



## Release Notes: CPRD GOLD July 2021

### Summary

An updated version of the CPRD GOLD database (data sourced from GP practices using Vision® GP software) is now available, which contains the following:

Metric	Coverage
Total number of acceptable patients <sup>1</sup> (including transferred out and deceased patients):	20,418,567
Current acceptable patients (i.e. registered at currently contributing practices, excluding transferred out and deceased patients):	3,133,918
Percentage UK population coverage <sup>2</sup> (current patients only):	3,133,918 of 66,796,800 (4.69%)
Total patients eligible for linkage (Set 21):	9,284,356
Available follow-up time in years since 1 <sup>st</sup> January 1995 <sup>3</sup> (all patients including transferred out and deceased): Mean (Standard Deviation): Median (25 <sup>th</sup> and 75 <sup>th</sup> Percentile):	8.37 (7.73) 5.59 (1.98 – 13.43)
Available follow-up time in years (current patients only): Mean (Standard Deviation): Median (25 <sup>th</sup> and 75 <sup>th</sup> Percentile):	13.97 (9.57) 13.00 (4.76 – 25.65)
Total number of practices (current and historic) included in the database:	971
Currently contributing practices:	405
Percentage coverage of UK general practices (currently contributing practices only):	405 of 8,961 (4.52%)
Regional distribution of currently contributing practices <sup>4</sup> England: Northern Ireland: Scotland: Wales:	40 (9.88%) 40 (9.88%) 212 (52.35%) 113 (27.90%)

<sup>1</sup> Permanent registrations only. Over 98% of permanent registrations are deemed to have 'acceptable' (or research quality) data based on CPRD metrics.

<sup>2</sup> [Based on latest UK population estimates from the Office of National Statistics.](#)

<sup>3</sup> Follow-up time stated here does not incorporate the up-to-standard (UTS) date and the database includes records pre-dating the 1<sup>st</sup> of January 1995.

<sup>4</sup> Expressed as a percentage of all practices currently contributing to CPRD GOLD.



## Lookups

Indication of monthly changes to the Lookup files is given below. These files can be requested from CPRD Enquiries at [enquiries@cprd.com](mailto:enquiries@cprd.com)

Lookup	Filename	Records
Medical Dictionary	medical.txt	115,707
Product Dictionary	product.txt	84,285
Entity	entity.xls	501
Scoring	scoremethod.txt	5,108
Batch Number	batchnumber.txt	4,449,676
Product Pack Type	packtype.txt	89,615
Product BNF Code	bnfcodes.txt	2,455
Common Dosage	common_dosages.txt	73,068
Practice Details	allpractices_JUL2021.txt	971
All Patient Details	all_patients_JUL2021.txt	23,230,380
Acceptable Patient Details	acceptablepats_from_utspracts_JUL2021.txt	20,418,567

## Other Updates

The following updates have been made to the July 2021 release.

As part of the gradual shift to SNOMED coding, Vision implemented automatic mapping of Read codes to SNOMED Concept IDs from early 2018. Some versions of the Vision software also allow direct coding using SNOMED description IDs.

To reflect this shift, CPRD is piloting the inclusion of 7 new variables in tables that include a medcode variable in [CPRD GOLD](#) from July 2021. These include a SNOMED CT Concept ID field and a SNOMED CT Description ID field. The CPRD GOLD Full Data Specification document will include details of each field but please note the following:

- The SNOMED description ID field will only be complete if the native (original) coding was in SNOMED CT.
- Events recorded using Read codes are mapped to SNOMED Concept IDs by Vision at the point of data collection. These are not validated by CPRD. Mapping has not been applied retroactively and mapping files are regularly updated. This means the mapping is not 1:1 across the database.



- Where information was directly coded in SNOMED CT and appropriate Read codes were not available, the mapped Read code is likely to be non-specific. It is in these instances that these additional variables will be most valuable.
- The online tools will continue to require medcode lists, which are mapped to Read codes.

Two further new variables will be present in the Therapy file: a mapped dmdcode and a prn variable to indicate an 'as required' dosage instruction. Neither field appears to be backfilled.

## DOI

Please cite in any publications using this version <https://doi.org/10.48329/yycw-n139>

## CPRD GOLD Data Specification

These files can be downloaded from the CPRD website: <https://www.cprd.com/primary-care>