README file for MIDUS Polygenic Risk Scores (P6)
November 29, 2021

The purpose of this memo is to provide basic information about the Polygenic Risk Scores (PRS) generated through the MIDUS Genomics project (P6) for participants that completed the MIDUS 2 or MIDUS Refresher Biomarker project (P4). This data and related documentation are publicly available via the MIDUS Colectica Portal, but are not currently available through ICPSR.

Data and documentation for other MIDUS projects are available through:
1. The public archive at ICPSR (http://www.icpsr.umich.edu/icpsrweb/NACDA/studies/36532/version/2)
2. The MIDUS Colectica Portal (http://midus.colectica.org/) which houses interactive codebooks for all the publicly available MIDUS projects. The Portal also includes search and explore functions, links to documentation, and a custom download function.

A. What data files are available?
A single data file containing PRS data for both M2 and MR participants has been generated and is available through the Portal as an SPSS dataset:

> ### M2MR_P6_PolygenicRiskScores_N2118_20211028.sav

B. What is the structure of the MIDUS PRScore Dataset?
The PRS datafile is a SPSS (ver. 26) dataset comprised of Polygenic Risk Scores along with relevant administrative variable and principal component and ancestry scores. The file contains 65 variables for 2118 cases that completed either the MIDUS 2 or the MIDUS Refresher Biomarker project (P4). See below for details about the ID system.

The dataset has been cleaned, meaning that value ranges and skip patterns have been checked. Variables have been named according to the MIDUS naming conventions. All variables include labels to aid interpretation. Value labels have been applied where appropriate and discrete missing values have been defined.

The following documents provide additional information about the data:

> ***M2MR_P6_PolygenicRiskScores_Documentation_20211129:*** provides general information about the Polygenic Risks Scores dataset along with technical and other details about how the scores are generated.

> ***MR_M3_Naming and Coding Conventions:*** is posted with the MIDUS Refresher Survey (Project 1) data and describes conventions for naming and coding variables. It is included with the MR Project 1 documentation in the Portal or at ICPSR: (http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/203/studies/36532?archive=ICPSR&sortBy=7

C. Other important information.

November 2021 UPDATE.

In September 2021 an additional individual PRS for Alzheimer's disease was computed using default parameters in PRS-CS. This new variable, BRA6DPRSCSALZ, is included in the above data file and computation details are included in the above documentation.

IMPORTANT ID system. The MIDUS PRScore dataset contains cases from both the MIDUS 2 and MIDUS Refresher samples, thus it uses a special 5-digit respondent identification variable called 'M2MRID' that combines the 'M2ID' and 'MRID' variable values for respective M2 and MR cases. The data file also includes a variable called 'M2MRCASE' that identifies whether the case is from the MIDUS 2 or the MIDUS Refresher sample. For all M2 cases, the value of the 'M2MRID' variable is identical to 'M2ID' and can be used to link to other MIDUS 1, 2, and 3 datasets. Similarly, for all MR cases the value of the 'M2MRID' variable is identical to 'MRID' and can be used to link to other MIDUS Refresher datasets.

****

Please report any errors or inconsistencies you find in the data or documentation to
glove@wisc.edu or bradler@wisc.edu

****