



Release Notes: CPRD Aurum January 2022

Summary

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

Metric	Coverage
Total number of acceptable patients ¹ (including transferred out and deceased patients):	40,833,738
Current acceptable patients (i.e. registered at currently contributing practices, excluding transferred out and deceased patients):	13,344,212
Percentage UK population coverage ² (current patients only):	13,344,212 of 67,081,000 (19.89 %)
Total patients eligible for linkage (Set 21):	37,505,167
Available follow-up time in years since 1 st January 1995 ³ (all patients including transferred out and deceased): Mean (Standard Deviation): Median (25 th and 75 th Percentile):	7.90 (7.92) 4.75 (1.82 – 11.76)
Available follow-up time in years since 1 st January 1995 (current patients only): Mean (Standard Deviation): Median (25 th and 75 th Percentile):	11.60 (9.43) 8.74 (3.22 – 19.87)
Total number of practices (current and historic) included in the database:	1,489
Currently contributing practices:	1,359
Percentage coverage of UK general practices (currently contributing practices only):	1,359 of 8,178 (16.62 %)
Regional distribution of currently contributing practices ⁴ England: Northern Ireland: Scotland: Wales:	1,346 (99.04 %) 13 (0.96 %) 0 (0.00 %) 0 (0.00 %)

¹ Permanent registrations only. Over 93% 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 Aurum.



Lookups

These files can be requested from CPRD Enquiries at enquiries@cprd.com

DOI

Please cite in any publications using this version <https://doi.org/10.48329/db7t-ay41>

Other Updates

Move to ONS Region in CPRD Aurum and CPRD GOLD practice files

With the January 2022 build of the CPRD Aurum and CPRD GOLD database we will be updating the 'region' variable in the Practice files from the former Strategic Health Authority to [ONS Regions \(former GORs\)](#) for practices within England, and, as now, the country i.e. Wales, Scotland, or Northern Ireland for practices in the devolved nations.

The only change to note is that the SHAs South Central and South East Coast are now represented by a single South East Region. Please consider what impact this might have on any in-house processing you apply.

Duplication of data between practices in the CPRD Aurum database

Duplication of data occurs when one practice is absorbed by another practice that also contributes (or goes on to contribute) to the CPRD Aurum database. This issue affects 29 practices, including 1% of the individuals recorded in the database, and all builds of the Aurum database up to and including this one.

When a practice is absorbed by another one it stops contributing data to the database, but its data remains in the database, in the same way as for all other practices that close or stop contributing data. Patients from the absorbed practice are assigned a new patient identifier in the absorbing practice, but carry across all their data to, and retain their original registration start date in, their new practice.

This is different from what happens when individual patients move between two practices that contribute to the database. In that case previous data is also carried across but, because it is before their new registration date, studies can avoid pseudoreplication by excluding data recorded before each individual's registration date. Such people are effectively split into two separate individuals within analyses.

We have identified 29 practices that appear likely to have merged into other contributing practices and suggest that you may want to exclude these from your studies, wherever possible. The practice identifiers are:

20024, 20036, 20091, 20202, 20254, 20389, 20430, 20469, 20487, 20552, 20554, 20734, 20790, 20803, 20868, 20996, 21001, 21078, 21118, 21172, 21173, 21277, 21334, 21390, 21444, 21451, 21553, 21558, 21585



Medicines & Healthcare products
Regulatory Agency



These practices had more than 95% of their patients sharing combinations of registration date, sex, and year of birth with patients in another practice within the same region. No other pairs of practices share more than 70% of values. Together these practices contain slightly less than 1% of the individuals in CPRD Aurum, so we expect the impact of this duplication on previous analyses to be small.

We intend to resolve this issue in future builds, but older builds will remain unchanged in order to allow for replication of previous analyses.

CPRD Aurum Data Specification

Available on the CPRD website: <https://www.cprd.com/primary-care>