Release Notes: CPRD GOLD January 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[1] (including transferred out and deceased patients):	19,483,855
Current acceptable patients (i.e. registered at currently contributing practices, excluding transferred out and deceased patients):	3,020,680
Percentage UK population coverage[2] (current patients only):	3,020,680 of 66,796,800 (4.52%)
Total patients eligible for linkage (Set 20):	9,209,834
Available follow-up time in years since 1st January 1995[3] (all patients including transferred out and deceased): Mean (Standard Deviation): Median (25th and 75th Percentile):	8.35 (7.65) 5.62 (1.98 – 13.45)
Available follow-up time in years (current patients only): Mean (Standard Deviation): Median (25th and 75th Percentile):	13.79 (9.35) 12.92 (4.79 – 25.26)
Total number of practices (current and historic) included in the database:	945
Currently contributing practices:	394
Percentage coverage of UK general practices (currently contributing practices only):	394 of 8,961 (4.40%)
Regional distribution of currently contributing practices[4] England: Northern Ireland: Scotland: Wales:	54 (13.71%) 39 (9.90%) 197 (50.00%) 104 (26.40%)

---

[1] Permanent registrations only. Over 98% of permanent registrations are deemed to have 'acceptable' (or research quality) data based on CPRD metrics.
[2] Based on latest UK population estimates from the Office of National Statistics.
[3] Follow-up time stated here does not incorporate the up-to-standard (UTS) date and the database includes records pre-dating the 1st of January 1995.
[4] 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

Lookup	Filename	Records
Medical Dictionary	medical.txt	115,344
Product Dictionary	product.txt	82,938
Entity	entity.xls	501
Scoring	scoremethod.txt	5,026
Batch Number	batchnumber.txt	4,317,418
Product Pack Type	packtype.txt	88,346
Product BNF Code	bnfcodes.txt	2,444
Common Dosage	common_dosages.txt	73,068
Practice Details	allpractices_JAN2021.txt	945
All Patient Details	all_patients_JAN2021.txt	22,226,930
Acceptable Patient Details	acceptablepats_from_utspracts_JAN2021.txt	19,483,855

DOI

Please cite in any publications using this version https://doi.org/10.11581/v8hp-qa82

Other Updates

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

Three Character Lookups

The following lookups have been added to the IMT three character lookup file: FLUQIVCOHP, FLUTIVOHP, COVCOURAGEOUS, FLUQIVCPHA, FLUQIVEOHP, FLUTIV, FLUTIVPHA, COVOTHER, FLUATIVOHP.

Changes to CPRD GOLD numbering convention

From the January 2021 build, CPRD GOLD will use the new five digit practice ID numbering convention, which will also affect the patient ID number. Customers are advised to check that any standard processing scripts used when processing data extracts will handle the updated version of pracid and patid.

Digital object identifier (DOI)

CPRD database releases for CPRD GOLD and CPRD Aurum now include a digital object identifier (DOI). The DOI links to information about the database version and how to cite the data in research findings. This update supports transparent, reproducible research and provides additional assurance of data provenance. In future, if you are using CPRD data that has a DOI, please cite this in any publications.

CPRD GOLD Data Specification

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