



Medicines & Healthcare products
Regulatory Agency



CPRD Aurum Lookups Specification

Version 1.0

Date: 26 February 2021



Documentation Control Sheet

It may be necessary to issue amendments or clarifications to parts of this document. This form must be updated whenever changes are made.

| Version | Summary of Change | Prepared By | Date | Reviewed By | Date |
|---------|-------------------|-----------------|------------|------------------|------------|
| 1.0 | Initial draft | Hilary Shepherd | 09/09/2020 | Arlene Gallagher | 11/03/2021 |
| | | | | | |

Summary of Changes

Version 1.0

- Initial draft

Table of Contents

| | |
|--|-------------------------------------|
| Dictionaries | 4 |
| Medical Dictionary (YYYYMM_EMISMedicalDictionary.txt) | 4 |
| Product Dictionary (YYYYMM_EMISProductDictionary.txt) | 4 |
| Other lookups | 6 |
| Common Dosage (common_dosages.txt) | 6 |
| Consultation Source Identifier (ConsSource.txt) | 7 |
| EMIS Code Category Identifier (EmisCodeCat.txt) | 7 |
| Gender (Gender.txt) | 7 |
| Job Category (JobCat.txt) | 7 |
| Numeric Unit Identifier (NumUnit.txt) | 8 |
| Observation Type Identifier (ObsType.txt) | 8 |
| Organisation Identifier (OrgType.txt) | 8 |
| Parent Problem Relationship Identifier (ParentProbRel.txt) | 8 |
| Patient Type (PatientType.txt) | 8 |
| Problem Status Identifier (ProbStatus.txt) | 9 |
| Quantity Unit Identifier (QuantUnit.txt) | 9 |
| Referral Mode Identifier (RefMode.txt) | 9 |
| Referral Service Type Identifier (RefServiceType.txt) | 9 |
| Referral Urgency Identifier (RefUrgency.txt) | 10 |
| Region (Region.txt) | 10 |
| Significance (Sign.txt) | 10 |
| Migrators file (VisionToEmisMigrators.txt) | Error! Bookmark not defined. |

All files are provided in tab-delimited text format.

Dictionaries

Medical Dictionary (YYYYMM_EMISMedicalDictionary.txt)

This file contains the descriptions of all medical codes referenced in the data files as 'medcodeid'.

| Column name | Description | Mapping | Type | Format |
|-----------------------|---|--------------------------|---------|-----------------------------|
| medcodeid | CPRD code to describe the observation. Links to the observation table. This is the primary key for this table. | None | TEXT | 6-18 numeric characters |
| term | Description of the observation associated with the codeid | None | TEXT | 255 characters |
| originalreadcode | The original Read code text as provided in the EMIS® dictionary (contains codes which are not valid Read codes) | None | TEXT | 100 characters |
| cleansedreadcode | CPRD-cleaned and validated version of the originalreadcode | None | TEXT | 7 characters |
| snomedctconceptid | The SNOMED CT concept identifier associated with the observation | None | TEXT | Up to 19 numeric characters |
| snomedctdescriptionid | The SNOMED CT description identifier associated with the observation | None | TEXT | Up to 19 numeric characters |
| release | Reference data version. [Not currently populated] | None | TEXT | 100 characters |
| emiscodecategoryid | Category identifier in EMIS® that describes the observation | Lookup: EMISCode Cat.txt | INTEGER | 2 |

Product Dictionary (YYYYMM_EMISProductDictionary.txt)

This file contains descriptions of all product codes referenced in the data files as 'prodcodeid'.

| Column name | Description | Mapping | Type | Format |
|-----------------------|---|---------|------|-----------------------------|
| prodcodeid | CPRD code to describe the treatment. Links to the Drug Issue table. This is the primary key for this table. | None | TEXT | 6-18 numeric characters |
| dmdid | The DM+D code associated with the treatment | None | TEXT | Up to 19 numeric characters |
| termfromemis | Description of the treatment provided by EMIS® associated with the prodcodeid | None | TEXT | 255 characters |
| productname | Name of the product | None | TEXT | Up to 999 characters |
| formulation | Formulation of the product | None | TEXT | Up to 999 characters |
| routeofadministration | Route of administration for the product | None | TEXT | Up to 999 characters |
| drugsubstanceid | Active ingredient(s) included in the | None | TEXT | Up to 999 |

| | | | | |
|-------------------|--|------|------|----------------------|
| | product. For combination therapies, each component is listed, separated by / | | | characters |
| substancestrength | Strength of each active ingredient listed in the drugsubstancecolumn, including units. Separated by / if more than 1 | None | TEXT | Up to 999 characters |
| bnfchapter | BNF chapter to which the product belongs | None | TEXT | Up to 999 characters |
| release | Reference data version. [Not currently populated] | None | | |

Other lookups

Common Dosage (common_dosages.txt)

This file contains dosage information for drug events and is recorded in the Drug Issue table.

| <i>Column name</i> | <i>Description</i> | <i>Type</i> | <i>Format</i> |
|--------------------|---|-------------|-----------------------|
| dosageid | Identifier that allows dosage text information in therapy data to be retrieved | TEXT | 64 |
| dosage_text | Anonymised dosage text associated with dosageid | TEXT | Up to 1000 characters |
| daily_dose | Numerical daily dose extracted from dosage text | NUMERIC | 15.3 |
| dose_number | Amount in each dose | NUMERIC | 15.3 |
| dose_unit | Unit of each dose | TEXT | Up to 7 characters |
| dose_frequency | How often a dose is taken in a day | NUMERIC | 15.3 |
| dose_interval | Time span in days that the dose is over, e.g. 1 every 2 weeks = 14, 4 a day = 0.250 | NUMERIC | 15.3 |
| choice_of_dose | Indicates if there is a choice the user can make as to how much they can take | INTEGER | 1 |
| dose_max_average | If dose was averaged, value = 2, if maximum was taken, value = 1, otherwise 0 | INTEGER | 1 |
| change_dose | If an option between 2 parts of the dose was available, indicates the part used | INTEGER | 1 |
| dose_duration | If specified, the number of days the prescription is for | NUMERIC | 15.3 |

Consultation Source Identifier (ConsSource.txt)

This file contains descriptions of the most common locations of consultations (surgery clinic, nurse surgery) as recorded by EMIS® and is recorded in the Consultation table.

| Column name | Description | Type | Format |
|-------------|--|---------|----------------------|
| Id | Unique identifier for the source or location of the consultation | INTEGER | Up to 10 digits |
| Description | A description of the type of consultation | TEXT | Up to 255 characters |

EMIS Code Category Identifier (EmisCodeCat.txt)

This file describes category identifiers in EMIS® that are used to describe observations (health management, family history) as recorded in the medical dictionary.

| Column name | Description | Type | Format |
|---------------|---|---------|---------------------|
| emiscodecatid | Unique identifier for the category that describes the observation | INTEGER | Up to 5 digits |
| Description | A description of the category | TEXT | Up to 64 characters |

Gender (Gender.txt)

This file describes the gender of patients in the data and is recorded in the Patient table.

| Column name | Description | Type | Format |
|-------------|---|---------|-------------|
| genderid | Unique identifier for the gender | INTEGER | 1 |
| Description | A description of the identifier. M=male, F=female, I=indeterminate, U=unknown | TEXT | 1 character |

Job Category (JobCat.txt)

This file describes the job category of the member of staff who created the event (consultant, midwife) and is recorded in the Staff table.

| Column name | Description | Type | Format |
|-------------|---|---------|----------------------|
| jobcatid | Unique identifier for the job category of the member of staff who created the event | INTEGER | Up to 5 digits |
| Description | A description of the job category | TEXT | Up to 100 characters |

Numeric Unit Identifier (NumUnit.txt)

This file provides units for measurements recorded in the “value” field of the Observation table (per day, kg, beats per minute).

| Column name | Description | Type | Format |
|-------------|---|---------|----------------------|
| numunitid | Unique identifier for the unit of measurement | INTEGER | Up to 10 digits |
| Description | A description of the identifier | TEXT | Up to 200 characters |

Observation Type Identifier (ObsType.txt)

This file describes the type of observation (allergy, immunisation) and is recorded in the Observation table.

| Column name | Description | Type | Format |
|-------------|---|---------|----------------------|
| obstypeid | Unique identifier for the type of observation | INTEGER | Up to 5 digits |
| Description | A description of the identifier | TEXT | Up to 255 characters |

Organisation Identifier (OrgType.txt)

This file is not currently populated. It will describe the source organisation of the referral in the Referral table.

| Column name | Description | Type | Format |
|-------------|-----------------|------|--------|
| orgtypeid | Awaiting review | | |
| Description | Awaiting review | | |

Parent Problem Relationship Identifier (ParentProbRel.txt)

This file provides a description of the current problem to another problem and is recorded in the Problem table.

| Column name | Description | Type | Format |
|-----------------|--|---------|----------------------|
| parentprobrelid | Unique identifier for the relationship of the problem to another problem | INTEGER | Up to 5 digits |
| Description | A description of the identifier | TEXT | Up to 255 characters |

Patient Type (PatientType.txt)

This file describes the category that a patient has been assigned to (temporary, regular) and is recorded in the Patient table.

| Column name | Description | Type | Format |
|---------------|---------------------------|---------|----------------|
| patienttypeid | Unique identifier for the | INTEGER | Up to 5 digits |

| | | | |
|-------------|--|------|---------------------|
| | category that the patient has been assigned to | | |
| Description | A description of the category | TEXT | Up to 50 characters |

Problem Status Identifier (ProbStatus.txt)

This file describes the status of a problem (active problem, past problem) and is recorded in the Problem table. The file is provided in tab-delimited text format.

| Column name | Description | Type | Format |
|--------------|--|---------|----------------------|
| probstatusid | Unique identifier for the problem status | INTEGER | Up to 5 digits |
| Description | A description of the problem status | TEXT | Up to 255 characters |

Quantity Unit Identifier (QuantUnit.txt)

This file describes the unit of the treatment (capsule, tablet) and is recorded in the Drug Issue table.

| Column name | Description | Type | Format |
|-------------|---|---------|----------------------|
| quantunitid | Unique identifier for the quantity of treatment | INTEGER | Up to 5 digits |
| Description | A description of the quantity | TEXT | Up to 255 characters |

Referral Mode Identifier (RefMode.txt)

This file describes the mode by which a referral was made (telephone, written) and is recorded in the Referral table.

| Column name | Description | Type | Format |
|-------------|--|---------|----------------------|
| refmodeid | Unique identifier for the type of referral | INTEGER | Up to 5 digits |
| Description | A description of the referral | TEXT | Up to 255 characters |

Referral Service Type Identifier (RefServiceType.txt)

This file describes the type of service that the referral relates to (assessment, management) and is recorded in the Referral table.

| Column name | Description | Type | Format |
|------------------|---|---------|----------------------|
| refservicetypeid | Unique identifier for the type of service the referral relates to | INTEGER | Up to 5 digits |
| Description | A description of the type of the service | TEXT | Up to 255 characters |

Referral Urgency Identifier (RefUrgency.txt)

This file describes the urgency of the referral (routine, urgent) and is recorded in the Referral table.

| <i>Column name</i> | <i>Description</i> | <i>Type</i> | <i>Format</i> |
|--------------------|---|-------------|----------------------|
| refurgencyid | Unique identifier for the urgency of referral | INTEGER | Up to 5 digits |
| Description | A description of the urgency | TEXT | Up to 255 characters |

Region (Region.txt)

This file describes the region in the UK where the GP practice is based (North East, Midlands) and is recorded in the Practice table. The region denotes the Strategic Health Authority for English practices.

| <i>Column name</i> | <i>Description</i> | <i>Type</i> | <i>Format</i> |
|--------------------|--|-------------|----------------------|
| regionid | Unique identifier for a region in the UK | INTEGER | Up to 2 digits |
| Description | A description of the region | TEXT | Up to 100 characters |

Significance (Sign.txt)

This file describes the significance of the problem (minor, significant) and is recorded in the Problem table.

| <i>Column name</i> | <i>Description</i> | <i>Type</i> | <i>Format</i> |
|--------------------|---|-------------|----------------------|
| signid | Unique identifier for the significance of the problem | INTEGER | Up to 5 digits |
| Description | A description of the significance | TEXT | Up to 255 characters |