**JASA ACS Reproducibility Initiative - Author Contributions Checklist Form**

The purpose of the Author Contributions Checklist (ACC) Form is to document the artifacts associated with a manuscript (e.g., code and data supporting the computational findings), and describe how to reproduce the findings. The final version of this document will be included as online supplemental material with the published paper and referenced in the abstract.

As of Sept. 1, 2016, the ACC Form must be included with all submissions to JASA ACS.

This document is the template that will be provided to authors; please replace the (non-bold) text below that provides guidance on how to fill out each item with the actual information for your manuscript.

## Data

**Abstract (Mandatory)**

The R scripts used to simulate the data for the simulation studies in Section 4 are provided along with the code for each of the methods being compared. The data used for the two real data analyses in Section 5 are publicly available. The first dataset from the Nurse Family Partnership program is provided by one of the authors of the manuscript (David Olds). The second dataset from the Study to Understand Prognoses Preferences Outcomes and Risks of Treatment (SUPPORT) is publicly available from the Vanderbilt University Department of Biostatistics website.

**Availability (Mandatory)**

Both real datasets have been uploaded to my GitHub repository which also hosts the R package. They can be downloaded directly from the public GitHub repository.

**Description (Mandatory if data available)**

The datasets are available at <https://github.com/sahirbhatnagar/sail/tree/jasa/manuscript/raw_data>

1. Nurse Family Partnership program data consists of three files. They are merged together using the script <https://github.com/sahirbhatnagar/sail/blob/jasa/manuscript/bin/PRS_bootstrap.R>

* [Gen\_3PC\_scores.txt](https://github.com/sahirbhatnagar/sail/blob/jasa/manuscript/raw_data/Gen_3PC_scores.txt)
* [IQ\_and\_mental\_development\_variables\_for\_Sahir\_with\_study\_ID.txt](https://github.com/sahirbhatnagar/sail/blob/jasa/manuscript/raw_data/IQ_and_mental_development_variables_for_Sahir_with_study_ID.txt)
* [NFP\_170614\_INFO08\_nodup\_hard09\_noambi\_GWAS\_EduYears\_Pooled\_beta\_withaf\_5000pruned\_noambi\_16Jan2018.score](https://github.com/sahirbhatnagar/sail/blob/jasa/manuscript/raw_data/NFP_170614_INFO08_nodup_hard09_noambi_GWAS_EduYears_Pooled_beta_withaf_5000pruned_noambi_16Jan2018.score)

2. The SUPPORT data consists of a single file:

* <https://github.com/sahirbhatnagar/sail/blob/jasa/manuscript/raw_data/support2.csv>

All datasets are in .txt format. Code used to read in the datasets are provided in the section below. All output from this project published online is available according to the conditions of the Creative Commons License (https://creativecommons.org/licenses/by-nc-sa/2.0/)

Code

**Abstract (Mandatory)**

The software which implements our algorithm is available in an R package published on CRAN. The paper itself is written in knitr format, and therefore includes both the code and text in the same .Rnw file.

**Description (Mandatory)**

The R package implementing our method is available from CRAN at <https://cran.r-project.org/package=sail>. Package version 0.1.0 with MIT license.

The scripts and data used to produce the results in the manuscript are available at <https://github.com/sahirbhatnagar/sail/tree/jasa/manuscript>.This is the jasa branch with commit number 7be7ecd1b5404796a88936026bf680a7e2529798.

The knitr file which contains both the main text and code is available at: <https://github.com/sahirbhatnagar/sail/blob/jasa/manuscript/source/sail_manuscript_v2.Rnw>

The manuscript was compiled using R version 3.6.1 with knitr version 1.25.

**Optional Information (complete as necessary)**

* The bootstrap analysis was run in parallel on a compute cluster with 40 cores. Though this is not necessary to reproduce the results, it definitely speeds up the computation time.

## Instructions for Use

**Reproducibility (Mandatory)**

All tables and figures from the paper can be reproduced by compiling the knitr file. The easiest way to reproduce the results is to download the GitHub repository and compile the knitr file from within an R session as follows:

1. Download the GitHub repository <https://github.com/sahirbhatnagar/sail/archive/jasa.zip>
2. From within an R session, run the command:

knitr::knit2pdf(‘sail\_manuscript\_v2.Rnw’)

Note that to speed up compilation time, we have saved the simulation and bootstrap results in .RData files available at <https://github.com/sahirbhatnagar/sail/tree/jasa/manuscript/results>. These .RData files are called directly by the knitr file.

Note also that the R scripts used to generate the results are called from the knitr file using the ‘code externalization’ functionality of knitr (<https://yihui.org/knitr/demo/externalization/>). That is, the actual R code is stored in R scripts and not within the knitr file. These R scripts are available at <https://github.com/sahirbhatnagar/sail/tree/jasa/manuscript/bin>.

The expected run time to compile the manuscript is about 5 minutes on a standard desktop machine, assuming that you are using the pre-run simulation and bootstrap results.

**Replication (Optional)**

A website with two vignettes has been created for our sail package available at <https://sahirbhatnagar.com/sail/>

The 2 vignettes are:

1. <https://sahirbhatnagar.com/sail/articles/introduction-to-sail.html>
2. <https://sahirbhatnagar.com/sail/articles/user-defined-design.html>

## Notes

None.