Introduction of example files

**R packages used in this paper:**

**R package “Binary logistic regression” for Binary logistic regression analysis**

***Code:***

Binary logistic regression.R

***Example file:***

Binary logistic regression-input-demo.xlsx,

Binary logistic regression-output-demo.xlsx

***Introduction:***

Binary logistic regression-input-demo.xlsx → “predictor” represents zscore concentration of exposome in samples. Age and gender were corrected confounding factors. BMI represents the group classified by BMI threshold or other diagnosis.

Binary logistic regression-output-demo.xlsx → The column “term” represents the name of each chemical. The column “estimate” represents the odd ration “conf.low” represents the lower limit of confidence interval, “conf.high” represents the higher limit of confidence interval, “p.value” represents the significant of two tails. “OR”, “LowCI”, “HighCI”and“Pvalue”are all derived from the binary logistic regression model.

**R package “MedEffectAnalysis” for Mediation effect analysis**

***Code:***

MedEffectAnalysis.R

***Example file:***

MedEffect-input-demo.xlsx

MedEffect-output-demo.txt

***Introduction:***

MedEffect-input-demo.xlsx → The column “Filename” represents chronic disease samples. The column “group” represents the group classified by diagnosis results. The “M\_number” represents zscore concentration of different metabolites. The “E\_number” represents zscore concentration of different chemical exposures.

MedEffect-output-demo.txt → “E\_15” was used as a demo to calculate the mediating effects of differential metabolites. “Total Effect” represents the total risk effect of “E\_15” to disease. The “ACME (average)” represents the average causal mediated effect from the independent variable (“E\_15”) to the dependent variable (group) through the mediator (“M\_number”). “ADE (average)” represents the average direct effect of the independent variable (“E\_15”) on the dependent variable (group), which is not transmitted through the mediating variable (“M\_number”). “Prop. Mediated” indicates the proportion of the mediating effect in the total effect. \*\*\* represents p <0.001, \*\* represents p <0.01, \* represents p <0.05, .respents p<0.1.

**Codes used for plot in this paper:**

**Donut Bar Chart**

***Code:***

Donut Bar Chart.R

.Rhistory

***Example file:***

Donut Bar Chart-data.tsv

***Introduction:***

Column “geneSymbol” represents the name of chemical exposures. Column “Rho\_deplete” represents quantitative frequency/40 of chemical exposures. Column “med\_LFC” represents the ensemble mean/2 of chemical exposures in all samples.

**Ring Network Diagram**

***Code:***

Ring Network Diagram.R

***Example file:***

edges.csv

nodes.csv

***Introduction:***

The “edges.csv” contains the classification of chemical exposures. Column “population” in “nodes.csv” represents the detection frequency of chemical exposures.

**Sunrise plot**

***Code:***

Sunrise plot.R

pie\_donut.R

***Example file:***

Unnecessary

***Introduction:***

The exposure classification information is already included in the R code, just run it directly.