |
| procedure\_source\_value | diagnoosi\_selite |  |  |
| procedure\_source\_concept\_id | diagnoosi\_koodi |  |  |
| modifier\_source\_value |  |  | Not available in source data. Use NULL. |

## Table name: location

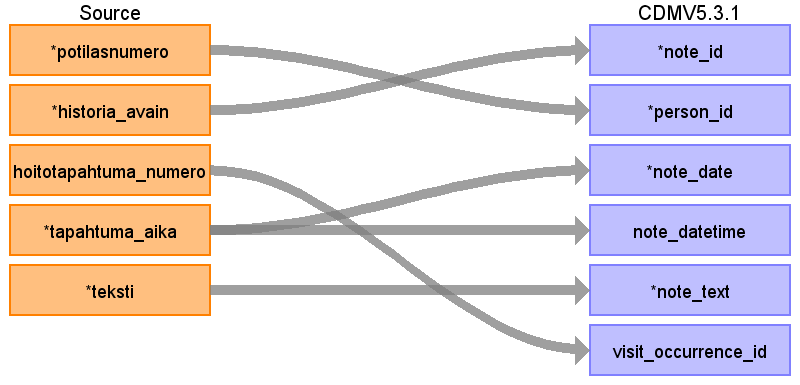
## Table name: device\_exposure

## Table name: fact\_relationship

## Table name: note

Includes sensitive data that cannot be transrerred. To be removed from ETL? Will not be implemented.

### Reading from potilaskertomus



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| note\_id |  |  | Generic ID |
| person\_id | potilasnumero |  | Retrieved from lookup table |
| note\_date | tapahtuma\_aika |  |  |
| note\_datetime | tapahtuma\_aika |  |  |
| note\_type\_concept\_id |  |  | Use 32817 EHR |
| note\_class\_concept\_id |  |  |  |
| note\_title |  |  | Not available in source data. Use NULL. |
| note\_text | teksti |  |  |
| encoding\_concept\_id |  |  |  |
| language\_concept\_id |  |  |  |
| provider\_id |  |  | Specialty. Not available in source data. Use NULL. |
| visit\_occurrence\_id | hoitotapahtuma\_numero |  |  |
| visit\_detail\_id |  |  | Not used. |
| note\_source\_value |  |  |  |

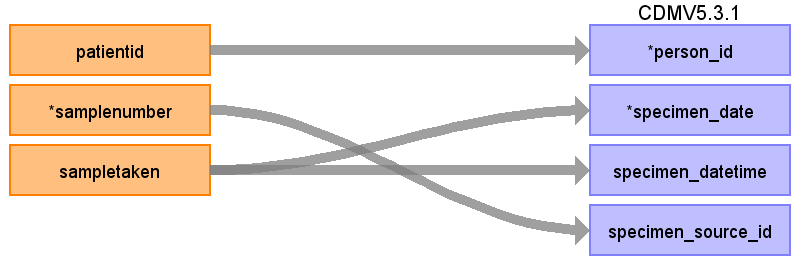
## Table name: note\_nlp

Will not be implemented.

## Table name: specimen

### Reading from qpati\_sampledata

The problem with finnish laboratory codes that cannot be mapped directly to SNOMED-CT etc. At the moment THL in Finland is going to standardize codes used in Finnish laboratories so in the future we can have data that can be mapped to SNOMED-CT. All the history data in HUS is at the moment not mapped to SNOMED-CT. Only limited part of history codes are mapped and have to be considered if it is useful for the project to continue with this at the moment.



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| specimen\_id |  |  | Generic ID |
| person\_id | patientid |  | Retrieved from lookup table |
| specimen\_concept\_id |  |  | Pitäisi saada T\_SNOMED mapattua THL:n uudistettuun SNOMEDiin ja sitä kautta löydettyä concept\_id? |
| specimen\_type\_concept\_id |  |  | Use standard: 32817 EHR |
| specimen\_date | sampletaken |  |  |
| specimen\_datetime | sampletaken |  |  |
| quantity |  |  | Not available in source data. Use NULL. |
| unit\_concept\_id |  |  | Not available in source data. Use NULL. |
| anatomic\_site\_concept\_id |  |  | T\_SNOMED?  Vaatii anatomisten paikkojen mappauksen, mikä on iso urakka |
| disease\_status\_concept\_id |  |  | Ei lähdedatassa rakenteisessa muodossa? |
| specimen\_source\_id | samplenumber |  |  |
| specimen\_source\_value |  |  | T\_SNOMED from view qpati\_diagnosis |
| unit\_source\_value |  |  | Not available in source data. Use NULL. |
| anatomic\_site\_source\_value |  |  | T\_SNOMED from view qpati\_diagnosis |
| disease\_status\_source\_value |  |  | Not available in source data. Use NULL. |

## Table name: visit\_detail

## Table name: cohort

## Table name: cohort\_attribute

## Table name: condition\_era

Populated using standard srcipt from OHDSI found here: <https://ohdsi.github.io/CommonDataModel/sqlScripts.html#condition_eras>.

## Table name: dose\_era

## Table name: drug\_era

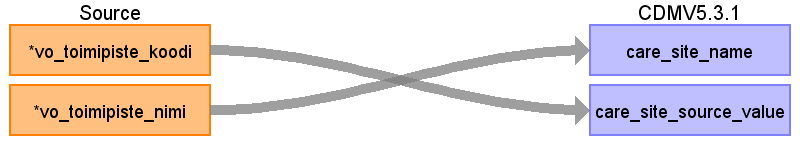
Populated using standard srcipt from OHDSI found here: <https://ohdsi.github.io/CommonDataModel/sqlScripts.html#Drug_>Eras

## Table name: cost

## Table name: payer\_plan\_period

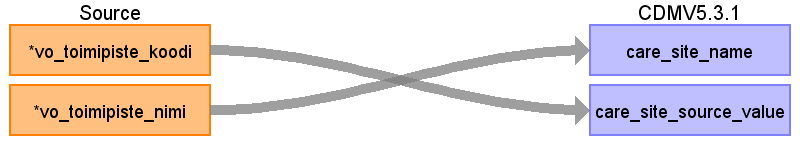
## Table name: care\_site

### Reading from kaynti



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| care\_site\_id |  |  | generic ID (lookup table is used to store generated care\_site\_id and corresponding vo\_toimipiste\_koodi) |
| care\_site\_name | vo\_toimipiste\_nimi |  |  |
| place\_of\_service\_concept\_id |  |  | Not available in source data. Use NULL. |
| location\_id |  |  | Not available in source data. Use NULL. |
| care\_site\_source\_value | vo\_toimipiste\_koodi |  |  |
| place\_of\_service\_source\_value |  |  | Not available in source data. Use NULL. |

### Reading from osastohoito



|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| care\_site\_id |  |  | generic ID (lookup table is used to store generated care\_site\_id and corresponding vo\_toimipiste\_koodi) |
| care\_site\_name | vo\_toimipiste\_nimi |  |  |
| place\_of\_service\_concept\_id |  |  | Not available in source data. Use NULL. |
| location\_id |  |  | Not available in source data. Use NULL. |
| care\_site\_source\_value | vo\_toimipiste\_koodi |  |  |
| place\_of\_service\_source\_value |  |  | Not available in source data. Use NULL. |

## Table name: provider

Not yet implemented. Mapping for specialties.

## Table name: cdm\_source

|  |  |  |  |
| --- | --- | --- | --- |
| Destination Field | Source Field | Logic | Comment |
| cdm\_source\_name |  |  | 'HUS-Tietoallas' |
| cdm\_source\_abbreviation |  |  | 'HUS-Tietoallas' |
| cdm\_holder |  |  | 'HUS' |
| source\_description |  |  | 'HUS-Tietoaltaan tiedot' |
| source\_documentation\_reference |  |  | Not yet implemented |
| cdm\_etl\_reference |  |  | 'https://ohdsi.github.io/CommonDataModel/cdm531.html#CDM\_SOURCE' |
| source\_release\_date |  |  | current timestamp |
| cdm\_release\_date |  |  | current timestamp |
| cdm\_version |  |  | '5.3.1' |
| vocabulary\_version |  |  | '20201211' |

## Table name: metadata

## Table name: attribute\_definition

## Table name: cohort\_definition

# Appendix: Data Dictionary

The tables and fields included in the OMOP mapping process are listed below.

## Source tables

|  |  |  |
| --- | --- | --- |
| **Source table** | **English name** | **Description** |
| aitteamview.asiakas | client/patient | patient gender, date of birth, date of death |
| aitteamview.cressida\_laakitys | in-hospital prescriptions and electronic prescriptions | drugs prescribed/delivered at the hospital and electronic prescriptions from HUS, coded in ATC |
| aitteamview.diagnoosi | diagnoses | ICD10 diagnoses associated with patient visit, 1 diagnosis per row. |
| aitteamview.kaynti | outpatient visit | people with health problems who visit the hospital for diagnosis or treatment, but do not at this time require a bed or to be admitted for overnight care |
| aitteamview.multilab\_lisatiedot | extra information on laboratory results |  |
| aitteamview.multilab\_mikrobitulos | microbe results |  |
| aitteamview.osastohoito | inpatient visit | Inpatient care requires overnight hospitalization. Patients must stay at the medical facility where their procedure was done for at least one night. |
| aitteamview.potilaskertomus | medical report |  |
| aitteamview.qpati\_diagnoses | pahtological diagnoses | data on pathology samples and tests |
| aitteamview.qpati\_sampledata | pathological samples | data on pathology samples and tests |
| aitteamview.toimenpide | procedures | surgical and clinical procedures, coded in Finnish procedure coding or local codes |
| aitteamview.multilab\_tulos | laboratory results | laboratory test and results |
| uranus.hoitotaulukko\_ods |  |  |
| uranus.potilaan\_hoitokuurit\_ods | anticancer drug treatments | Anticancer drug treatments, 1 cycle per row |
| uranus.potilaan\_syopalaakkeet\_ods | anticancer drugs | anticancer drug data, coded in active ingredient names (geneerinen nimi) |

## Source fields

|  |  |  |  |
| --- | --- | --- | --- |
| **Source Table** | **Field** | **Type** | **Description** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |