

EXPoSE

Explaining **Po**population trends in cardiovascular risk: A comparative analysis of health transitions in **S**outh Africa and **E**ngland

Harmonisation Code

This folder contains the code used within the scope of the ExPoSE Project (www.exposeproject.net) to import and harmonise microdata on cardiovascular risk factors and other information from a set of population surveys conducted in South Africa and England between 1997 and 2017.

The harmonisation code for the England datasets (partly in Stata™ and partly in SPSS™ format) is included in the folder EXTRACT_EN. The main files are EXTRACT_EN_1.SPS and EXTRACT_EN_2.do, which must be run in sequence and include instructions on how to set up the directory structure for the input data. The subfolders TEMP and OUT are initially empty and are used by the code to store intermediate and final results, respectively.

The harmonisation code for the South African datasets (in Stata™ format) is included in the folder EXTRACT_SA. The main file is EXTRACT_SA.do, which includes instructions on how to set up the directory structure for the input data. The subfolder AUXILIARY includes geographic and demographic information needed to complete the process. The subfolders TEMP and OUT are initially empty and are used by the code to store intermediate and final results, respectively.

The R file CONSOLIDATE.R imports the South African and English extracted data, reorders and renames the variables consistently, calculates cardiovascular risk scores for individuals aged 40 to 74 years (GloboRisk, WHO and Framingham) and recalibrates sampling weights for the South African dataset.

The ExPoSE project was funded by:



**Economic
and Social
Research Council**