**Supplementary Material for Integrated exposomic analysis of lipid phenotypes: leveraging GE.db in environment by environment association studies**

ANDRE LUIS GARAO RICO and NICOLE PALMIERO

Department of Genomics, University of Pennsylvania, 3700 Hamilton Walk  
Philadelphia, PA 19104, USA  
Email: andreluis.rico@pennmedicine.upenn.edu

MARYLYN D. RITCHIE

Department of Genomics, University of Pennsylvania, 3700 Hamilton Walk  
Philadelphia, PA 19104, USA  
Email: marylyn@pennmedicine.upenn.edu

MOLLY A. HALL

Department of Genomics, University of Pennsylvania, 3700 Hamilton Walk  
Philadelphia, PA 19104, USA  
Email: molly.hall@pennmedicne.upenn.edu

We change the file here!!!

**Contents of This File**

* Introduction
* Figures S-1 through S-X
* Tables S-1 through S-X

**Additional Files Uploaded Separately**

* Captions for Data Sets S-1 through S-X
* Captions for Videos S-1 through S-X
* Captions for Audio S-1 through S-X

**Introduction**

This document provides supplementary materials for the paper titled "Integrated Exposomic Analysis of Lipid Phenotypes: Leveraging GE.db in Environment by Environment Association Studies." The supplementary materials include additional figures and tables that support the findings presented in the main paper. These materials provide further details on the data, methods, and results, including extended descriptions of results and full details of models used in the study. The data were collected from the National Health and Nutrition Examination Survey (NHANES) spanning from 1999 to 2018. Known imperfections and anomalies in the data are discussed, including the limitations of self-reported data and adjustments made for cholesterol-lowering medications.

**Figures**

\*\*Fig. S-1\*\*—[Insert descriptive caption here]. Example: Interaction plot showing the relationship between docosapentaenoic acid (22:5n-3) and arachidic acid (20:0) on HDL-cholesterol levels. The plot includes data points from both the discovery and replication datasets.

\*\*Fig. S-2\*\*—[Insert descriptive caption here]. Example: Scatter plot depicting the distribution of blood benzene levels and HDL-cholesterol across different age groups in the NHANES dataset.

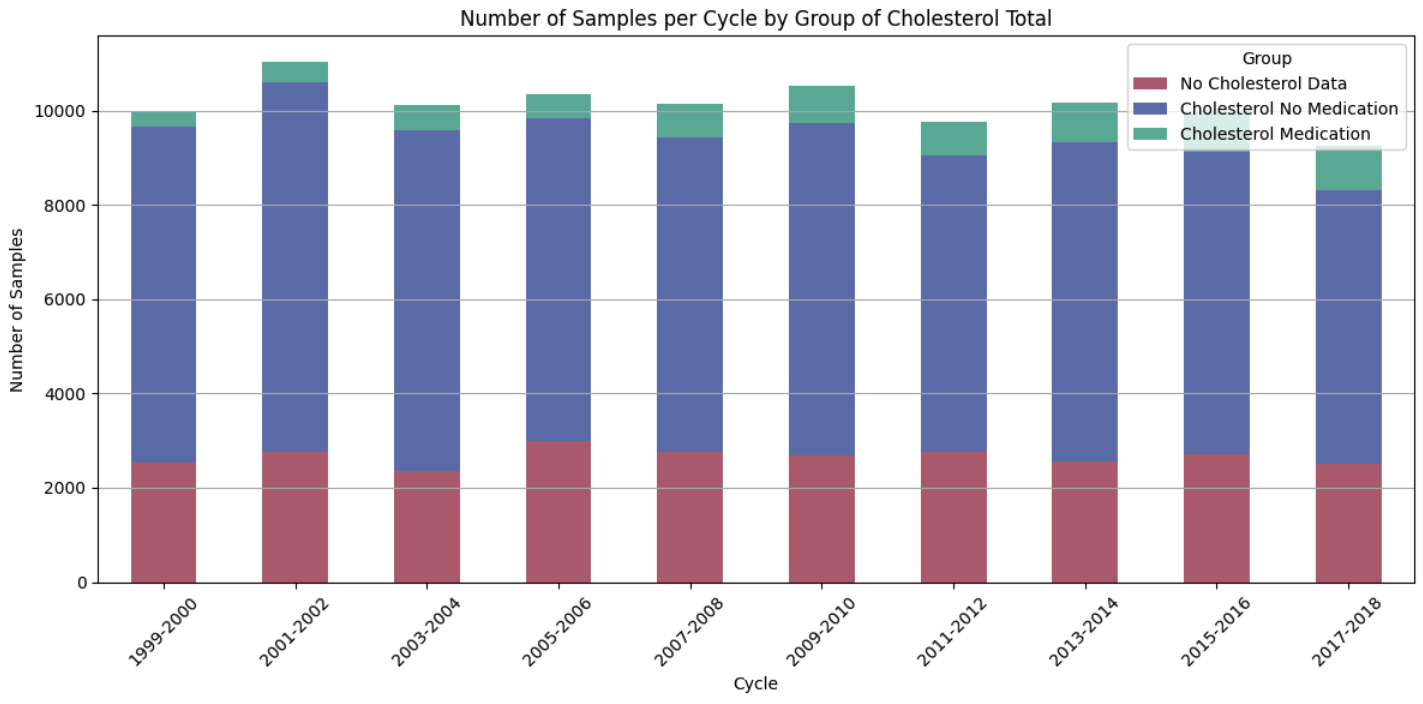
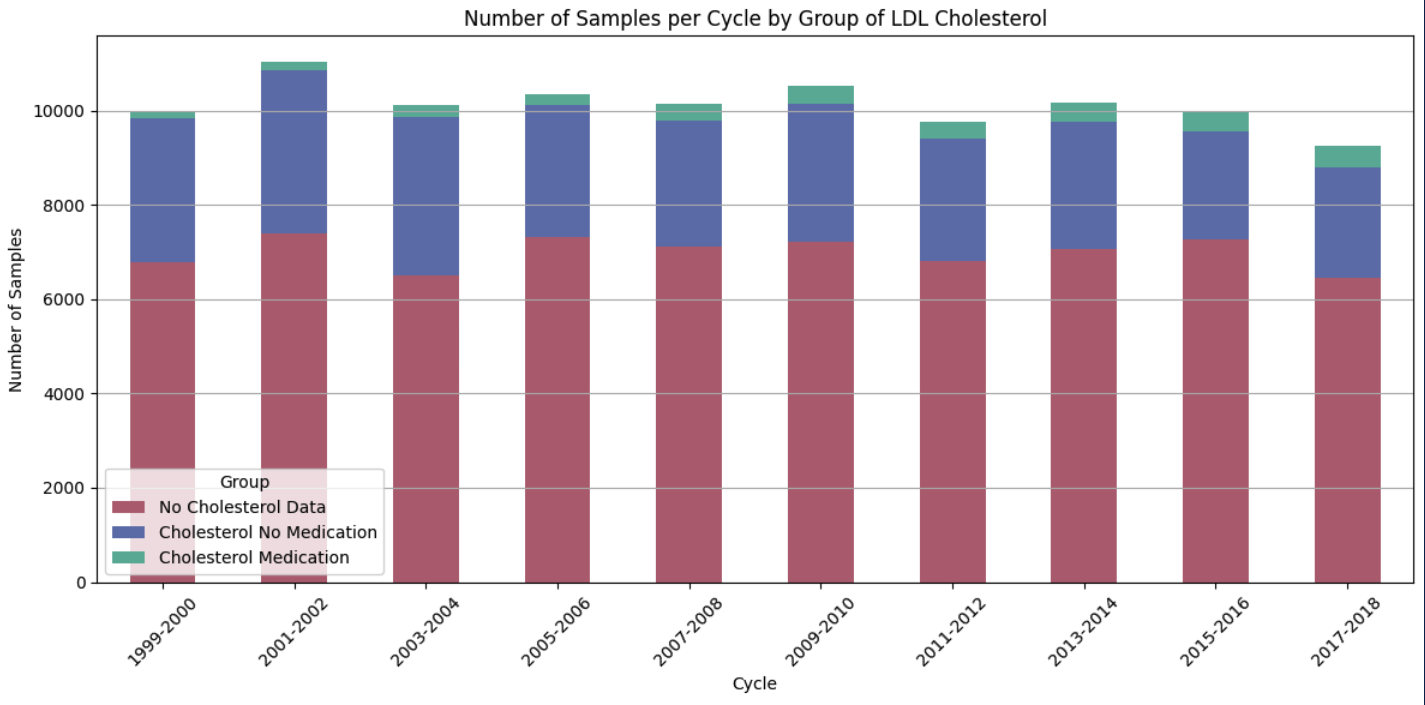
**Tables**

\*\*Table S-1\*\*—[Insert descriptive caption here]. Example: Summary of the significant ExE interactions identified in the discovery dataset with an FDR-adjusted p-value < 0.1.

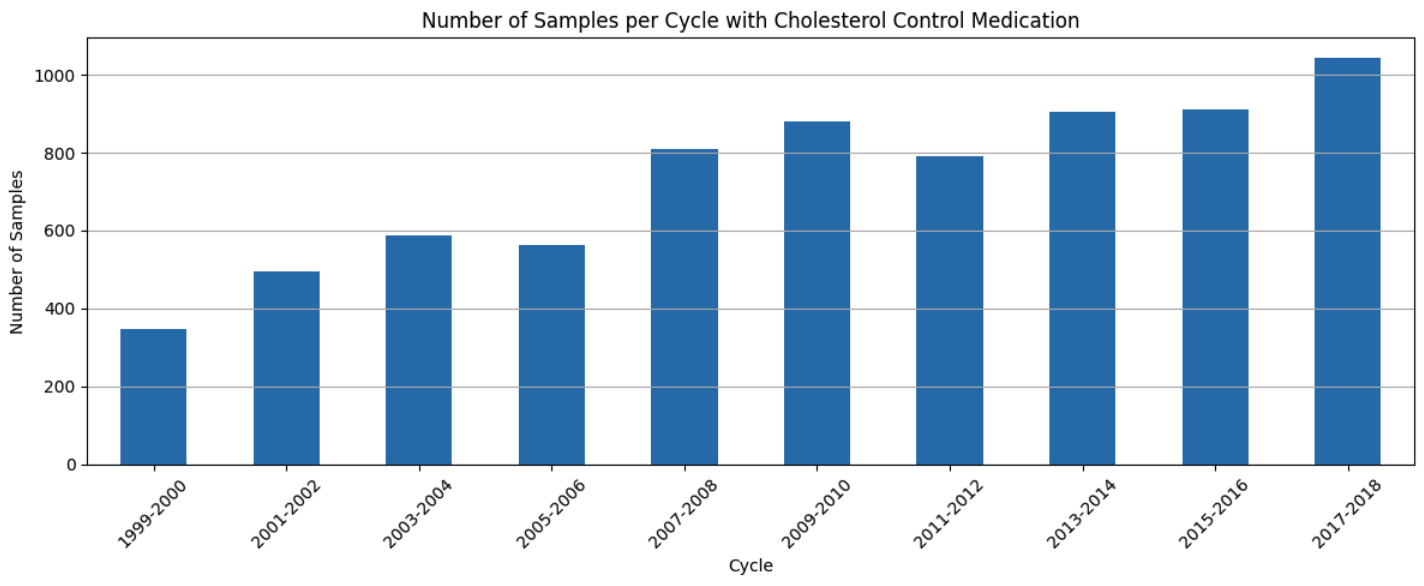
\*\*Table S-2\*\*—[Insert descriptive caption here]. Example: Detailed description of the lipid phenotypes and confounders used in the study, including their NHANES IDs and cycles.

**Data Sets**

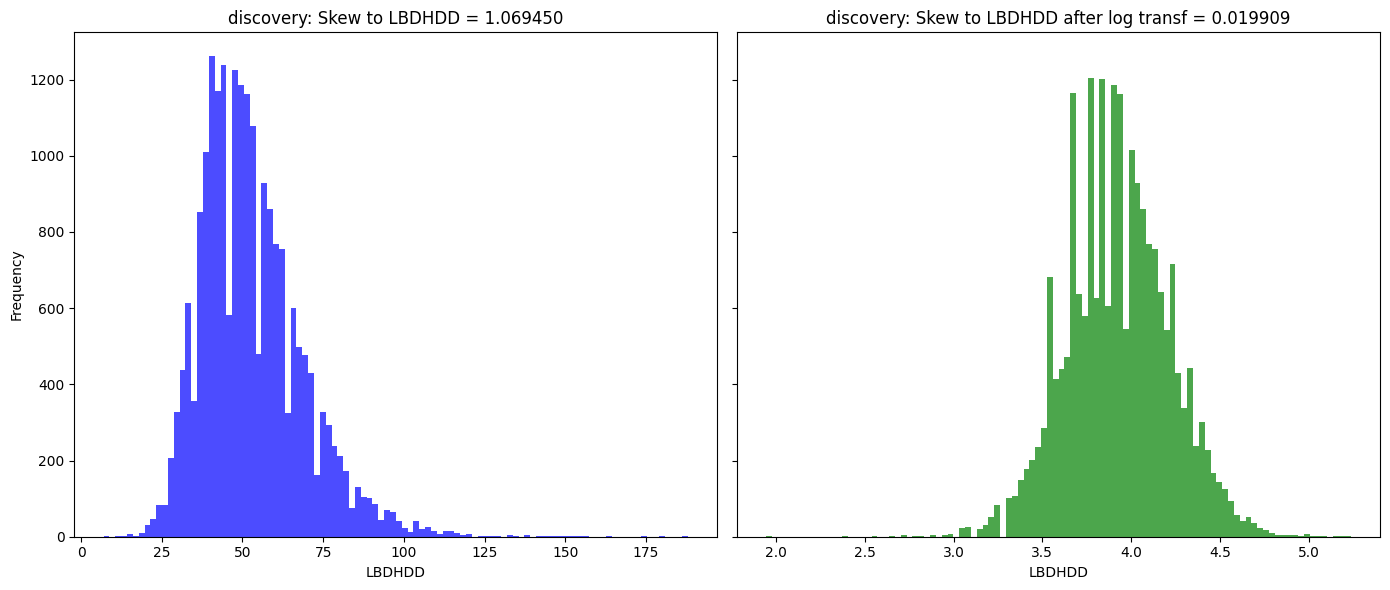
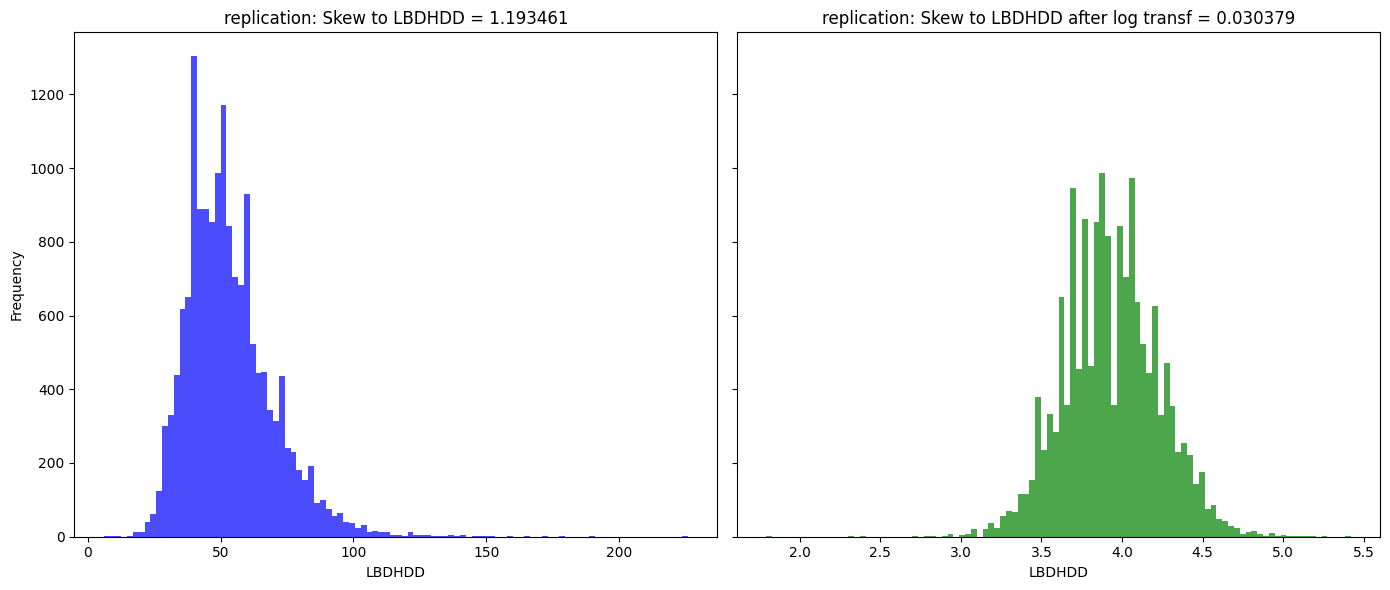
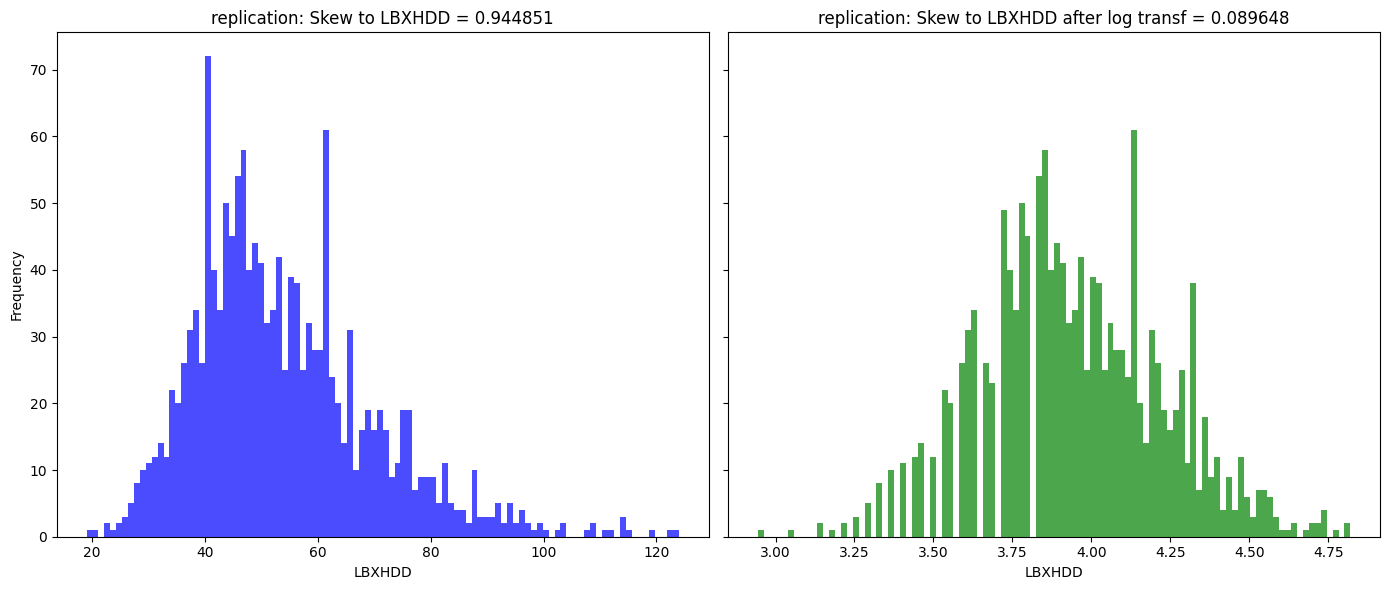
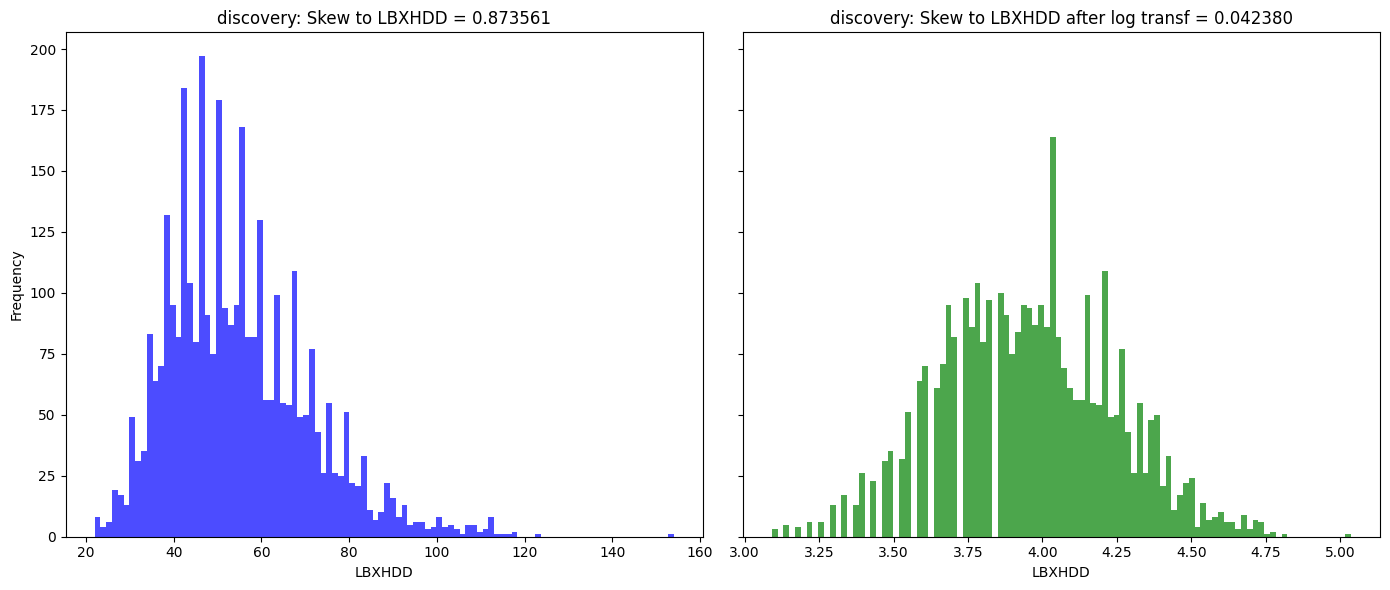
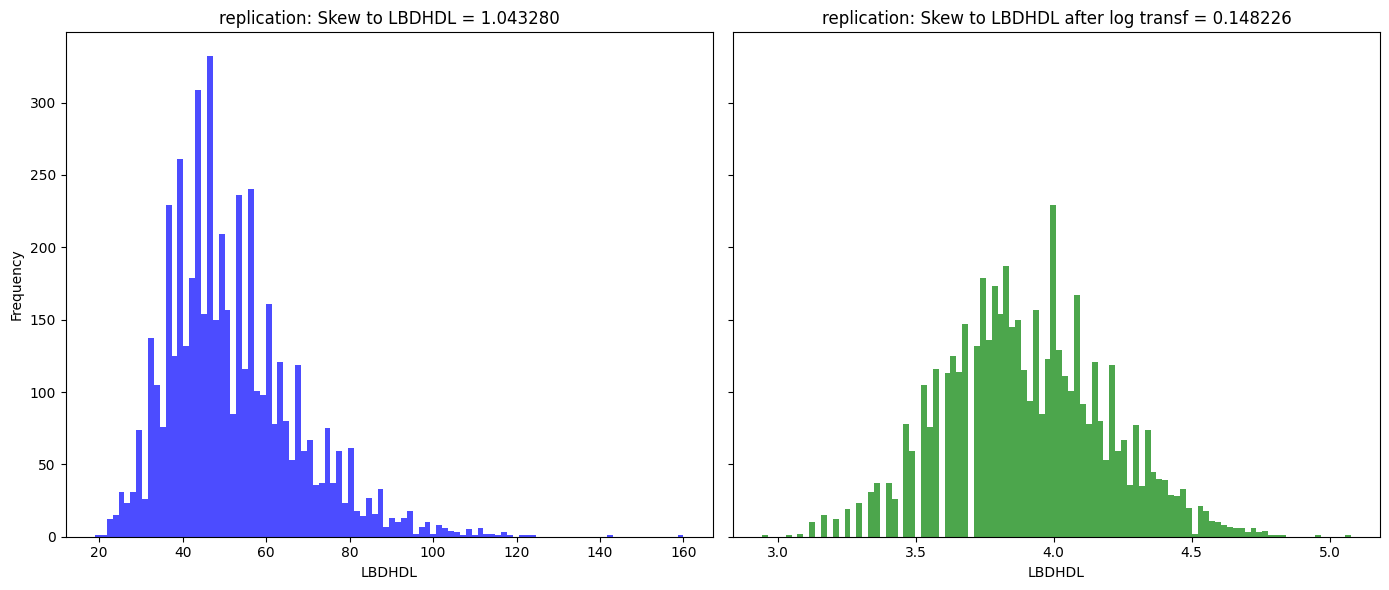
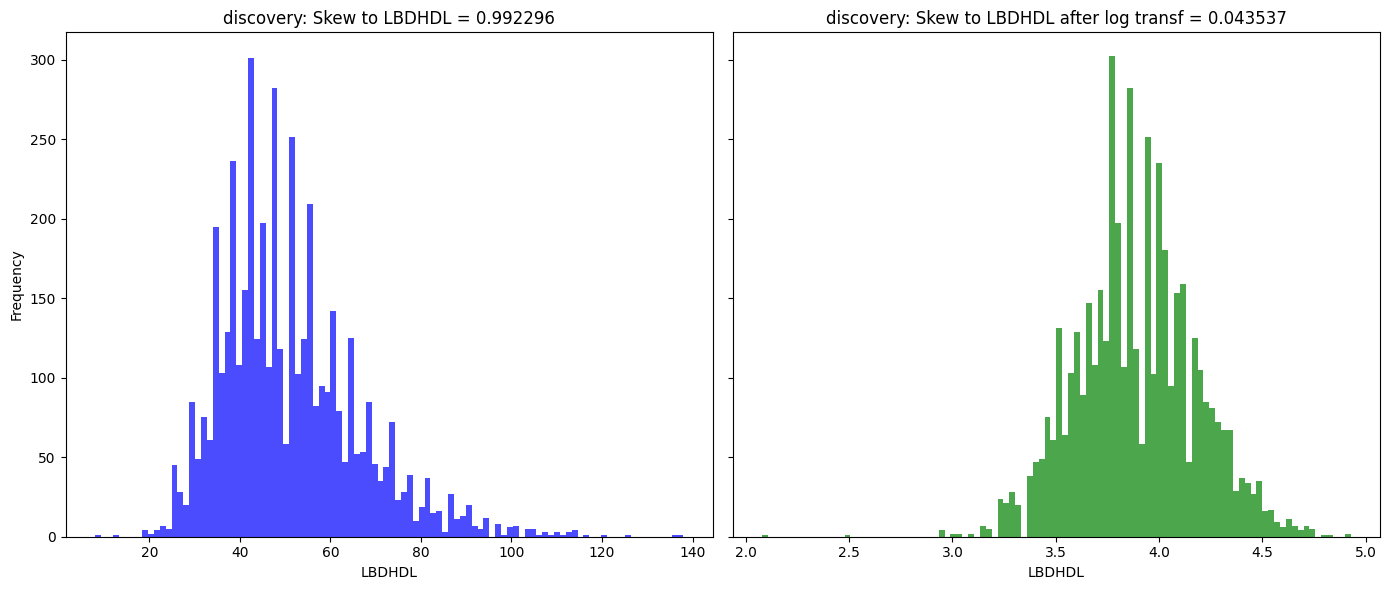
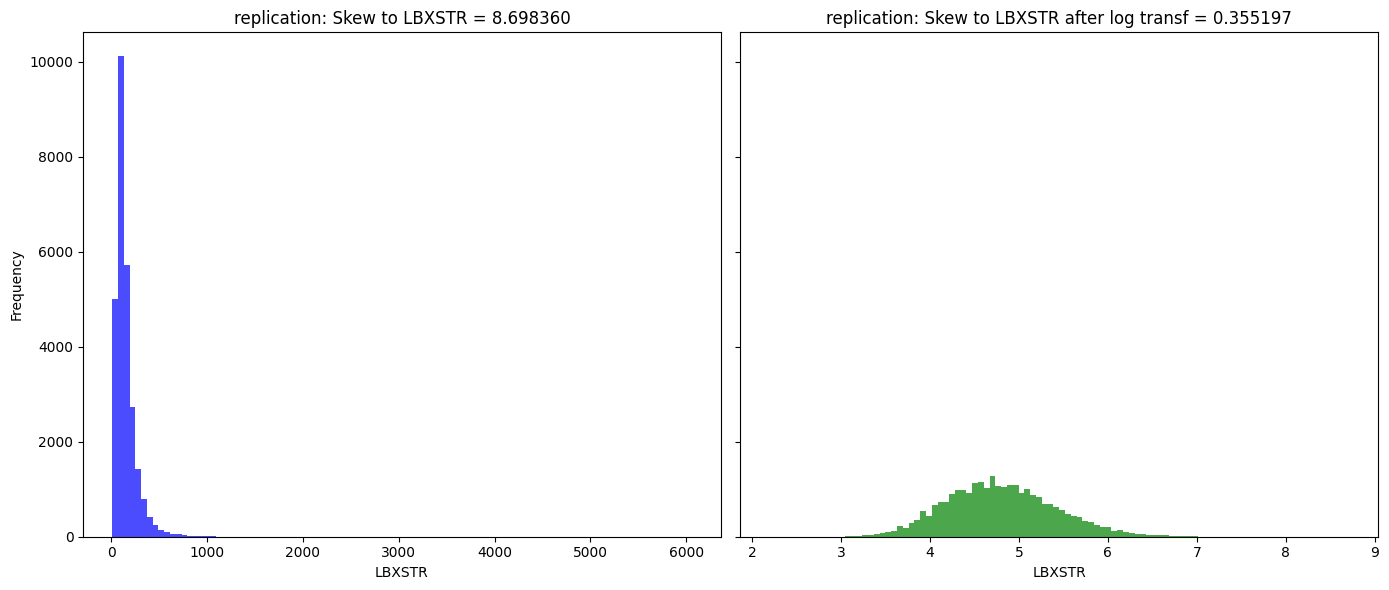
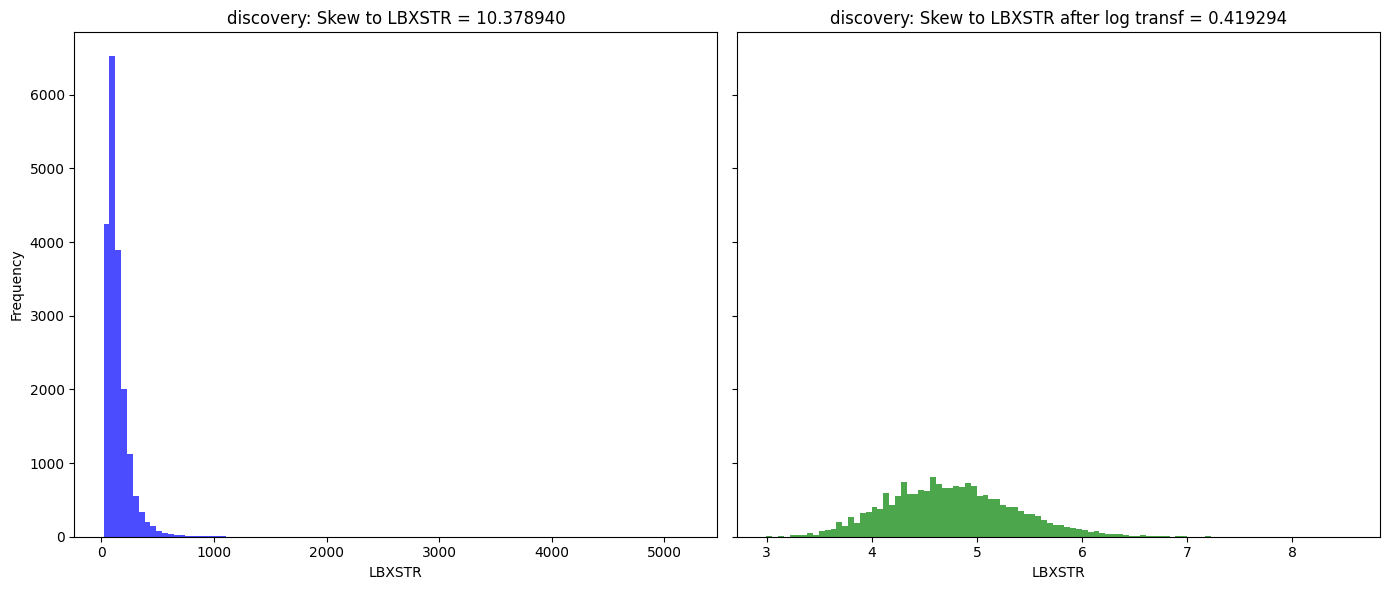
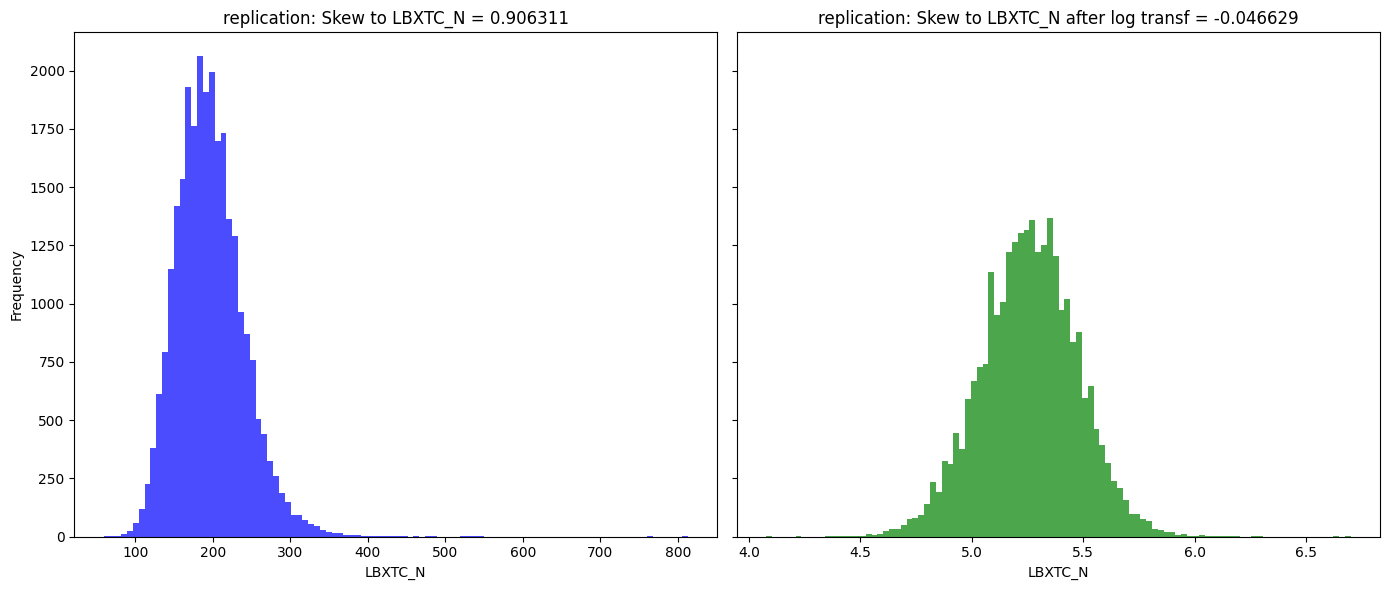
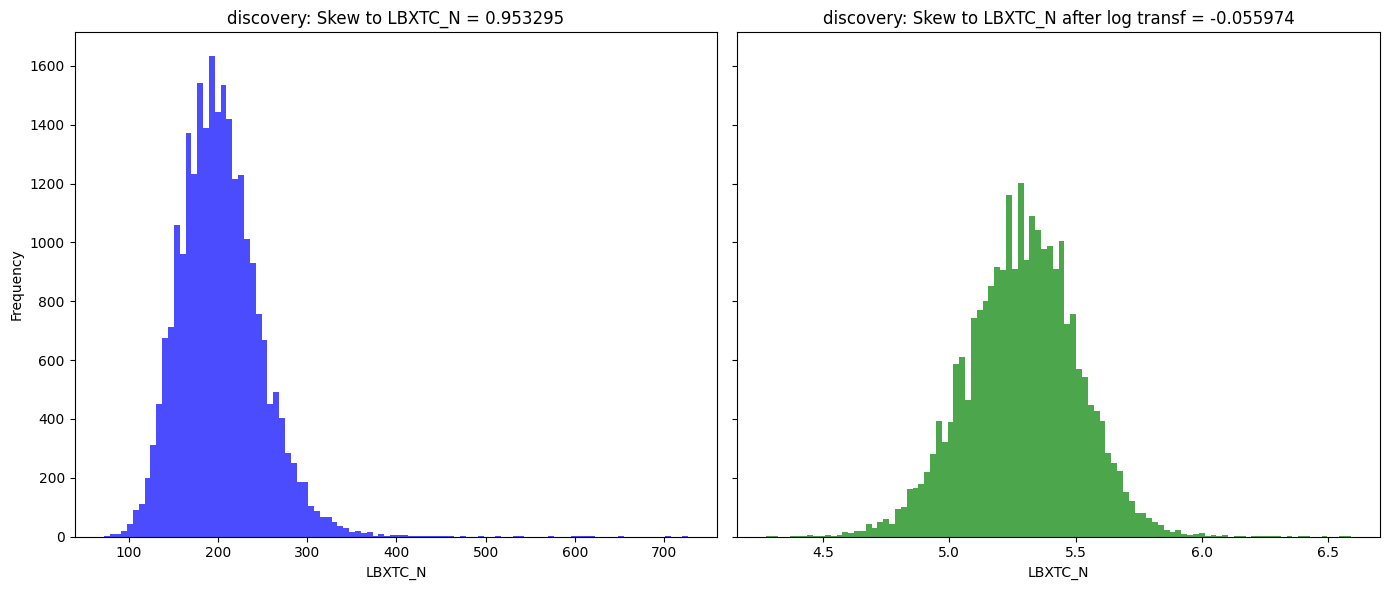
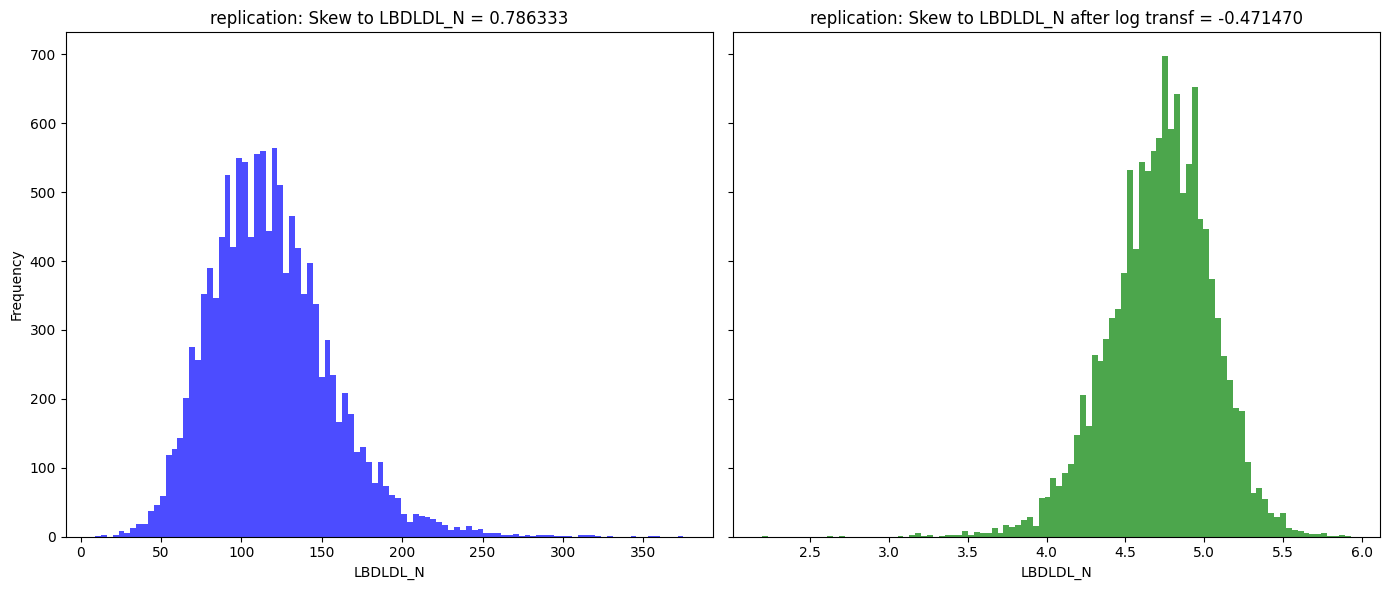
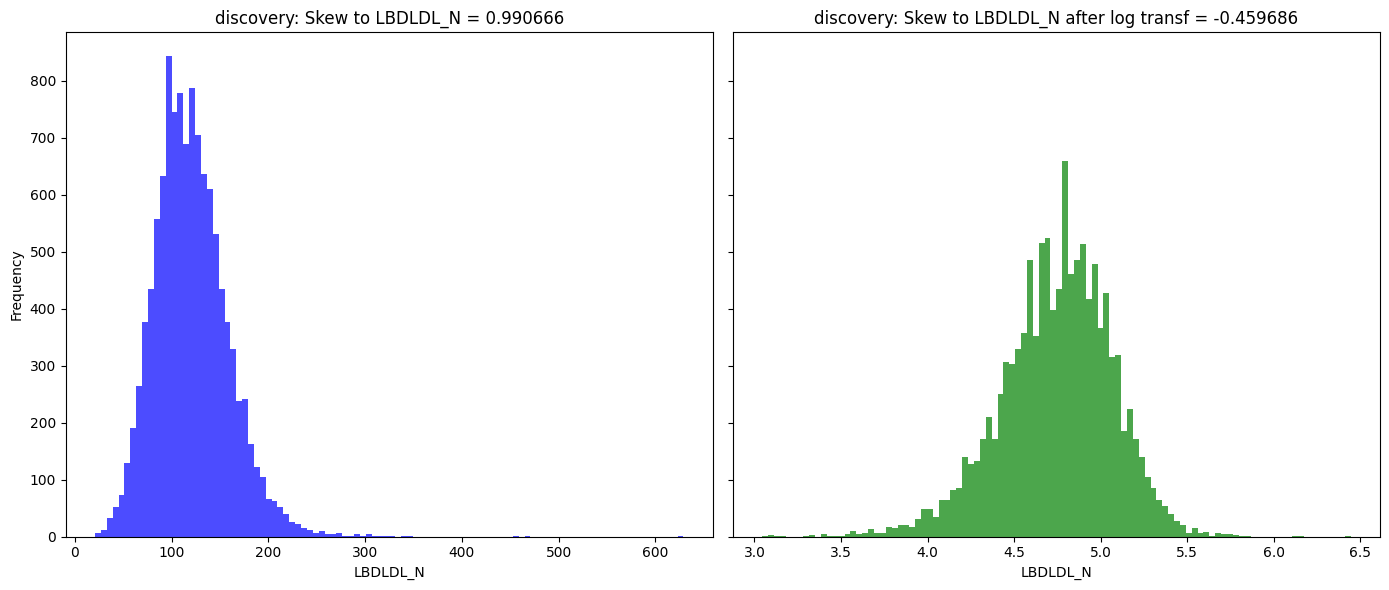
\*\*Data Set S-1\*\*—[Insert caption here]. Example: Dataset containing the raw NHANES data used in the analysis, including participant demographics, dietary information, and laboratory results.



Supplementary Figure 1. Bar charts showing the sample counts by survey cycle of participants who either did or did not have cholesterol data and if they took cholesterol medications.



Supplementary Figure 2. Bar chart featuring only the samples who used cholesterol medication organized by survey cycle.

Supplementary Figure 3. Histograms displaying the frequencies before and after log transformation of the lipid phenotypes in the discovery and replication datasets.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **Outcome** | **Term 1** | **Term 2** | **N** | **p-value** | **p-value bonferroni** | **Full Var1 Var2 beta** | **Full Var1 beta** | **Full Var2 beta** |
| discovery | HDL-C | docosapentaenoic acid (22:5n-3) | arachidic acid (20:0) | 2377 | 1.28075E-16 | 8.43247E-13 | -0.000144898 | 0.002680318 | 0.014532194 |
| replicate | HDL-C | blood 2,5-dimethylfuran | blood benzene | 8398 | 3.17923E-14 | 4.48271E-12 | 0.781909446 | -0.586413251 | -0.033567756 |
| discovery | HDL-C | stearic acid (18:0) | arachidic acid (20:0) | 2415 | 1.34941E-15 | 8.8845E-12 | -7.79411E-06 | 9.92465E-06 | 0.015849325 |
| replicate | HDL-C | blood furan | blood benzene | 8385 | 9.62712E-11 | 1.35742E-08 | 1.420982688 | -1.323958524 | -0.007526433 |
| replicate | HDL-C | linoleic acid (18:2n-6) | arachidic acid (20:0) | 2395 | 1.96582E-10 | 2.7718E-08 | -3.57503E-06 | 3.46268E-05 | 0.021099434 |
| discovery | HDL-C | linoleic acid (18:2n-6) | arachidic acid (20:0) | 2411 | 7.39935E-12 | 4.87173E-08 | -1.41019E-06 | -9.71676E-06 | 0.015430215 |
| replicate | HDL-C | linoleic acid (18:2n-6) | eicosadienoic acid (20:2n-6) | 2369 | 1.06132E-09 | 1.49647E-07 | -2.1948E-06 | 6.64321E-05 | 0.004222916 |
| replicate | HDL-C | palmitic acid (16:0) | arachidic acid (20:0) | 2395 | 1.59601E-09 | 2.25037E-07 | -2.95707E-06 | -1.45341E-05 | 0.019949596 |
| discovery | HDL-C | blood 2,5-dimethylfuran | blood benzene | 9475 | 4.18242E-11 | 2.75371E-07 | 0.976923752 | -0.64427957 | -0.024175285 |
| discovery | HDL-C | docosapentaenoic acid (22:5n-3) | linoleic acid (18:2n-6) | 2467 | 4.25142E-11 | 2.79914E-07 | -7.3444E-07 | 0.003027217 | 3.64004E-05 |
| replicate | HDL-C | stearic acid (18:0) | arachidic acid (20:0) | 2380 | 2.46234E-09 | 3.47191E-07 | -1.25774E-05 | 1.72283E-05 | 0.018372048 |
| discovery | HDL-C | linoleic acid (18:2n-6) | arachidonic acid (20:4n-6) | 2522 | 2.60311E-10 | 1.71389E-06 | -6.93702E-08 | 3.24419E-05 | 0.000491034 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-3) | eicosapentaenoic acid (20:5n-3) | 2404 | 1.82598E-08 | 2.57463E-06 | -1.1111E-05 | -0.002002029 | 0.002018101 |
| discovery | HDL-C | palmitic acid (16:0) | arachidic acid (20:0) | 2415 | 4.1245E-10 | 2.71557E-06 | -1.20948E-06 | -2.67145E-05 | 0.015212688 |
| discovery | HDL-C | arachidic acid (20:0) | eicosadienoic acid (20:2n-6) | 2231 | 6.89813E-10 | 4.54173E-06 | -0.000175853 | 0.012111312 | 0.000217524 |
| replicate | HDL-C | stearic acid (18:0) | eicosadienoic acid (20:2n-6) | 2353 | 5.97817E-08 | 8.42922E-06 | -6.60251E-06 | 0.000195983 | 0.000984914 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-6) | linoleic acid (18:2n-6) | 2423 | 9.84717E-08 | 1.38845E-05 | -2.09496E-06 | 0.006502545 | 3.19393E-05 |
| replicate | HDL-C | linoleic acid (18:2n-6) | arachidonic acid (20:4n-6) | 2440 | 7.49201E-07 | 0.000105637 | -6.84412E-08 | 1.7679E-05 | 0.000438513 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-3) | linoleic acid (18:2n-6) | 2415 | 7.75244E-07 | 0.000109309 | -6.25946E-07 | 0.001790377 | 2.19337E-05 |
| discovery | HDL-C | docosapentaenoic acid (22:5n-3) | arachidonic acid (20:4n-6) | 2471 | 2.0292E-08 | 0.000133603 | -2.99454E-06 | 0.000966904 | 0.000404193 |
| discovery | HDL-C | docosapentaenoic acid (22:5n-3) | docosahexaenoic acid (22:6n-3) | 2461 | 2.6022E-08 | 0.000171329 | -8.02995E-06 | 0.000252371 | 0.000954311 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-3) | arachidic acid (20:0) | 2367 | 2.31352E-06 | 0.000326207 | -0.000129134 | 0.001472715 | 0.012442521 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-6) | stearic acid (18:0) | 2407 | 3.51408E-06 | 0.000495485 | -7.12041E-06 | 0.004383999 | 8.95905E-05 |
| discovery | HDL-C | docosapentaenoic acid (22:5n-3) | eicosapentaenoic acid (20:5n-3) | 2471 | 9.57152E-08 | 0.000630189 | -1.13301E-05 | -0.000901038 | 0.001638041 |
| replicate | HDL-C | eicosapentaenoic acid (20:5n-3) | docosahexaenoic acid (22:6n-3) | 2431 | 5.88022E-06 | 0.000829111 | -2.1792E-06 | 0.001044047 | 0.000166351 |
| replicate | HDL-C | blood benzene | lead (ug/dl) | 7636 | 7.29013E-06 | 0.001027908 | 0.060487832 | -0.152786878 | -0.000918697 |
| replicate | HDL-C | arachidic acid (20:0) | eicosadienoic acid (20:2n-6) | 2321 | 1.13147E-05 | 0.001595377 | -0.00026477 | 0.013139333 | 0.000527346 |
| discovery | HDL-C | blood furan | blood benzene | 6815 | 2.55057E-07 | 0.001679298 | 1.713079939 | -1.735628183 | 0.064637015 |
| discovery | HDL-C | linoleic acid (18:2n-6) | eicosadienoic acid (20:2n-6) | 2254 | 3.76226E-07 | 0.002477071 | -1.00678E-06 | 3.95061E-05 | 0.001524265 |
| discovery | HDL-C | stearic acid (18:0) | eicosadienoic acid (20:2n-6) | 2258 | 4.25208E-07 | 0.002799567 | -3.89678E-06 | 0.000190283 | -0.000173272 |
| discovery | HDL-C | docosapentaenoic acid (22:5n-6) | stearic acid (18:0) | 2448 | 7.76617E-07 | 0.005113246 | -4.072E-06 | 0.002819707 | 5.4422E-05 |
| replicate | HDL-C | blood 2,5-dimethylfuran | cotinine | 8585 | 3.94045E-05 | 0.00555604 | 0.000249928 | -0.008133922 | -0.000304585 |
| discovery | HDL-C | stearic acid (18:0) | docosahexaenoic acid (22:6n-3) | 2516 | 1.16001E-06 | 0.007637476 | -7.26193E-07 | 4.04553E-05 | 0.000853798 |
| replicate | HDL-C | stearic acid (18:0) | docosahexaenoic acid (22:6n-3) | 2426 | 5.85865E-05 | 0.008260698 | -9.31559E-07 | -3.1578E-06 | 0.000970907 |
| discovery | HDL-C | blood benzene | lead (ug/dl) | 9746 | 1.41682E-06 | 0.009328359 | 0.047836334 | -0.194836053 | 0.005451476 |
| replicate | HDL-C | cadmium | lead (ug/dl) | 4945 | 0.003514931 | 0.014059726 | 0.009193112 | -0.040713526 | -0.003505994 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-3) | docosahexaenoic acid (22:6n-3) | 2393 | 0.000118217 | 0.016668644 | -5.83236E-06 | -0.000815409 | 0.000824604 |
| discovery | HDL-C | cadmium | lead (ug/dl) | 4570 | 2.05225E-05 | 0.022903097 | 0.01305091 | -0.064211971 | -0.004896294 |
| discovery | HDL-C | eicosapentaenoic acid (20:5n-3) | docosahexaenoic acid (22:6n-3) | 2516 | 3.73808E-06 | 0.024611518 | -1.80857E-06 | 0.001080847 | 9.69117E-05 |
| discovery | HDL-C | docosapentaenoic acid (22:5n-6) | linoleic acid (18:2n-6) | 2444 | 4.0915E-06 | 0.02693845 | -1.04972E-06 | 0.003715053 | 1.49125E-05 |
| replicate | HDL-C | docosapentaenoic acid (22:5n-3) | arachidonic acid (20:4n-6) | 2402 | 0.000223822 | 0.03155897 | -2.45749E-06 | -0.0003768 | 0.000369318 |
| discovery | HDL-C | blood 2,5-dimethylfuran | cotinine | 10030 | 5.05096E-06 | 0.033255528 | 0.000270204 | -0.02241935 | -0.000329177 |

Supplementary Table 1. All twenty-one Bonferroni adjusted p < 0.05 significant interactions listing their p-values, sample sizes, full interaction betas, and full main-effect betas.