**PrePost\_MRI.R\_README**

Robert Toms, robert.toms@utdallas.edu

README SECTIONS:

1. Purpose of Code
2. Researcher-facing Explanation of Code
3. Explanation of Variables
4. Methods Section Summary
5. References

PURPOSE OF CODE:

The purpose of this code is to match Pre & Post menarche timepoints for each subject with available neuroimaging data.

RESEARCHER-FACING EXPLANATION OF CODE:

1. Derivative ABCD 6.0 menarche data were imported for youth (Youth\_PrePost\_Menarche\_9\_15\_25.csv) and parent (Parent\_PrePost\_Menarche\_9\_15\_25.csv) reports. These were used rather than raw data because they contain predetermined variables for when participants switched from premenarche to postmenarche (last\_pre\_session, first\_post\_session). Additionally, MRI QC data (mr\_y\_qc\_\_incl.tsv) and interview date data (ab\_g\_dyn.tsv) were imported.
2. Clean and Prepare Data
   1. Useful columns were kept and renamed for human readability, if necessary.
   2. Timepoints lacking any usable T1, rsfMRI, and nBack data were removed.
      1. All 3 must be missing ( = 0) to be removed:
         1. mr\_y\_qc\_\_incl\_smri\_t1\_indicator = 0
         2. mr\_y\_qc\_incl\_rsfmri\_indicator = 0
         3. mr\_y\_qc\_incl\_tfmri\_nback\_indicator = 0
      2. n = 789 timepoints removed
3. Make Human-readable list vector of timepoints with available scan data
   1. This is done for consolidation purposes, and so that the list vectors can be algorithmically searched through to determine which scans best align with premenarche and postmenarche reports. They are saved to “has\_scan”.
   2. e.g. has\_scan = c(“0”, “2”, “6”) ; has scans for ses-00A, ses-02A, and ses-06A
4. Simplify and merge Menarche and MRI/Dates data
   1. Only useful columns are kept for mri and menarche data
   2. List vectors are constructed of menarche reports, in chronological order. This is done for consolidation purposes, and for human readability. They will not need to be algorithmically searched through in this code. They are saved to “menarche\_reports”.
      1. e.g. menarche\_reports = c(“0”, “0”, NA, “0”, “1”, “1”, “1”) =

c(Pre, Pre, NA, Pre, Post, Post, Post)

* 1. MRI data and scan data are merged into parent\_img and youth\_img (parent and youth menarche reports, respectively)

1. Determine Pre/Post Menarche Scan Timepoints
   1. Wrote 2 functions:
      1. find\_premenarche\_scan: finds the index in the has\_scan list vector (1-4) with the session (with scan data) that matches or most closely precedes a participant’s last premenarche report, saves it to PreScanIndex
         1. Example
            1. last\_pre\_session = 3
            2. has\_scan = c(“0”, “4”, “6”)
            3. finds “0” value, 1st item in vector, PreScanIndex = 1
      2. find\_postmenarche\_scan: finds the index in the has\_scan list vector (1-4) session (with scan data) that matches or most closely follows a participant’s first postmenarche report, saves it to PostScanIndex
         1. Example
            1. first\_post\_session = 4
            2. has\_scan = c(“0”, “4”, “6”)
            3. finds “4” value, 2nd item in vector, PostScanIndex = 2
   2. Applied the find\_premenarche\_scan & find\_postmenarche\_scan functions to youth and parent reports
   3. Used the indexes (PreScanIndex, PostScanIndex) to identify the timepoints with scan data most closely associated with premenarche and postmenarche reports
      1. Example
         1. has\_scan = c(“0”, “4”, “6”)
         2. PreScanIndex = 1, 1st item is “0”, ses-00A has the best-fit premenarche scan
         3. PostScanIndex = 2, 2nd item is “4”, ses-04A has the best-fit postmenarche scan
2. Calculate Days between Data collection and/or days between menarche report and data collection
   1. Interview date data were pivoted wide and merged with menarche and imaging data, making new dataframes “parent\_img\_dates” and “youth\_img\_dates”
   2. A list vector was made of the interview dates in chronological order, so they could be systematically pulled and subtracted from each other to determine the difference in days. This was saved to “dates”.
   3. Calculating Date Differences
      1. Number of days between best-fit premenarche and postmenarche scans, saved to “between\_scans”
      2. Number of days between last premenarche and first postmenarche reports, saved to “PreMen->PostMen”
      3. Number of days from best-fit premenarche Scan to last premenarche report, saved to “PreMenScan->PreMenReport”
      4. Number of days from first postmenarche report to best-fit postmenarche scan, saved to “PostMenReport->PostMenScan”
3. Clean and Reorder dataframes for CSV export
   1. Itemized interview dates and itemized menarche reports were renamed to delineate between them, they will be kept for reference
   2. has\_scan was converted to a string to retain readability while affording the ability to be written into a csv
   3. Necessary columns are kept and reordered
4. Write to CSV
   1. parent\_img\_dates -> Parent\_PrePost\_MRI\_Menarche\_9\_16\_25.csv
   2. youth\_img\_dates -> Youth\_PrePost\_MRI\_Menarche\_9\_16\_25.csv

EXPLANATION OF VARIABLES

|  |  |
| --- | --- |
| participant\_id | ABCD Participant Identifier |
| session\_id | ABCD Visit Identifier |
| mr\_y\_qc\_\_incl\_smri\_t1\_indicator | MRI QC data for T1 scans. Usable = 1, unusable = 0 |
| mr\_y\_qc\_incl\_rsfmri\_indicator | MRI QC data for resting-state fMRI scans. Usable = 1, unusable = 0 |
| mr\_y\_qc\_incl\_tfmri\_nback\_indicator | MRI QC data for nBack scans. Usable = 1, unusable = 0 |
| ab\_g\_dyn\_\_visit\_dtt  interview\_date | Datetime for interview |
| session\_num | Numeric version of session\_id |
| has\_scan | List vector of session numbers with usable scan data; e.g. c(“0”, “4”, “6”) = usable ses-00A, ses-04A, ses-06A scans |
| menarche\_reports | List vector of menarche reports: “0” = Premenarche, “1” = Postmenarche;  e.g. c(“0”, “0”, NA, “0”, “1”, “1”, “1”) =  c(Pre, Pre, NA, Pre, Post, Post, Post) |
| last\_pre\_session | Numeric version of session with last premenarche report |
| first\_post\_session | Numeric version of session with first postmenarche report |
| PostMenarche\_at\_Baseline\_Y1N0 | Flags everyone who was postmenarche at first visit, Yes = 1 and No = 0 |
| PreMenarche\_at\_LastReport\_Y1N0 | Flags everyone who was premenarche through 6.0, Yes = 1 and No = 0 |
| Inconsistent\_Reporting\_Y1N0 | Flags everyone who reported being premenarche after reporting bein postmenarche  e.g. c(“0”, “0”, NA, “1”, **“0”,** “1”, “1”) = c(Pre, Pre, NA, Post, **Pre**, Post, Post) |
| find\_premenarche\_scan | finds the index in the has\_scan list vector (1-4) with the session (with scan data) that matches or most closely precedes a participant’s last premenarche report, saves it to PreScanIndex |
| find\_postmenarche\_scan | finds the index in the has\_scan list vector (1-4) session (with scan data) that matches or most closely follows a participant’s first postmenarche report, saves it to PostScanIndex |
| PreScanIndex | Location in has\_scan list vector of the value that matches or is the best-fit for last premenarche report |
| PostScanIndex | Location in has\_scan list vector of the value that matches or is the best-fit for last postmenarche report |
| PreScan | PreMenarche Scan: Numeric version of the session with scan data best-fit matched to last premenarche report |
| PostScan | PostMenarche Scan: Numeric version of the session with scan data best-fit matched to first postmenarche report |
| dates | List vector of each participant’s interview dates |
| between\_scans | Number of days between PreMenarche scan (PreScan) and PostMenarche scan (PostScan) |
| PreMen->PostMen | Number of days between PreMenarche report and PostMenarche report |
| PreMenScan->PreMenReport | Number of days from PreMenarche scan to PreMenarche report |
| PostMenReport->PostMenScan | Number of days from PostMenarche report to PostMenarche scan |

Methods Section Summary:

To facilitate difference calculations of hippocampal volumes from premenarche to postmenarche, we compared ABCD MRI quality control data (mr\_y\_qc\_\_incl.tsv) to predetermined last premenarche reports and first postmenarche reports to find matching sessions or sessions of best-fit. To accomplish this for each participant, lists of sessions with usable MRI data were compiled. Scans matching and/or preceding the participant’s last premenarche report were separated out, and the latest was deemed the “best-fit” premenarche scan. Similarly, scans matching and/or following the participant’s first postmenarche report were separated out, and the first of these was deemed the “best-fit” postmenarche scan.

References: