

==== Concerning the dataset “hxx.tsv” ====

==== Background ====

This is a material from a Swedish primary health care centre (vårdcentral) with details about 2886 patients with diabetes mellitus type 2.

The goal of treating diabetes type 2 (apart from alleviating acute effects of high blood glucose, such as dry mouth and thirst) is to prevent long-term cardiovascular effects, notably myocardial infarction and stroke, also blindness and renal failure. The mean level of blood glucose correlates fairly well with the risk of developing these diseases, and the level of glycosylated haemoglobin in blood (B-HbA1c) reflects the blood glucose level over the previous few months. For this reason, HbA1c is a measure of major interest in diabetes care, and a level >48 is considered valid for diagnosing diabetes type 2. In general terms, the lower the HbA1c, the better, but depending on the age and possible dangers of lowering the blood sugar excessively, higher levels of HbA1c (e.g. 60–70) are sometimes tolerated in frail patients with short life expectancy.

The Covid19 pandemic meant that many older people led a more passive life, which is likely to worsen diabetes type 2. We wanted to know if the HbA1c changed to any interesting degree (defined as >5% from previous levels) and, if so, this was more pronounced in any particular group of patients defined by age and sex. Also, was any change dependent on previous trends in a patient’s HbA1c, so that those with already poor (high) levels fared worse?

==== Data export and cleaning ====

The dataset was exported in 6-monthly intervals, using the data extraction tool “MedRave”, by selecting patients with a diagnosis of diabetes type 2 and exporting the following variables:

- Namn = name
- Personnr = personnummer
- (OLD)B-HbA1c = latest HbA1c during this period, with date and name of staff ordering test

The 6-monthly exports were joined in a spreadsheet where a variable for the time period was added:

- Källa = semester (vt2015 = first half of 2015; ht2015 = second half, etc)

Variables for HbA1c value and date were extracted, as well as derivations of personnummer: year of birth, sex.

See “material_dump1_anon.png” for illustration of the dataset at this point.

Some patients had temporary or unavailable personnummer, these were excluded, reducing the number of lines from 29 704 to 28 812. Next, the personnummer was replaced with a 6-letter code, e.g. “WPVQNB”.

==== Variables ====

In the resulting dataset, “hxx.tsv”, the following variables remain:

- kod : Code replacing personnummer, levels “MLNJEH”,..
- kon : Sex, levels “K”, “M”
- yob : Year of birth, 1919–2003
- term : Semester, levels “VT2015”,..“HT2020”
- hba1c : HbA1c, 22–151
- hba1cdt: Date for HbA1c, 2015-01-07–2020-12-30

=== Work ===

This project would be done in collaboration with Staffan Svensson (University of Gothenburg). A preliminary analysis (comparison of means and variances for different periods) can be found at <https://pharm.nu/hamid/>.