Example Centre : Digione Onboarding Document

Example Author

2025-06-16

# CDM Details

Table 1.1 : CDM overview

| Level | Variable | Value |
| --- | --- | --- |
| General | Snapshot date | 2025-06-16 |
| General | Person count | 1,735,134 |
| General | Vocabulary version | v5.0 09-SEP-22 |
| Observation period | N | 1,735,134 |
| Observation period | Start date | 1936-01-01 |
| Observation period | End date | 2021-12-30 |
| Cdm | Source name | full\_202112\_source\_data |
| Cdm | Version | V5.4.0 |
| Cdm | Holder name | IQVIA |
| Cdm | Release date | 2022-03-30 |
| Cdm | Description | NotAv |
| Cdm | Documentation reference | NotAv |
| Cdm | Source type | snowflake |

# Summary of clinical tables

Table 2.1 : Condition occurrence

| Variable | Response | Measurement | Value |
| --- | --- | --- | --- |
| In observation | Yes | N (%) | 31,373,274 (100.00%) |
| Number records | - | N | 31,373,274.00 |
| Number subjects | - | N (%) | 1,735,075 (100.00%) |
| Records per person | - | Mean (SD) | 18.08 (32.24) |
|  |  | max | 662.00 |
|  |  | median | 6.00 |
|  |  | min | 0.00 |
|  |  | q25 | 3.00 |
|  |  | q75 | 17.00 |
| Source vocabulary | Icd10cm | N (%) | 18,425,053 (58.73%) |
|  | Icd9cm | N (%) | 11,602,431 (36.98%) |
|  | None | N (%) | 1,345,790 (4.29%) |
| Standard concept | S | N (%) | 31,371,426 (99.99%) |
|  | - | N (%) | 1,848 (0.01%) |
| Type concept id | Ehr | N (%) | 2,360,061 (7.52%) |
|  | Ehr problem list | N (%) | 29,013,210 (92.48%) |
|  | Lab | N (%) | 3 (0.00%) |

Table 2.2 : Drug exposure

| Variable | Response | Measurement | Value |
| --- | --- | --- | --- |
| In observation | Yes | N (%) | 151,874,418 (100.00%) |
| Number records | - | N | 151,874,418.00 |
| Number subjects | - | N (%) | 1,158,017 (66.74%) |
| Records per person | - | Mean (SD) | 87.53 (355.23) |
|  |  | max | 47,501.00 |
|  |  | median | 7.00 |
|  |  | min | 0.00 |
|  |  | q25 | 0.00 |
|  |  | q75 | 55.00 |
| Source vocabulary | None | N (%) | 151,874,418 (100.00%) |
| Standard concept | S | N (%) | 141,354,691 (93.07%) |
|  | - | N (%) | 10,519,727 (6.93%) |
| Type concept id | Ehr administration record | N (%) | 108,180,130 (71.23%) |
|  | Ehr prescription issue record | N (%) | 43,694,288 (28.77%) |

Table 2.3 : Measurement table summary

| Variable | Response | Measurement | Value |
| --- | --- | --- | --- |
| In observation | Yes | N (%) | 788,114,086 (100.00%) |
| Number records | - | N | 788,114,086.00 |
| Number subjects | - | N (%) | 1,501,604 (86.54%) |
| Records per person | - | Mean (SD) | 454.21 (848.14) |
|  |  | max | 39,907.00 |
|  |  | median | 131.00 |
|  |  | min | 0.00 |
|  |  | q25 | 11.00 |
|  |  | q75 | 543.00 |
| Source vocabulary | Icd10cm | N (%) | 476,536 (0.06%) |
|  | Icd9cm | N (%) | 486,796 (0.06%) |
|  | None | N (%) | 787,150,754 (99.88%) |
| Standard concept | S | N (%) | 774,910,260 (98.32%) |
|  | - | N (%) | 13,203,826 (1.68%) |
| Type concept id | Ehr | N (%) | 4,837,023 (0.61%) |
|  | Ehr physical examination | N (%) | 48,745,090 (6.19%) |
|  | Ehr problem list | N (%) | 921,301 (0.12%) |
|  | Lab | N (%) | 733,610,672 (93.08%) |

# Summary of observation tables

Table 3.1 : Observation table summary

| Observation Period | Variable | Measurement | Value |
| --- | --- | --- | --- |
| all | Number records | N | 1,735,134 |
|  | Number subjects | N | 1,735,134 |
|  | Records per person | mean (sd) | 1.00 (0.00) |
|  |  | median [Q25 - Q75] | 1 [1 - 1] |
|  | Duration in days | mean (sd) | 1,936.98 (2,523.45) |
|  |  | median [Q25 - Q75] | 1,101.00 [189.00 - 2,769.00] |
| 1st | Number subjects | N | 1,735,134 |
|  | Duration in days | mean (sd) | 1,936.98 (2,523.45) |
|  |  | median [Q25 - Q75] | 1,101.00 [189.00 - 2,769.00] |

# Vocabulary Mapping

Table 4.1 : Vocabulary mapping completedness

| Domain | #Codes Source | #Codes Mapped | %Codes Mapped | #Records Source | #Records Mapped | %Records Mapped |
| --- | --- | --- | --- | --- | --- | --- |
| condition | 44,184 | 44,073 | 99.748778 | 31,373,274 | 31,371,426 | 99.994110 |
| condition status | 1 | 0 | 0.000000 | 31,373,274 | 0 | 0.000000 |
| death cause | 4 | 0 | 0.000000 | 443,496 | 0 | 0.000000 |
| drug exposure | 135,116 | 126,795 | 93.841588 | 151,874,418 | 141,354,691 | 93.073404 |
| drug route | 62 | 55 | 88.709677 | 151,874,418 | 66,703,233 | 43.919993 |
| measurement | 224,401 | 10,461 | 4.661744 | 788,114,086 | 774,910,260 | 98.324630 |
| measurement unit | 184 | 172 | 93.478261 | 788,114,086 | 382,990,217 | 48.595784 |
| measurement value | 11,989 | 473 | 3.945283 | 788,114,086 | 7,604,734 | 0.964928 |
| observation | 2,676 | 2,676 | 100.000000 | 8,330,594 | 8,330,594 | 100.000000 |
| observation unit | 1 |  |  | 8,330,594 |  |  |
| procedure occurrence | 1,682 | 1,682 | 100.000000 | 8,097,829 | 8,097,829 | 100.000000 |
| provider speciality | 254 | 168 | 66.141732 | 28,418 | 24,331 | 85.618270 |
| visit | 2 | 2 | 100.000000 | 59,661,220 | 59,661,220 | 100.000000 |

# MEDOC to OMOP mapping

Table 5.1 : MEDOC concept has valid concept id and locations

| medoc\_concept | omop\_table | check | result |
| --- | --- | --- | --- |
| date\_of\_birth | person | variable is present in cdm | TRUE |
| sex | person | concepts present in omop tables | TRUE |
| healthcare\_id | person | variable is present in cdm | TRUE |
| primary\_cancer\_diagnosis | condition\_occurrence | concepts present in omop tables | TRUE |
| primary\_diagnosis\_date | condition\_occurrence | variable is present in cdm | TRUE |
| comorbidities | condition\_occurrence | concepts present in omop tables | TRUE |
| date\_of\_death | death | variable is present in cdm | TRUE |
| date\_of\_visit | visit\_occurrence | variable is present in cdm | TRUE |
| participation\_in\_clinical\_trial | observation | concepts present in omop tables | FALSE |
| date\_of\_trial\_consent | observation | filtered by participation\_in\_clincial\_trial | FALSE |
| menopausal\_status | observation | concepts present in omop tables | FALSE |
| weight | measurement | concepts present in omop tables | FALSE |
| height | measurement | concepts present in omop tables | TRUE |
| performance\_status | measurement | concepts present in omop tables | FALSE |
| biomarker\_name | measurement | concepts present in omop tables | TRUE |
| biomarker\_measure | measurement | variable is present in cdm | TRUE |
| biomarker\_sample\_id | measurement | variable is present in cdm | TRUE |
| metastasis\_presence | measurement | concepts present in omop tables | FALSE |
| metastasis\_location | measurement | concepts present in omop tables | TRUE |
| extent\_of\_debulking | measurement | filtered by extent\_of\_debulking | FALSE |
| method\_of\_primary\_diagnosis | procedure\_occurrence | concepts present in omop tables | FALSE |
| radiotherapy\_type | procedure\_occurrence | concepts present in omop tables | FALSE |
| radiotherapy\_start\_date | procedure\_occurrence | filtered by radiotherapy\_type | FALSE |
| radiotherapy\_end\_date | procedure\_occurrence | filtered by radiotherapy\_type | FALSE |
| radiotherapy\_dose | see\_radiotherapy\_dose |  |  |
| surgery\_type | procedure\_occurrence | concepts present in omop tables | FALSE |
| surgery\_date | procedure\_occurrence | filtered by surgery\_type | FALSE |
| anti\_cancer\_treatment\_name | see\_summary\_drugs |  |  |
| molecule\_generic\_name | see\_summary\_drugs |  |  |
| start\_date\_drug | see\_summary\_drugs |  |  |
| treatment\_dose | see\_summary\_drugs |  |  |
| end\_drug\_treatment | see\_summary\_drugs |  |  |
| location | location | variable is present in cdm | TRUE |
| care\_site | care\_site | variable is present in cdm | TRUE |
| provider | provider | variable is present in cdm | TRUE |
| observation\_period | observation\_period | variable is present in cdm | TRUE |
| legal\_basis | derived |  |  |
| line\_of\_therapy | derived |  |  |
| time\_to\_treatment | derived |  |  |
| vital\_status | derived |  |  |
| disease\_stage | measurement | concepts present in omop tables | FALSE |
| histological cell type | observation | concepts present in omop tables | FALSE |
| histological cell type | condition\_occurrence | concepts present in omop tables | TRUE |

Table 5.2 : Episode table checks

| Table | Present |
| --- | --- |
| EPISODE | TRUE |
| EPISODE\_EVENT | TRUE |

# Cancer phenotype

## Cancer diagnosis

Table 6.1 : Primary diagnosis - concept ids

| concept\_name | concept\_id | total\_patient\_count |
| --- | --- | --- |
| Malignant neoplasm of upper lobe of lung | 4,092,216 | 89755 |
| Malignant neoplasm of corpus uteri, excluding isthmus | 194,286 | 23999 |
| Malignant neoplasm, overlapping lesion of breast | 4,092,513 | 17615 |
| Secondary malignant neoplasm of brain and spinal cord | 378,087 | 15373 |
| Malignant neoplasm of pituitary gland and craniopharyngeal duct | 4,095,304 | 2736 |
| Malignant neoplasm of connective and soft tissue of hip and lower limb | 4,089,777 | 2720 |
| Malignant melanoma of ear and/or external auditory canal | 4,095,589 | 1516 |
| Hodgkin's disease, nodular sclerosis of intrathoracic lymph nodes | 136,928 | 1180 |
| Lymphosarcoma | 200,662 | 464 |
| Hodgkin's disease, lymphocytic-histiocytic predominance of lymph nodes of multiple sites | 321,522 | 293 |
| Burkitt's lymphoma of intra-abdominal lymph nodes | 4,097,560 | 195 |
| Hodgkin's paragranuloma of intrathoracic lymph nodes | 438,709 | 120 |
| Kaposi's sarcoma of soft tissue | 375,479 | 119 |
| Hodgkin's disease, mixed cellularity of intrathoracic lymph nodes | 136,655 | 116 |
| Malignant neoplasm of auditory tube, middle ear and mastoid air cells | 377,229 | 37 |
| Lymphosarcoma of lymph nodes of axilla and upper limb | 441,803 | 32 |
| Malignant neoplasm of peripheral nerve of thorax | 4,094,262 | 19 |
| Hodgkin's paragranuloma of intrapelvic lymph nodes | 198,986 | 10 |
| Gleason grade score 4 out of 10 | 4,086,970 | 5 |
| Malignant neoplasm of liver and intrahepatic bile ducts | 4,095,432 | 5 |

## Metastasis diagnosis

Table 6.2 : Metastasis diagnosis - Measurement concept ids

| concept\_name | concept\_id | total\_patient\_count |
| --- | --- | --- |
| Distant metastases.clinical [Class] Cancer | 3,006,575 | 674940 |
| Metastasis | 36,769,180 | 19368 |
| Metastasis to spinal cord | 35,225,743 | 15373 |
| Metastasis to lung | 36,770,283 | 13197 |
| Metastasis to left lung | 35,225,630 | 6111 |
| Metastasis to bone marrow | 35,226,074 | 2797 |
| Metastasis to large intestine | 35,225,543 | 2377 |
| Metastasis to pleura | 35,226,258 | 1809 |
| Metastasis to digestive organ | 35,226,222 | 1538 |
| Metastasis to adrenal gland | 35,225,568 | 1176 |
| Metastasis to right adrenal gland | 35,226,288 | 1034 |
| Metastasis to breast | 35,225,556 | 785 |
| Metastasis to small intestine | 35,225,719 | 432 |
| Metastasis to respiratory tract | 35,226,280 | 421 |
| Metastasis to ovary | 35,226,230 | 385 |
| Metastasis to right ovary | 35,225,656 | 354 |
| Metastasis to left ovary | 35,225,631 | 333 |
| Metastasis to left renal pelvis | 712,646 | 222 |
| Metastasis to cerebral meninges | 35,226,099 | 176 |
| Metastasis to gastrointestinal tract | 36,770,040 | 163 |

Table 6.3 : Metastasis diagnosis - Condition Occurrence concept ids

| concept\_name | concept\_id | total\_patient\_count |
| --- | --- | --- |
| Distant metastasis present | 4,163,446 | 318958 |
| Secondary malignant neoplasm of liver | 198,700 | 24600 |
| Metastatic carcinoid tumor | 45,767,695 | 2844 |
| Disseminated malignancy of unknown primary | 438,701 | 2826 |
| Secondary malignant neoplasm of large intestine | 200,348 | 1355 |
| Secondary malignant neoplasm of small intestine | 198,371 | 1226 |
| Secondary malignant neoplasm of cerebral meninges | 4,312,290 | 176 |
| Leukemic reticuloendotheliosis of lymph nodes of head, face and neck | 132,570 | 26 |

# Treatment checks

## Immunotherapy drugs

Table 7.1 : Immunotherapy drug completedness

| ingredient | ingredient\_concept\_id | n\_records | n\_patients | proportion\_of\_records\_with\_dose\_form | median\_drug\_exposure\_days\_q05\_q95 | proportion\_of\_records\_with\_negative\_drug\_exposure\_days | result\_obscured |
| --- | --- | --- | --- | --- | --- | --- | --- |
| atezolizumab | 42,629,079 | 36,289 | 4,536 | 18317 (50.5%) | 1 (1-1) | 0 (0%) | FALSE |
| avelumab | 1,593,273 | 4,162 | 283 | 1745 (41.9%) | 1 (1-1) | 0 (0%) | FALSE |
| cemiplimab | 35,200,783 | 2,956 | 325 | 1302 (44%) | 1 (1-1) | 0 (0%) | FALSE |
| durvalumab | 1,594,034 | 34,707 | 2,824 | 17181 (49.5%) | 1 (1-1) | 0 (0%) | FALSE |
| nivolumab | 45,892,628 | 270,812 | 20,032 | 142422 (52.6%) | 1 (1-98) | 0 (0%) | FALSE |
| pembrolizumab | 45,775,965 | 191,389 | 18,607 | 67272 (35.1%) | 1 (1-22) | 0 (0%) | FALSE |
| ipilimumab | 40,238,188 | 20,048 | 5,524 | 9053 (45.2%) | 1 (1-63) | 0 (0%) | FALSE |

## Chemotherapy drugs

Table 7.2 : Chemotherapy drug completedness

| ingredient | ingredient\_concept\_id | n\_records | n\_patients | proportion\_of\_records\_with\_dose\_form | median\_drug\_exposure\_days\_q05\_q95 | proportion\_of\_records\_with\_negative\_drug\_exposure\_days | result\_obscured |
| --- | --- | --- | --- | --- | --- | --- | --- |
| cyclophosphamide | 1,310,317 | 435,277 | 91,499 | 210734 (48.4%) | 1 (1-88) | 0 (0%) | FALSE |
| dacarbazine | 1,311,409 | 48,846 | 5,662 | 29467 (60.3%) | 1 (1-129.049999999999) | 0 (0%) | FALSE |
| docetaxel | 1,315,942 | 362,603 | 74,194 | 215881 (59.5%) | 1 (1-105) | 0 (0%) | FALSE |
| doxorubicin | 1,338,512 | 358,475 | 75,181 | 358475 (100%) | 1 (1-85) | 0 (0%) | FALSE |
| etoposide | 1,350,504 | 373,426 | 34,275 | 195909 (52.5%) | 1 (1-63) | 0 (0%) | FALSE |
| fluorouracil | 955,632 | 1,030,528 | 88,310 | 465044 (45.1%) | 1 (1-85) | 0 (0%) | FALSE |
| gemcitabine | 1,314,924 | 439,429 | 48,048 | 207788 (47.3%) | 1 (1-77.0499999999993) | 0 (0%) | FALSE |
| ifosfamide | 19,078,187 | 37,808 | 3,448 | 12394 (32.8%) | 1 (1-30) | 0 (0%) | FALSE |
| irinotecan | 1,367,268 | 249,586 | 26,729 | 140166 (56.2%) | 1 (1-99) | 0 (0%) | FALSE |
| methotrexate | 1,305,058 | 110,556 | 14,694 | 14684 (13.3%) | 1 (1-270.049999999999) | 0 (0%) | FALSE |
| mitomycin | 1,389,036 | 17,222 | 7,833 | 5674 (32.9%) | 1 (1-30) | 0 (0%) | FALSE |
| paclitaxel | 1,378,382 | 950,914 | 112,788 | 566397 (59.6%) | 1 (1-64) | 0 (0%) | FALSE |
| pemetrexed | 1,304,919 | 118,887 | 17,236 | 43296 (36.4%) | 1 (1-97) | 0 (0%) | FALSE |
| topotecan | 1,378,509 | 39,717 | 3,968 | 23463 (59.1%) | 1 (1-106) | 0 (0%) | FALSE |
| vinblastine | 19,008,264 | 64,622 | 7,902 | 23562 (36.5%) | 1 (1-70) | 0 (0%) | FALSE |
| vincristine | 1,308,290 | 167,711 | 26,743 | 62857 (37.5%) | 1 (1-63) | 0 (0%) | FALSE |
| vinorelbine | 1,343,346 | 43,831 | 5,088 | 22979 (52.4%) | 1 (1-106) | 0 (0%) | FALSE |
| carboplatin | 1,344,905 | 683,349 | 119,091 | 412873 (60.4%) | 1 (1-64.0999999999985) | 0 (0%) | FALSE |
| cisplatin | 1,397,599 | 293,410 | 52,038 | 153950 (52.5%) | 1 (1-49) | 0 (0%) | FALSE |
| oxaliplatin | 1,318,011 | 424,319 | 58,008 | 237908 (56.1%) | 1 (1-112) | 0 (0%) | FALSE |

## Targeted therapy drugs

Table 7.3 : Targeted therapy drug completedness

| ingredient | ingredient\_concept\_id | n\_records | n\_patients | proportion\_of\_records\_with\_dose\_form | median\_drug\_exposure\_days\_q05\_q95 | proportion\_of\_records\_with\_negative\_drug\_exposure\_days | result\_obscured |
| --- | --- | --- | --- | --- | --- | --- | --- |
| bevacizumab | 1,397,141 | 440,443 | 40,228 | 273492 (62.1%) | 1 (1-151.099999999999) | 0 (0%) | FALSE |
| necitumumab | 35,606,215 | 1,197 | 128 | 803 (67.1%) | 1 (1-3) | 0 (0%) | FALSE |
| ramucirumab | 44,818,489 | 22,158 | 3,406 | 11818 (53.3%) | 1 (1-84.0499999999993) | 0 (0%) | FALSE |
| rituximab | 1,314,273 | 415,417 | 55,391 | 262092 (63.1%) | 1 (1-112) | 0 (0%) | FALSE |
| trastuzumab | 1,387,104 | 571,947 | 33,412 | 571947 (100%) | 1 (1-173) | 0 (0%) | FALSE |
| afatinib | 43,533,090 | 2,960 | 1,352 | 2340 (79.1%) | 31 (1-459.3) | 0 (0%) | FALSE |
| alectinib | 35,604,657 | 2,320 | 440 | 1042 (44.9%) | 21 (1-189) | 0 (0%) | FALSE |
| brigatinib | 1,594,187 | 168 | 71 | 81 (48.2%) | 10.5 (1-223.3) | 0 (0%) | FALSE |
| cabozantinib | 43,012,292 | 5,237 | 1,797 | 2719 (51.9%) | 30 (1-226.2) | 0 (0%) | FALSE |
| ceritinib | 44,818,466 | 723 | 184 | 287 (39.7%) | 12 (1-210.9) | 0 (0%) | FALSE |
| crizotinib | 40,242,675 | 3,450 | 927 | 1149 (33.3%) | 27 (1-457.099999999999) | 0 (0%) | FALSE |
| dabrafenib | 43,532,299 | 5,226 | 1,250 | 1703 (32.6%) | 18 (1-313) | 0 (0%) | FALSE |
| dasatinib | 1,358,436 | 20,465 | 3,962 | 8371 (40.9%) | 29 (1-432.099999999999) | 0 (0%) | FALSE |
| entrectinib | 37,496,447 | 20 | 12 | 17 (85%) | 30 (1-181.9) | 0 (0%) | FALSE |
| erlotinib | 1,325,363 | 16,667 | 6,385 | 9701 (58.2%) | 30 (1-888) | 0 (0%) | FALSE |
| gefitinib | 1,319,193 | 591 | 217 | 316 (53.5%) | 31 (1-328.6) | 0 (0%) | FALSE |
| larotrectinib | 1,355,705 | 32 | 25 | 28 (87.5%) | 30 (1-175.5) | 0 (0%) | FALSE |
| lorlatinib | 35,201,733 | 263 | 70 | 58 (22.1%) | 30 (1-137) | 0 (0%) | FALSE |
| nintedanib | 45,775,396 | 1,005 | 101 | 20 (2%) | 1 (1-309) | 0 (0%) | FALSE |
| osimertinib | 35,605,522 | 6,385 | 1,714 | 3568 (55.9%) | 30 (1-222.799999999999) | 0 (0%) | FALSE |
| pazopanib | 40,167,554 | 10,130 | 3,491 | 5581 (55.1%) | 30 (1-474) | 0 (0%) | FALSE |
| trametinib | 43,532,497 | 4,739 | 1,453 | 1591 (33.6%) | 29 (1-352.199999999999) | 0 (0%) | FALSE |
| vandetanib | 40,238,052 | 198 | 73 | 83 (41.9%) | 31 (1-849.25) | 0 (0%) | FALSE |
| dacomitinib | 35,200,803 | 33 | 15 | 16 (48.5%) | 1 (1-35.8) | 0 (0%) | FALSE |
| capmatinib | 1,145,637 | 20 | 11 | 0 (0%) | 29 (1-183.25) | 0 (0%) | FALSE |

## Radiotherapy treatment checks

Table 7.4 : Radiotherapy completedness

| name | procedure\_concept\_id | unique\_person\_count | unique\_record\_count | negative\_procedure\_days\_count | after\_death\_count | proportion\_of\_records\_with\_negative\_procedure\_days | proportion\_of\_after\_death\_procedures | type |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |

Table 4.5 : Radiotherapy dosage

| radiation dose | status |
| --- | --- |
| radiation dose info available? | FALSE |

# Procedures check

Table 8.1 : All procedures check

| unique\_person\_count | unique\_record\_count | negative\_procedure\_days\_count | after\_death\_count | proportion\_of\_records\_with\_negative\_procedure\_days | proportion\_of\_after\_death\_procedures |
| --- | --- | --- | --- | --- | --- |
| 863,670 | 8,089,314 |  | 104 | (%) | 104 (0.01%) |

# TNM check

Table 9.1 : TNM code check

| staging stored as value | value\_as\_concept\_id contains tumour stage | tumour stage stored in measurement\_concept\_id |
| --- | --- | --- |
| FALSE | FALSE | FALSE |