

Diabetes Cohort

- Any Diabetes codes OR
- At least 6 months of Insulin data OR
- At least 2 consecutive high HbA1c

Step 1
Any Diabetes Other codes

YES

**Diabetes
Other**

N = 31,215 (0.43%)

NO

Step 2
Prescribing data coverage
available
*Not available if died before
prescribing coverage starts*

NO

YES

Step 3
Not currently on Insulin
*No prescription within the last
6 months prior to last
observable date*
AND
>3 years from diagnosis
to last observable date

YES

Step 3.1
Any Diabetes Type 2 codes

YES

NO

NO

Step 4
Diabetes Type 1 codes AND
NO Diabetes Type 2 codes

YES

**Diabetes
Type 1**
N = 400,150 (5.50%)

NO

Step 5
Diabetes Type 2 codes AND
NO Diabetes Type 1 codes

YES

NO

Step 6
Diabetes Type 1 codes AND
Diabetes Type 2 codes

YES

NO

YES

Step 7
Diagnosed age <35 years
AND
on Insulin within 1 year of
diagnosis

NO

Step 8
Any Diabetes NOS codes

YES

NO

Step 9
>= 2 of the following:

- Antidiabetic/
Metformin
medication
prescription ever
- >= 2 consecutive high
HbA1c
- Insulin prescribed
ever

YES

**Diabetes
Not Otherwise Specified
(NOS)**
N = 236,790 (3.25%)

NO

**Diabetes
Unlikely**
N = 116,070 (1.59%)

Notes

- Diabetes codes are derived from primary (SNOMED) and secondary (ICD 10) care.
- To be classified as Type 1 or Type 2 a person must have at least 1 Diabetes Code.
- If a person died prior to prescribing and HbA1c coverage they must have at least 1 Diabetes Code to be classified as Type 1 or Type 2 or NOS, otherwise they are classified as Unlikely as there is insufficient supporting data.
- A person does not need a diabetes code to be classified as NOS.
- Projects can choose to classify NOS with Type 2 to create a broader Type 2 categorization.

Definitions

Date of Diagnosis

- Defined as the date of a person's first diabetes code (of any type) or the first high HbA1c if more than a year before first diabetes code.

Last observable date

- For patients who are alive, this will be the current archived_on date of primary dataset, secondary dataset and prescribing dataset minus the max coverage lag across these datasets (likely 3 or 4 months).
- For patients who are dead this will be their date of death.