

Supplementary File S1. Inclusion and exclusion criteria

Inclusion criteria

The following inclusion criteria were applied:

- Being 18 years of age or older
- Subjects acceptable for research as per Clinical Practice Research Database (CPRD) definition
- Subjects eligible for Hospital Episode Statistics (HES) linkage
- Subjects registered in a general practitioner (GP) practice classified as “up to standard”.

Start of study eligibility was the latest of the following events:

- Start of study period (January 1st, 2000)
- HES admitted patient care (APC) set 14 coverage start date (April 1st, 1997)
- Turning 18 years old (based on the assumption of the birth date being the 1st January of the recorded year or birth in CPRD “yob” in Patient)
- Being registered in the CPRD-GOLD database for at least one year (current registration date “crd” in Patient + 365 days)
- Being registered in a GP practice classified as “up to standard” for at least one year (“uts” in Practice + 365 days).

The cohort inclusion (and start of follow up) date was the latest of date of registration in CPRD + 365 days (as described above) and the index date(s) of the T2DM and/or HBV-infection exposure(s) that needed to be present to enter the cohort. For sub-cohorts including patients with NAFLD, NAFLD exposure influenced eligibility for inclusion but did not influence the cohort inclusion date.

The end of follow-up was defined as the earliest of the following events:

- End of study period (December 31, 2015)
- CPRD-GOLD January 2017 release date (December 31st, 2016)
- HES APC set 14 coverage end date (March 31st, 2016)
- Transfer out of the practice date (“tod” in Patient)
- Death date (“death” in Patient)
- Practice last data collection date (“lcd” in Practice).

Subjects contributed time-at-risk from their cohort inclusion date until the end of follow-up.

Exclusion criteria

The following exclusion criteria were applied:

- Subjects with indeterminate sex (“gender”>2)
- Subjects for which the HES records are linked with more than one CPRD patient (“n_patid_hes”>1)
- Subjects for which end of follow-up precedes cohort inclusion date
- Subjects with Type 1 or unclassifiable diabetes
- Subjects with any missing event dates for diabetes-related diagnoses, treatments or “abnormal” tests
- Subjects with exactly one diabetes-related treatment prescription and no diabetes-related diagnoses or “abnormal” tests
- Subjects with any missing event dates for confirmatory (or A703.00) HBV Read code, International Classification of Disease (ICD) code, or HBV test “present”
- Subjects who pass through the entire HBV classification algorithm including step 13 without being classified

- Subjects with the following liver diseases **prior to cohort inclusion date** or with missing event date: liver transplantation, liver cancer, liver failure
- Subjects with other liver diseases, **anytime during their entire medical record** including: alcoholic liver disease, autoimmune liver disease, any viral hepatitis other than HBV, hemochromatosis.