README for MIDUS DNA Methylation Age Data
August, 2023

The purpose of this memo is to provide basic information about the DNA Methylation Age data generated through the MIDUS Genomics project for participants that completed the MIDUS 2 or MIDUS Refresher 1 Biomarker project. 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 DNA Methylation Age Scores for both M2 and MR1 participants has been generated and is available through the MIDUS Colectica Portal as an SPSS dataset:

M2MR1_GEN_DNAmAge_N2118_20230822.sav

B. What is the structure of the MIDUS Methylation Age Dataset?
The data file is a SPSS dataset comprised of DNA Methylation Epigenetic Age scores along with relevant administrative variable and quality control metrics. The file contains 21 variables for 2,118 cases that completed either the MIDUS 2 or the MIDUS Refresher 1 Biomarker project. Among the 2,118 cases in this data file, Epigenetic Age score are available for 1,310 cases.

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:

> **M2MR1_GEN_DNAmAge_Documentation_20230822:** provides general information about the Epigenetic Age 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 MR1 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.

IMPORTANT **ID system.** The Methylation Age dataset contains cases from both the MIDUS 2 and MIDUS Refresher1 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 MR1 cases. 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 MR1 cases the value of the 'M2MRID' variable is identical to 'MRID' and can be used to link to other MIDUS Refresher1 datasets. The data file also includes a variable called 'M2MRCASE' that identifies whether the case is from the MIDUS 2 or the MIDUS Refresher 1 sample.