**Lookups Data Specification**

Table of Contents

[3-Character Lookup Files 2](#_Toc428270032)

[Combined Lookup File (LookupsXXX.doc) 2](#_Toc428270033)

[Individual Lookup Files (XXX.txt) 2](#_Toc428270034)

[Dictionaries 3](#_Toc428270035)

[Medical Dictionary (medical.txt) 3](#_Toc428270036)

[Product Dictionary (product.txt) 3](#_Toc428270037)

[Clinical Details Specific Lookups 4](#_Toc428270038)

[Entity (entity.xls) 4](#_Toc428270039)

[Scoring (scoremethod.txt) 4](#_Toc428270040)

[Product Specific Lookups 5](#_Toc428270041)

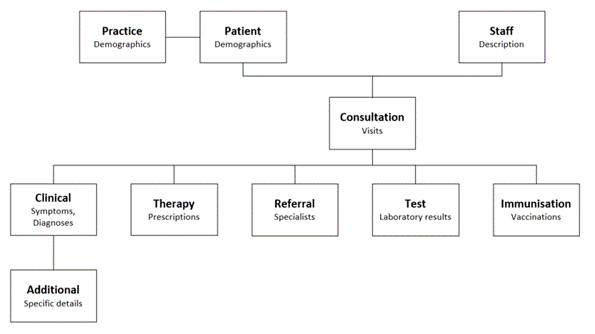
[Product Pack Type (packtype.txt) 5](#_Toc428270042)

[Product BNF Code (bnfcodes.txt) 5](#_Toc428270043)

[Common Dosage (common\_dosage.txt) 6](#_Toc428270044)

[Immunisation Specific Lookups 7](#_Toc428270045)

[Batch Number (batchnumber.txt) 7](#_Toc428270046)



# 3-Character Lookup Files

## Combined Lookup File (LookupsXXX.doc)

This file contains the lookup values for all 3-character lookups. The combined lookup file is provided in .doc format.

## Individual Lookup Files (XXX.txt)

Each file contains the lookup code and text description for the corresponding 3-character lookup. These files are provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| code | The coded value associated with lookup | INTEGER | 5 |
| text | A description | CHAR | 100 |

# Dictionaries

## Medical Dictionary (medical.txt)

This file contains the descriptions of all medical codes referenced in the data files as ‘medcode’. The medical dictionary file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| medcode | CPRD unique code for the medical term selected by the GP | INTEGER | 20 |
| read\_code | Read Code | CHAR | 7 |
| desc | Description of the medical term | CHAR | 100 |

## Product Dictionary (product.txt)

This file contains descriptions of all product codes referenced in the data files as ‘prodcode’. The product dictionary file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| prodcode | CPRD unique code for the treatment selected by the GP | INTEGER | 20 |
| dmdcode | Unique product identifier from the NHS Dictionary of Medicines and Devices (dm+d) – the NHS standard dictionary for products licensed in the UK | CHAR | 100 |
| gemscriptcode\* | Gemscript product code for the corresponding product name - should be treated as a string field as it contains leading ‘0’’s | CHAR | 8 |
| productname | Product name as entered at the practice | CHAR | 500 |
| drugsubstance | Drug substance | CHAR | 1500 |
| strength | Strength of the product | CHAR | 500 |
| formulation | Form of the product e.g. tablets, capsules etc | CHAR | 100 |
| route | Route of administration of the product | CHAR | 100 |
| bnfcode | British National Formulary (BNF) code | CHAR | 100 |
| bnfchapter | British National Formulary (BNF) chapter | CHAR | 500 |

# Clinical Details Specific Lookups

## Entity (entity.xls)

This file contains detailed information on all Entity Types referenced in the Clinical, Additional Clinical Details (Additional), and Test data files as ‘enttype’. Entity.xls is an excel spreadsheet, and contains information on the number of data fields associated with each entity, along with the description of the data within. The lookups needed to decode the information are also detailed.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| enttype | Number identifying the structured data area inVision where data was entered | INTEGER | 5 |
| description | Description of entity type | CHAR | 60 |
| filetype | Type of data file (‘Test’, or ‘Clinical’) that references entity type | CHAR | 8 |
| category | Broad category for entity type description | CHAR | 30 |
| data\_fields | Number of data fields associated with the entity type | INTEGER | 1 |
| data1\_desc | Description of data in data field 1 | CHAR | 40 |
| data1\_lkup\* | Lookup file needed to decode values in data field 1 | CHAR | 20 |
| data2\_desc | Description of data in data field 2 | CHAR | 40 |
| data2\_lkup\* | Lookup file needed to decode values in data field 2 | CHAR | 20 |
| data3\_desc | Description of data in data field 3 | CHAR | 40 |
| data3\_lkup\* | Lookup file needed to decode values in data field 3 | CHAR | 20 |
| data4\_desc | Description of data in data field 4 | CHAR | 40 |
| data4\_lkup\* | Lookup file needed to decode values in data field 4 | CHAR | 20 |
| data5\_desc | Description of data in data field 5 | CHAR | 40 |
| data5\_lkup\* | Lookup file needed to decode values in data field 5 | CHAR | 20 |
| data6\_desc | Description of data in data field 6 | CHAR | 40 |
| data6\_lkup\* | Lookup file needed to decode values in data field 6 | CHAR | 20 |
| data7\_desc | Description of data in data field 7 | CHAR | 40 |
| data7\_lkup\* | Lookup file needed to decode values in data field 7 | CHAR | 20 |
| data8\_desc | Description of data in data field 8 | CHAR | 40 |
| data8\_lkup\* | Lookup file needed to decode values in data field 8 | CHAR | 20 |
| data9\_desc | Description of data in data field 9 | CHAR | 40 |
| data9\_lkup\* | Lookup file needed to decode values in data field 9 | CHAR | 20 |
| data10\_desc | Description of data in data field 10 | CHAR | 40 |
| data10\_lkup\* | Lookup file needed to decode values in data field 10 | CHAR | 20 |
| data11\_desc | Description of data in data field 11 | CHAR | 40 |
| data11\_lkup\* | Lookup file needed to decode values in data field 11 | CHAR | 20 |
| data12\_desc | Description of data in data field 12 | CHAR | 40 |
| data12\_lkup\* | Lookup file needed to decode values in data field 12 | CHAR | 20 |

## Scoring (scoremethod.txt)

This file is specifically for the data3 field of entity type 372 in the clinical details data files, which refers to scoring methodology (e.g. Framingham). It contains details of the scoring methodology used. The scoring file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| code | Coded value associated with the scoring methodology used | INTEGER | 5 |
| scoringmethod | Scoring methodology | CHAR | 20 |

# Product Specific Lookups

## Product Pack Type (packtype.txt)

This file contains descriptions of the pack size or type referenced in the therapy files as ‘packtype’. The product pack type file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| packtype | Coded value associated with the pack size or type of the prescribed product | INTEGER | 10 |
| packtype\_desc | Pack size or type of the prescribed product | CHAR | 21 |

## Product BNF Code (bnfcodes.txt)

This file contains the British National Formulary code referenced in the therapy files, and details the chapter & section in the BNF. The product BNF code file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| bnfcode | Coded value for the actual BNF code representation of the prescribed product | INTEGER | 5 |
| bnf | BNF code representing the chapter and section in the British National Formulary (BNF) for the prescribed product (nn.nn.nn.nn is coded as nnnnnnnn) | CHAR | 8 |

## Common Dosage (common\_dosage.txt)

This file contains dosage information (sorted alphabetically) for therapy events that can be referenced in the therapy files via the 64 character dosageid values. The file also contains extracted numerical information from the dosage text such as the daily dose. Over 95% of dosage text occurring in the data should be found in this file. The common dosage file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| dosageid | Identifier that allows dosage text information in therapy data to be retrieved | CHAR | 64 |
| dosage\_text | Anonymised dosage text associated with dosageid | CHAR | 1000 |
| 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 | CHAR | 7 |
| 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 |

# Immunisation Specific Lookups

## Batch Number (batchnumber.txt)

This file contains the immunisation batch number for the encoded field referenced in the immunisation files as ‘batch’. The batch number file is provided in tab-delimited text format.

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **Description** | **Type** | **Format** |
| batch | Coded value associated with the immunisation batch number | INTEGER | 20 |
| batch\_number | Immunisation batch number | CHAR | 200 |