**PrePost\_Menarche\_TPs.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 find Pre & Post menarche timepoints for each subject, for both parent and youth reports.

RESEARCHER-FACING EXPLANATION OF CODE:

1. ABCD 6.0 menarche data were imported for youth (ph\_y\_pds.tsv) and parent (ph\_p\_pds.tsv) reports.
2. All columns of interest are enumerated below in the Explanation of Variables section, in order of appearance.
3. Clean and Prepare Data
   * 1. Unnecessary columns and males were removed
     2. Ambiguous responses (999 = I don't know; 777= refuse to answer) were set to NA (in line with recommendations in Herting et al., 2021)
     3. Youth and parent dataframes were pivoted wide so each subject would be in a single row, to facilitate the next step.
     4. A vector called menarche\_reports listing Pre/Post menarche reports were constructed using the timepoint-specific menarche self-report variables **ph\_y\_pds\_\_f\_002** and **ph\_p\_pds\_\_f\_002** (youth & parent reports, respectively).
        1. Have you/your child begun to menstruate (started to have your/their period)? No = 0; Yes = 1; Decline to Answer = 777, I don't know = 999
        2. e.g. c(“0”, “0”, NA, “1”, “1” , “1”, “1”) = c(Pre, Pre, NA, Post, Post, Post, Post)
4. Finding Last Premenarche and First Postmenarche Timepoints
   * 1. Last Premenarche: we iterated through our list vectors and found the last occurrence of a “0”, then converted the location in the vector (1-7) of that “0” to a human-readable format (e.g. ses-00A = 1, ses-01A = 2, ses-02A = 3, etc.). These were then reduced by 1 to match the timepoint count (ses-00A = 0, ses-01A = 1, etc.) and saved to last\_pre\_session.
     2. First Postmenarche: we iterated through our list vectors and found the first occurrence of a “1”, then converted the location in the vector (1-7) of that “1” to a human-readable format (e.g. ses-00A = 1, ses-01A = 2, ses-02A = 3, etc.). These were then reduced by 1 to match the timepoint count (ses-00A = 0, ses-01A = 1, etc.) and saved to first\_post\_session.
5. Flag for PreMenarche at Last report, Postmenarche at first report, and Inconsistent reporting (premenarche after postmenarche)
   * 1. PostMenarche at First Report: If no premenarche reports, postmenarche at all reports (including first) are assumed [PostMenarche\_at\_Baseline\_Y1N0]
     2. PreMenarche at Last Report: If no postmenarche reports, premenarche at all reports (including last) are assumed [PreMenarche\_at\_LastReport\_Y1N0]
     3. Inconsistent Reporting: If premenarche status is reported after a postmenarche report (0 after a 1), this is considered Inconsistent reporting
        1. e.g. c(“0”, “0”, NA, “1”, “0”, “1”, “1”) = c(Pre, Pre, NA, Post, Pre, Post, Post)
6. Find adjusted premenarche timepoints for Inconsistent reporters
   1. Theoretical justification: In these cases, some young women may experience irregular cycling during menarche onset, where bleeding occurs sporadically, leading to confusion as to whether menarche has truly occurred. It may be the case that the onset of this irregular cycling is the result of the cascading biological effects that lead to menarche onset, thus—though irregular—these may potentially be treated as menarche events. Hence, analysis with and without to check for effects. In such cases it is likely that the large organizational effects of rising estradiol would have still occurred with the first menarche even if the acute changes related to menstrual cycle fluctuations do not occur. If so, then we would expect no differences in organizational effects of estradiol from the consistent versus inconsistent reporters.
   2. Defined a function to find the premenarche timepoint most closely preceding the first postmenarche timepoint – “find\_premenarche”
   3. Applied the function to youth and parent, saved corrected last premenarche timepoint to “Inconsistent\_Pre”
   4. Replaced the last\_pre\_session with the corrected “Inconsistent\_Pre” session number for those with inconsistent reports

1. Clean for Export
   1. Kept and reordered useful columns
2. Data are then exported to Parent and Youth-specific csv’s.
   1. youth  ->  Youth\_PrePost\_Menarche\_9\_15\_25.csv
   2. parent  ->  Parent\_PrePost\_Menarche\_9\_15\_25.csv

EXPLANATION OF VARIABLES

|  |  |
| --- | --- |
| participant\_id | ABCD Participant ID |
| session\_id | ABCD Visit ID |
| ph\_y\_pds\_\_f\_002 | Menarche status (youth report)  Have you begun to menstruate (started to have your period)? No = 0; Yes = 1; Decline to Answer = 777, I don't know = 999 |
| ph\_p\_pds\_\_f\_002 | Menarche status (parent report)  Has your child begun to menstruate? No = 0; Yes = 1; Decline to Answer = 777, I don't know = 999 |
| ses-00A | Menarche status report for ses-00A |
| ses-01A | Menarche status report for ses-01A |
| ses-02A | Menarche status report for ses-02A |
| ses-03A | Menarche status report for ses-03A |
| ses-04A | Menarche status report for ses-04A |
| ses-05A | Menarche status report for ses-05A |
| ses-06A | Menarche status report for ses-06A |
| menarche\_reports | List vector of menarche reports, in chronological order;  e.g. c(“0”, “0”, NA, “1”, “1” , “1”, “1”) = c(Pre, Pre, NA, Post, Post, Post, Post) |
| last\_pre\_index | Location in the vector (1-7) menarche\_reports of the last pre-menarche endorsement (“0”) |
| first\_post\_index | Location in the vector (1-7) menarche\_reports of the first post-menarche endorsement (“1”) |
| last\_pre\_session | Session of last premenarche report: last\_pre\_index – 1 |
| first\_post\_session | Session of first postmenarche report:  first\_post\_index - 1 |
| 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 | Function: finds closest previous premenarche report “0” prior to first postmenarche report “1”, applies to list vector menarche\_reports |
| Inconsistent\_Pre | Corrected version of last\_pre\_session for those with Inconsistent Reporting: made using find\_premenarche, used when Inconsistent\_Reporting\_Y1N0 = 1 |

Methods Section Summary:

To determine menarche status in terms of pre-menarche and post-menarche timepoints, the self-reported menarche status single items of the Pubertal Development Scale (PDS) (ph\_y\_pds\_\_f\_002, youth report; ph\_p\_pds\_\_f\_002, parent report; Have you/your child begun to menstruate (started to have your/their period)?; 1 = Yes, 0 = No, 999 = I don't know, 777= refuse to answer) were used. Data for this item are gathered annually. This marker was chosen because menarche is a categorical indication that estradiol levels are sufficient to cause cascading biological effects (Cheng et al., 2021, Damme et al., 2020), the use of which obviates each individual’s profound variability in sensitivity to estrogens (Shirtcliff et al., 2000).

For Supplement:

Following the best practices of Herting et al., “I don’t know” and “Refuse to answer” responses were considered missing (Herting et al., 2021). For each adolescent, the timepoints associated with their last pre-menarche report (last “0” occurrence) and the first post-menarche report (first “1” occurrence) were accepted for their corresponding events.

Some adolescents reported inconsistently (a “0” following a “1”), reported being Post-menarche at first interview, or reported having not yet experienced menarche in the course of the study. Because our analyses of interest rely on menarche as a marker of transition into puberty, subjects who did not experience menarche during the course of data collection were excluded. Subjects whose occurrence of menarche could not be specified (due to irregular reporting) were flagged, so that effects of inclusion/exclusion could be assessed. In these cases, some young women may experience irregular cycling during menarche onset, where bleeding occurs sporadically, leading to confusion as to whether menarche has truly occurred (INSERT CITATION). It may be the case that the onset of this irregular cycling is the result of the cascading biological effects that lead to menarche onset, thus—though irregular—these can be treated as menarche events.

References:

Herting, M. M., Uban, K. A., Gonzalez, M. R., Baker, F. C., Kan, E. C., Thompson, W. K., Granger, D. A., Albaugh, M. D., Anokhin, A. P., Bagot, K. S., Banich, M. T., Barch, D. M., Baskin-Sommers, A., Breslin, F. J., Casey, B. J., Chaarani, B., Chang, L., Clark, D. B., Cloak, C. C., Constable, R. T., … Sowell, E. R. (2021). Correspondence Between Perceived Pubertal Development and Hormone Levels in 9-10 Year-Olds From the Adolescent Brain Cognitive Development Study. Frontiers in endocrinology, 11, 549928. https://doi.org/10.3389/fendo.2020.549928