

## **Analysis code**

# **Epigenetic Biomarkers of Socioeconomic Status are Associated with Age-Related Chronic Diseases and Mortality in Older Adults**

**May 2024**

Lauren L. Schmitz,<sup>1\*</sup> Lauren A. Opsasnick,<sup>2</sup> Scott M. Ratliff,<sup>2</sup> Jessica D. Faul,<sup>3</sup> Wei Zhao,<sup>2,3</sup> Timothy M. Hughes,<sup>4</sup> Jingzhong Ding,<sup>4</sup> Yongmei Liu,<sup>5</sup> and Jennifer A. Smith<sup>2,3\*</sup>

<sup>1</sup> Robert M. La Follette School of Public Affairs, University of Wisconsin-Madison, Madison, WI, USA

<sup>2</sup> Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

<sup>3</sup> Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

<sup>4</sup> Department of Gerontology and Geriatric Medicine, School of Medicine, Wake Forest University, Winston-Salem, NC, USA

<sup>5</sup> Department of Medicine, Divisions of Cardiology and Neurology, Duke University Medical Center, Durham, NC, USA

\*Corresponding Authors: Lauren L. Schmitz and Jennifer A. Smith

**Email:** llschmitz@wisc.edu, smjenn@umich.edu

# README

The following R, SAS, and bash codes were used for analysis in this manuscript. Each code file is annotated, but file and path names may change depending on data sources.

## 1. **HRS\_phenos.sas**

This is the main code for phenotype creation and processing of HRS data. It also includes code for descriptive statistics, SES biomarker/health outcome association models, and table creation.

## 2. **HRS\_socenv.R**

Along with HRS\_phenos.sas, creates the neighborhood social environment score, one of six measures included in the SES index.

## 3. **HRS\_socialepi\_EWAS\_row\_col\_plate.R**

This code performs general EWAS across multiple outcomes and strata. It was used for EWAS of the SES indices and their component parts.

## 4. **EWAS\_Rbatch.bash**

Bash code calling HRS\_socialepi\_EWAS\_row\_col\_plate.R for parallel processing of multiple outcomes in batches of 100,000 CpG sites.

## 5. **EWAS\_post\_processing.R**

Performs post-processing of EWAS results, including combining CpG batches and creating QQ and Manhattan plots.

## 6. **HRS\_methy\_resids.R**

This code residualizes the HRS methylation data for input to the elastic net models.

## 7. **resids.bash**

Bash code calling HRS\_methy\_resids.R for parallel processing in batches of 100,000 CpG sites.

## 8. **HRS\_ElasticNet\_XValid.R**

This code performs the cross-validated elastic net analysis and creates the SES biomarkers for both the EPIC chip and 450K subset in HRS and MESA.

## 9. **genome\_feature\_and\_gene\_set\_enrichment\_analysis.R**

This code performs the genomic feature and gene set enrichment analysis for the 850k and 450k adult and childhood SES biomarkers.

## 9. **MESA\_phenos.sas**

This is the main code for phenotype creation and processing of MESA data. It also includes code for descriptive statistics.

#### **10. MESA\_mclock\_models.sas**

This code performs the SES biomarker/health outcome association models in MESA. It also includes code for descriptive statistics and for analyzing the biomarker predictiveness in MESA.

#### **11. forest\_plots.R**

Creates forest plots for Figures 2 and 3.