A RANDOMIZED PLACEBO-CONTROLLED TRIAL OF PITAVASTIN CALCIUM TO TREAT COMBINED DYSLIPIDEMIA OF OBESITY IN ADOLESCENTS – THE PEDIATRIC HEART NETWORK DYSLEPIDEMIA OF OBESITY INTERVENTION IN TEENS (DO IT) TRIAL PUBLIC USE DATASET

Pediatric Heart Network (PHN)

ABOUT THE STUDY

The NHLBI-sponsored DO IT! Study was conducted by the Pediatric Heart Network (PHN), with enrollment at 15 North American centers.

The study objective was to assess the safety and efficacy of statin therapy on vascular and lipid measures in a population of youth affected by combined dyslipidemia of obesity (CDO). This was a 2-year double-blind trial of pitavastatin calcium 4 mg/day versus placebo for participants ages 10-19 years with body mass index (BMI) $\geq 85^{th}$ %ile and CDO defined as non-HDL-C $\geq$120 mg/dL and either low HDL-C or high triglyceride (TG):HDL-C ratio.

From July 2018 to April 2021, 59 participants were randomized to pitavastatin and 60 to placebo. Enrollment was stopped early due to slow recruitment. Subjects Pulse Wave Velocity (PWV), lipid profiles, and carotid stiffness were assessed at baseline, 1, 6, 12, 18 and 24 months.

Additional information about the DO IT! study can be found at https://www.pediatricheartnetwork.org/studies/do-it/

DATA AND DOCUMENTATION

The following datasets and descriptor files are available for download. A free login and password are required for download capability. The lock date used for creation of these public use datasets was June 30, 2023. Privacy protection of these data is described in Appendix A.

1. Study protocol
2. DOIT! published main results
3. Study data collection forms
4. SAS datasets
5. Excel datasets (with variable formats applied) – These data have a .csv extension, which means that the file may also be opened either in Excel, OR in a text editor, appearing as a comma-delimited file.
6. Codebooks for each dataset – These contain variable names, labels, and descriptive statistics for each variable on the data collection forms. Key created variables are

included as well. Categorical variables are listed first, with frequencies; continuous variables are listed next, with mean, SD, median, interquartile range, and range.

DATA USE POLICY

- REQUIRED ACKNOWLEDGEMENTS: All presentations and publications using these data must include the following statement: *"The NIH/NHLBI Pediatric Heart Network DO IT! dataset was used in preparation of this work. Data were downloaded from www.pediatricheartnetwork.org on mm/dd/yyyy"."*
- PAPER, ABSTRACT, and PRESENTATION TITLES: Titles may, at the authors' discretion, mention the PHN database but should not imply that the work is from the PHN. An example of an acceptable phrase would be, "an analysis of the Pediatric Heart Network public database." Whether or not the title makes mention of the PHN, acknowledgement should be made as described above.
- All users are requested to send a copy of published abstracts and articles to the PHN Data Coordinating Center at HealthCore, Inc (PHNpubs@healthcore.com) within one month of publication. This will allow the PHN and the NHLBI to document the continued impact of this study on the field.
- The login and password to access the public dataset is provided to a single user. If a colleague would like to access the public dataset for a different analysis topic, a separate request for login and password should be submitted via the www.pediatricheartnetwork.org website.
- As an approved user, you agree that you will not attempt to establish the identities of research participants through use of this dataset.
- As an approved user, you agree to not place these data in other public locations.

A RANDOMIZED PLACEBO-CONTROLLED TRIAL OF PITAVASTIN CALCIUM TO TREAT COMBINED DYSLIPIDEMIA OF OBESITY IN ADOLESCENTS – THE PEDIATRIC HEART NETWORK DYSLEPIDEMIA OF OBESITY INTERVENTION IN TEENS (DO IT) TRIAL PUBLIC USE DATASET

TIPS ON USING THESE DATA

- Identification numbers for study participants have been re-assigned for privacy protection. A blinded ID ranging from 1 to 119 (number of participants in randomization dataset) has been assigned.

- The study data are contained in a large number of individual forms. These may be used jointly by merging on *blind_id*. Sometimes it may be necessary to merge on both *blind_id* and *visit*.

- A blinded site variable has been included. Participants with the same blinded site ID are from the same site, but sites are not identified by name or PHN site number.

- An analytical dataset of variables created for the study is included.

- A dataset with calculated SES measures is included:

    - Childhood Opportunity Index (COI)
        - https://data.diversitydatakids.org/dataset?vocab_Topic=Child+Opportunity+Index&_ga=2.79944727.1853358601.1749737593-419558661.1749737593
        - Higher COI indicates higher SES
        - Both national and state-normed COI scores are available
        - The v3.0 2020 census tract COI dataset was used
        - For 11 participants, there was no match, so the 2010 tract COI dataset was used
        - For 4 participants the v2.0 with 2015 data aggregated to the 2020 census zipcodes COI dataset needed to be used, as census tracts were missing. However, the zipcode-level COI does not include a numeric format score for the state-normed overall COI score.

    - Environmental Justice Index (EJI)
        - https://www.atsdr.cdc.gov/place-health/php/eji/eji-data-download.html
        - Lower EJI indicates higher SES
        - Modules of the EJI:
            - Social vulnerability module (SVM)
            - Environmental burden module (EBM)
            - Health vulnerability module (HVM)
            - Climate burden module (CBM)
        - Each index is presented in both percentile and score formats. A percentile ranking represents the proportion of census tracts (or communities) that

may experience more or less severe cumulative impacts from environmental burdens than the community of interest. For example, a EJI ranking of 0.75 signifies that 75% of tracts in the nation likely experience less severe cumulative impacts from environmental burdens than the tract of interest, and that 25% of tracts in the nation likely experience more severe cumulative impacts from environmental burdens.

- EJI is based on census tract and not available by zipcode
- There are two sets of EJI data: one from 2022 and another from 2024. The Climate Burden Module (CBM) is only available in the 2024 EJI data.
- We primarily used the 2024 EJI file. For participants whose data did not merge using the 2024 data (N=11), we used the 2022 file.

ADDITIONAL ASSISTANCE

If you have questions or concerns about the study datasets that this documentation and the above resources (e.g., protocol, case report forms) have not answered, please email the PHN Mailbox at PHNmailbox@healthcore.com.

APPENDIX A

Implementation of Privacy Protection Rules for Public Use

Variables that could lead to subject identification were eliminated in the public dataset. Steps included:

1. Removal of original study ID number (replaced with *blind_id*, a random consecutive numbering ranging from 1 to 119). Of note, no patient names, addresses, zip code, or medical record numbers were ever contained in the original study dataset.

2. All dates in the original datasets were removed, and replaced with subject age on that date, in years.

3. Free (write-in) text variables were sometimes removed from the public datasets. When included as clinically relevant, they were first scanned for any identifying information (e.g., dates, places) and edited accordingly.

4. Extreme outliers were set to missing (waist circumference for one participant at screening)

5. The study incorporated a diverse array of racial backgrounds. The subject population consisted of four individuals identifying as American Indian or Alaska Native, two identifying as Asian, and eighteen identifying as Black or African American. Additionally, three subjects identify as being of Mixed race. All these individuals have been designated under the broader category of 'Non-White'.