CPRD Aurum Glossary of terms/Data definitions

Acceptable flag

Patients are labelled as 'acceptable' for use in research by a process that identifies and excludes patients with non-continuous follow up or patients with poor data recording that raises suspicion as to the validity of that patient's record. Patient data is checked for the following issues:

- Year of birth is empty
- Current registration date is empty
- Current registration date is greater than the practice's last collection date
- Current registration date is less than or equal to 01/01/1900
- Current registration date is equal to or greater than the registration end date
- Current registration date is prior to the birth year
- Gender other than male, female or indeterminate
- Age is greater than 115 at end of follow-up (based on registration end date, death or last collection date)
- All recorded health care episodes have empty event dates
- All recorded health care episodes have invalid events dates (less than or equal to 01/01/1900 or greater than last collection date)
- All recorded health care episodes have dates before the birth year
- Patients are not permanently registered

If any of these conditions are true, then the patient is labelled unacceptable and is not recommended for use in research.

UTS date

In CPRD GOLD, the overall quality of data in practices is mediated by use of an 'up to standard' (UTS) date, which is deemed as the date at which data in the practice is considered to have continuous high-quality data fit for use in research.

This flag is currently under development for CPRD Aurum.

Death Recording

Information on date of death is included in the source data used for CPRD Aurum (emis_ddate), but this may not always correspond to the date of occurrence. For instance, it may reflect the date of notification of the death to the GP, or when the deceased patients' registration record was updated. CPRD therefore provide a more realistic estimate of the date of death (cprd_ddate), based on an algorithm which uses additional information in the patient record. The algorithm is similar to that used in CPRD GOLD, and the resulting estimate should be sufficiently accurate for purposes such as censoring follow-

up time. For studies where accurate date, and/or cause of death are important, use of linked ONS mortality data is recommended.