ADNI Alzheimer's Disease Neuroimaging Initiative

UPLC/MS/MS Analysis of Human Serum Samples from the Alzheimer's disease Neuroimaging Initiative using the Biocrates AbsoluteIDQ p180 Kit

Drs. Rima Kaddurah-Daouk, Lisa St. John-Williams, Arthur Moseley, Will Thompson, Duke University

Alzheimer's Disease Metabolomics Consortium (ADMC)

Contents

Updated Aug 6, 2015

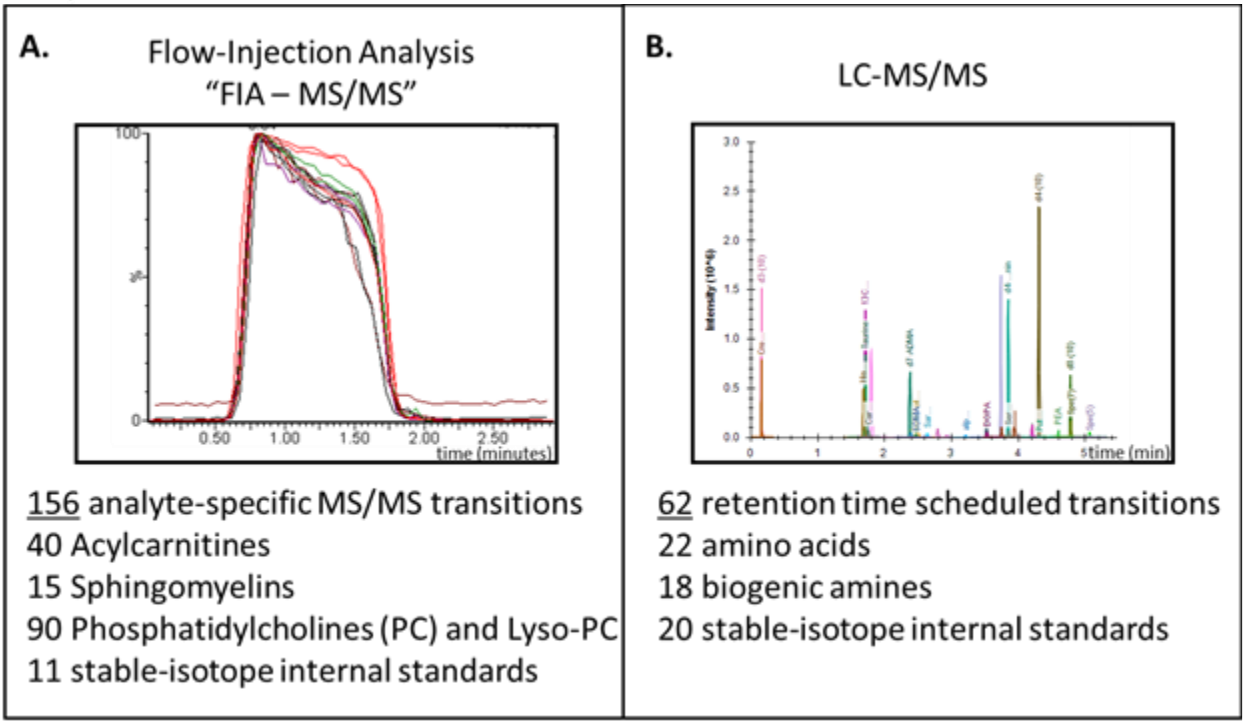
Objective: Measure the levels of selected metabolites in human serum samples using the Biocrates Absolute*IDQ* p180 kit.

Duke Proteomics Core Facility Contributors: Lisa St. John-Williams (sample preparation, data collection, data analysis, report writing), Laura Dubois (sample preparation), Will Thompson (study design, scientific oversight, report writing), and Arthur Moseley (scientific oversight, report writing).

Introduction:

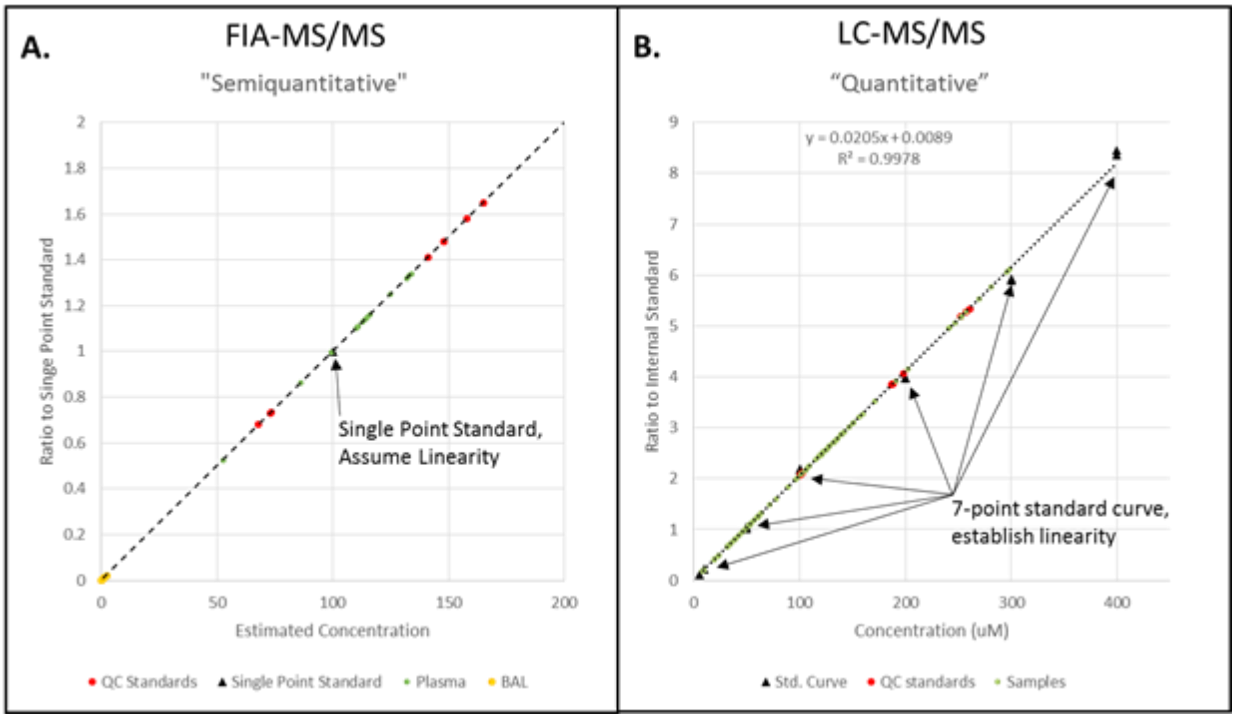
The Absolute*IDQ* p180 assay quantifies over 180 metabolites from five analyte groups: acylcarnities, amino acids, biogenic amines, glycerophospholipids,and sphingolipids. The p180 kit includes all requisite calibration standards, internal standards and QC samples. The use of these standards with a detailed analysis protocol which was validated in Biocrates' lab in Austria assures assay harmonization and standardization within a project, across projects, and across laboratories. Selective analyte detection is accomplished by use of a triple quadrupole tandem mass spectrometer operated in Multiple Reaction Monitoring (MRM) mode in which specific precursor to product ion transitions are measured for every analyte and stable isotope labeled internal standard. There are two separate tandem mass spectrometric analyses of each sample. For the analysis of acylcarnitines, glycerophospholipids, and sphingolipids samples are introduced using a Flow Injection Analysis method (FIA-MS/MS, **Figure 1 A**). Sample analysis of amino acids and biogenic amines are performed by a UPLC (ultra-high pressure liquid chromatography) tandem MS method using a reversed phase analytical column for analyte separation (LC-MS/MS, **Figure 1B**).

Figure 1. Schematic of data collection methodologies for Biocrates p180 kit, including Flow-Injection MS/MS (A) and LC-MS/MS (B).



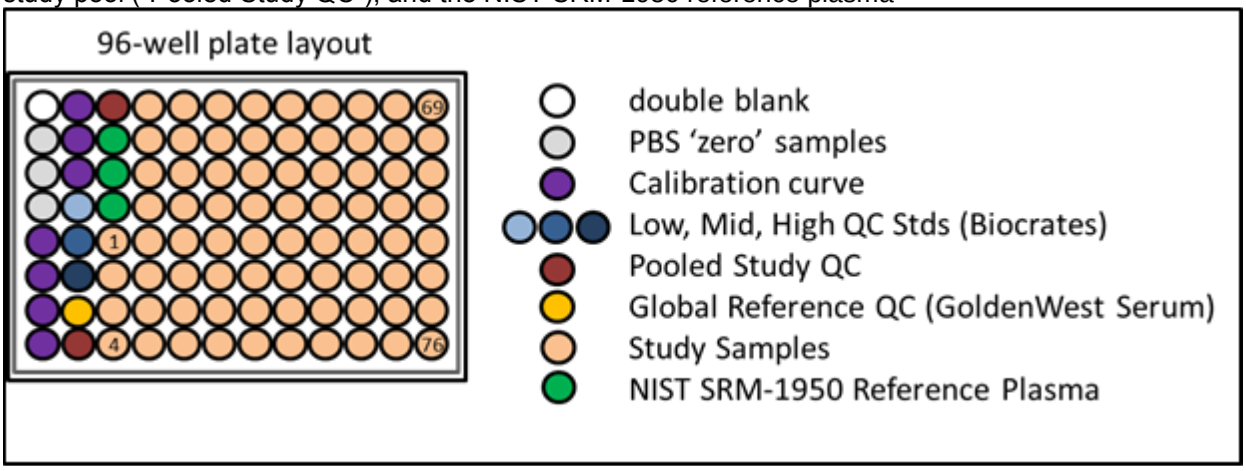
The calibration standards provided in the Biocrates Absolute*IDQ* kit were used for quantitation. A single point calibration standard is used for the reproducible quantitation of the acylcarnitines, glycerophospholipids, and sphingolipids. Therefore, the results for these analytes are described as semi-quantitative (schematic shown in **Figure 2A**). Seven calibration standards are used for highly accurate and reproducible quantitation of the amino acids and biogenic amines as shown in **Figure 2B**. **Figure 2** shows a schematic and representative examples of how the single-point calibrator or calibration curve would be used to back-calculate QC and sample concentrations for the different analyte classes.

Figure 2. Schematic depicting the quantitative methodologies used in the Biocrates p180 kit for flow-injection analysis (A) and LC-MS/MS analysis (B).



The samples were prepared in a 96-well plate format using the layout shown in **Figure 3**. Seventy-six study samples were analyzed on each of the first ten plates. The eleventh plate had 73 study samples.

Figure 3. Schematic depicting the 96-well plate layout for the analysis of the ADNI samples including: requisite blanks, calibration standards, and QC samples from Biocrates, Golden West Serum, the ADNI study pool ("Pooled Study QC"), and the NIST SRM-1950 reference plasma



Sample Preparation:

Samples were prepared using the Absolute*IDQ*® p180 kit (Biocrates Innsbruck, Austria) in strict accordance with their detailed protocol. After the addition of 10 µL of the supplied internal standard solution to each well of the 96-well extraction plate, 10 µL of each serum sample were added to the appropriate wells. The plate was then dried under a gentle stream of nitrogen. The samples were

derivatized with phenyl isothiocyanate then eluted with 5mM ammonium acetate in methanol. Samples were diluted with either 40% methanol in water for the UPLC analysis (15:1) or running solvent (a proprietary mixture provided by Biocrates) for flow injection analysis (20:1).

A pool of the remaining volume of the first 76 samples in Box 1 was created (Study Pool QC). This sample was frozen in aliquots of an appropriate volume and analyzed independently on all of the plates analyzed in this study. The pooled sample was prepared and analyzed in duplicate in the same way as the study samples. From each plate this sample was injected once before and once after the samples in order to measure the performance of the assay across the sample cohort. The analyses of this pool can be used to assess potential batch effects

Note that two samples had insufficient volume for sample analysis. Those samples are A-203 and A-254.

Sample Analysis:

UPLC separation of amino acids and biogenic amines was performed using a Waters (Milford, MA) Acquity UPLC with a Waters Acquity 2.1 mm x 50 mm 1.7 µm BEH C18 column fitted with a Waters Acquity BEH C18 1.7 µm Vanguard guard column. Analytes were separated using a gradient from 0.1% formic acid in water, to 0.1% formic acid in acetonitrile. Total UPLC analysis time was approximately 7 minutes per sample. Acylcarnitines, sphingolipids, and glycerophospholipids were analyzed by flow injection analysis (FIA) with total analysis time of approximately 3 minutes per sample. Using electrospray ionization in positive mode, samples for both UPLC and flow injection analysis were introduced directly into a Xevo TQ-S triple quadrupole mass spectrometer (Waters) operating in the Multiple Reaction Monitoring (MRM) mode. MRM transitions (compound-specific precursor to product ion transitions) for each analyte and internal standard were collected over the appropriate retention time. The UPLC-MS/MS data were imported into Waters application TargetLynx™ for peak integration, calibration, and concentration calculations. The UPLC-MS/MS data from TargetLynx™ and FIA-MS/MS data were analyzed using Biocrates Me*tIDQ*™ software.

Data Return Document Descriptions

This dataset includes 4 main files, and 4 supplementary files:

File	Description	Comments
admcdukep180_dict.csv	Data Dictionary	See below
ADNI_Methods_ADMC_4097_ Sample_Analysis_Summary.pdf	Methods/Summary/Report	This document
adni_duke_4097_uplc_p180.csv	Tab delimited results file	csv version of 4097_UPLC_p180_Data.xlsx with QC metadata removed
adni_duke_4097_fia_p180.csv	Tab delimited results file	csv version of 4097_FIA_p180_Data.xlsx with QC metadata removed
4097_UPLC_p180_Data.xlsx	Original Data File	Supplementary- see below, found in ADMC_supplement.zip - Supplementary materials
4097_UPLC_NIST_and_QC_Data.xlsx	Original QC Data File	Supplementary- see below, found in ADMC_supplement.zip - Supplementary materials
4097_FIA_p180_Data.xlsx	Original Data File	Supplementary- see below, found in ADMC_supplement.zip - Supplementary materials

4097_FIA_NIST_and_QC_Data.xlsx	Original QC Data File	Supplementary- see below, found in ADMC_supplement.zip - Supplementary materials

Duke p180 Data Dictionary.xlsx

This document describes the column headers in the data spreadsheets described below including p180 analytes using multiple naming conventions.

4097 UPLC p180 Data.xlsx

This is an Excel workbook containing two worksheets. The first worksheet contains the calculated concentration data (µM) acquired in the study for analyte classes Amino Acids and Biogenic Amines. The second row in this worksheet lists the analytes measured. Rows 4 through 14 list the lower limit of detection (LOD) for the analyte in each of the plates analyzed. The Biocrates-defined lowest calibration standard and highest calibration standard are listed in Rows 15 and 16. The barcode number of the plate on which each sample was analyzed is listed in column A. The Sample Bar Code number (column B) is assigned to every sample by the Biocrates MetIDQ software. The Sample Identification Number in column D is the unique sample identifier assigned by the Proteomics and Metabolomics Sample Submission System. Column G lists the Customer Sample Identification number which were listed on every sample tube provided for analysis. Table 1 below gives the unique plate barcode for the LC-MS/MS and Flow-Injection Analysis-MS (FIA-MS) plates for each sample box number. The results for the Study Pool QC sample are in rows 849 through 870. Row 872 has the %CV for the Study Pool QC sample results. It is our recommendation that if the inter-day variability of the Study Pool QC sample (as measured by the %CV) for a particular analyte is greater than 25%, then the data for that analyte in the study samples should be flagged as a potential outlier. In addition, some scaling to control for batch effects may be needed during statistical analysis; further investigation of batch effect is discussed below under Principal Components Analysis.

Table 1. Plate Barcodes for Sample Box Number delivered.

Sample Numbers	Plate Barcode for UPLC-MS/MS Analysis.	Plate Barcode for FIA-MS Analysis
A-001 through A-076	1007800073-1	1007800180-1
A-077 through A-152	1007800326-1	1007800428-1
A-153 through A-228	1007800432-1	1007800447-1
A-229 through A-304	1007800451-1	1007800466-1
A-305 through A-380	1007800471-1	1007800485-1
A-381 through A-456	1007800490-1	1007800501-1
A-457 through A-532	1007800515-1	1007800520-1
A-533 through A-608	1007800534-1	1007800549-1
A-609 through A-684	1007800553-1	1007800568-1
A-685 through A-760	1007800572-1	1007800587-1
A-761 through A-833	1007800591-1	1007800602-1

In the data document ***4097 UPLC p180 Data.xlsx*** the concentration data (µM) are color-coded as shown below in Table 2 in order to allow presentation of additional information regarding data quality. (Hovering over the cell in the spreadsheet will also reveal the indication of the color code.):

Table 2. Key to the Analyte Status Fields Used in the Data Table *4021 p180 Data.xlsx*

Valid	Calculated concentrations are based on a standard curve of the analyte listed. (Not all analytes are contained in the calibration standards provided by Biocrates. Therefore, results for some anlaytes will be coded as Semi-Quantitative.)
<Lowest CS: Lowest Calibration Standard>Value>LOD	The value is greater than the LOD but less than the Biocrates-defined lowest calibration standard. These values should be considered reliable except for analytes for which the %CV observed for the Pooled sample was greater than 25%.
>Highest CS	Greater than the highest calibration standard
<LOD	Less than the Lower Limit of Detection. The LODs were given by Biocrates, and were not determined using FDA guidelines or a Waters TQ-S mass spectrometer.
Internal Standard out of range	The internal standard peak area for this analyte was outside of the normal range. Data coded yellow should not be used. (Not all analytes have internal standards.)
Semi-Quantitative	Calculated concentration is not based on a calibration curve but on the peak area of an internal standard or that of a structurally similar analyte. This data is valid but is semi-quantitative instead of fully quantitative.
No Interception	No peak was detected in the chromatogram at the appropriate retention time for this analyte.

The second worksheet in the workbook contains a statistical summarization of the data from the Biocrates Met*IDQ*™ application. The Analyte Statistics worksheet lists these summary statistics (Min µM, Max µM, Mean µM, Median µM, 25th Percentile µM, 75th Percentile µM, STD µM, MAD µM, Skewness, Kurtosis, CV [%], CVRobust [%]) by metabolite for all of the samples. The number of results (e.g. samples with usable data) included for each statistical calculation is indicated by "n" in Column C.

4097 UPLC NIST and QC Data.xlsx

This workbook contains two worksheets. The first worksheet contains the Biocrates plasma QC concentration data for three levels of QCs for analyte classes amino acids and biogenic amines. The measurements for the QC samples are provided in order to confirm that quantification of the metabolites across a wide dynamic range performed in this analysis is generally accurate and reproducible. The second worksheet lists the results for the NIST SRM 1950 samples. Three replicates of this sample were prepared on every sample plate. One replicate was injected before the study samples, one in the middle, and one after the study samples as an additional measure pf assay performance.

4097 FIA p180 Data.xlsx

This is an Excel workbook containing two worksheets. The first worksheet contains the calculated concentration data (µM) acquired in the study for analyte classes Glycerophospholipids, Sphingolipids, and Acylcarnitines. The second row in this worksheet lists the analytes measured. Rows 4 through 14 list the lower limit of detection (LOD) for the analyte in each of the plates analyzed. The Biocrates-defined lowest calibration standard and highest calibration standard are listed in Rows 15 and 16 if applicable. The barcode number of the plate on which each sample was analyzed is listed in column A. The Sample Bar Code number (column B) is assigned to every sample by the Biocrates MetIDQ software.

The Sample Identification Number in column D is the unique sample identifier assigned by the Proteomics and Metabolomics Sample Submission System. Column G lists The Customer Sample Identification number which were listed on every sample tube provided for analysis. Table 1 above gives the unique plate barcode for the LC-MS/MS and Flow-Injection Analysis-MS (FIA-MS) plates for each sample box number.The results for the pooled sample are in rows 849 through 870. Row 872 has the %CV for the Study Pool QC sample results. It is our recommendation that if the inter-day variability of the Study Pool QC sample (as measured by the %CV) for a particular analyte is greater than 25%, then the data for that analyte in the study samples should be flagged as a potential outlier. In addition, some scaling to control for batch effects may be needed during statistical analysis; further investigation of batch effect is discussed below under Principal Components Analysis

In the data document ***4097 FIA p180 Data.xlsx*** the concentration data (µM) are color-coded as shown above in Table 2 in order to allow presentation of additional information regarding data quality. (Hovering over the cell in the spreadsheet will also reveal the indication of the color code.)

The second worksheet in the workbook contains a statistical summarization of the data from the Biocrates Met*IDQ*™ application. The Analyte Statistics worksheet lists these summary statistics (Min µM, Max µM, Mean µM, Median µM, 25th Percentile µM, 75th Percentile µM, STD µM, MAD µM, Skewness, Kurtosis, CV [%], CVRobust [%]) by metabolite for all of the samples. The number of results (e.g. samples with usable data) included for each statistical calculation is indicated by "n" in Column C.

4097 FIA NIST and QC Data.xlsx

This workbook contains two worksheets. The first worksheet contains the Biocrates plasma QC concentration data for three levels of QCs for analyte classes acylcarnitines, glycerophospholipids, and sphingolipids. QC concentrations are not provided for all the analytes detected, only for the subset of analytes within each class which have reference values available from Biocrates. The measurements for the QC samples are provided in order to confirm that quantification of the metabolites across a wide dynamic range performed in this analysis is accurate and reproducible. The second worksheet lists the results for the NIST SRM 1950 samples. Three replicates of this sample were prepared on every sample plate. One replicate was injected before the study samples, one in the middle, and one after the study samples as an additional measure pf assay performance.

Principal Components Analysis (PCA) Plots

In order to assess general variability for the samples within each analyte class, to assess batch effect, and to look for sample outliers; a Principal Components Analysis (PCA) was performed for amino acids, biogenic amines, acylcarnitines, using JMP® Pro v11.0 software (SAS, Cary, NC) and the data from tables ***4097 UPLC p180 Data.xlsx*** and ***4097 FIA p180 Data.xlsx***. Analytes with measurable concentrations for more than half of the samples were included in the analysis. For analytes for which there were missing values, missing values were replaced with the Biocrates LOD value for that analyte and the restricted maximum likelihood (REML) method was used for correlation.

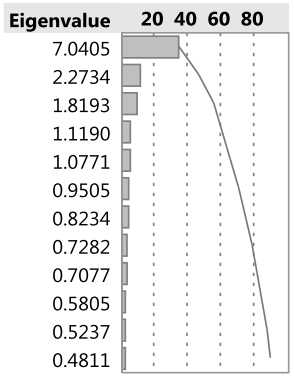
Figures below show PCA plots for each analyte class. The Study Pool values for each plate (same sample) as well as the individual samples for each plate as separate colors in order to globally examine plate batch effect. Even though the vast majority of analytes measured have Study Pool QC %CV less than 20% and biological variance (%CV) 50% or more, therefore, batch effects appear to be minimal based on this evaluation. For amino acids and biogenic amines, there is a single sample outlier, sample A-067. This appears to be largely due to outlier values in creatinine, ADMA and SDMA of this sample, as this sample was not a noticeable outlier with respect to lipid measurements.

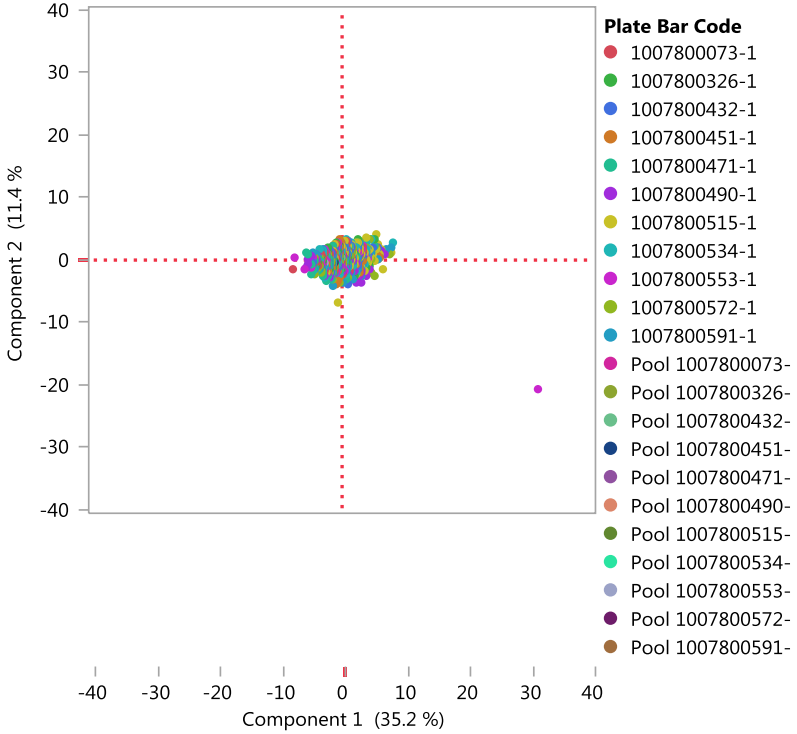
Whether or not correction based on batch effect will be needed during statistical analysis of this data is ultimately up to the determination of professional statisticians, however, if using batch correction in the statistical analysis the following should be considered:

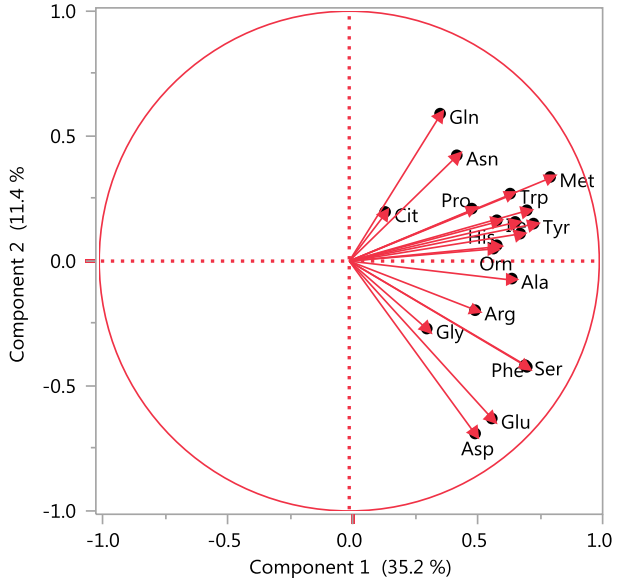
1. Only certain analytes exhibit batch effect so any batch corrections should be performed on an analyte-by-analyte basis.
2. The Study Pool QC samples in each table can provide one measure/method for correction for potential batch effect. Any batch effect may be mitigated by slightly adjusting the measured values for that plate so that the average value for the Pool is equal between plates on a per-analyte basis. The QC Concentration Data in the appropriate data file the low/mid/high QC samples for each plate (which were also the same) which can also be used to correct for batch effect. These measures suggest that there is no significant batch effect in this study.
3. For studies where sample cohorts are well-randomized between the plates, external statistical analysis has shown that using the entire plate-based average of a particular analyte can serve as a better correction for batch effect than using only one or a few wells (such as the Study Pool QC). This should be considered as an option for batch effect correction, should it be needed.

ADNI Alzheimer's Disease Neuroimaging Initiative

Principal Components Analysis of Amino Acids

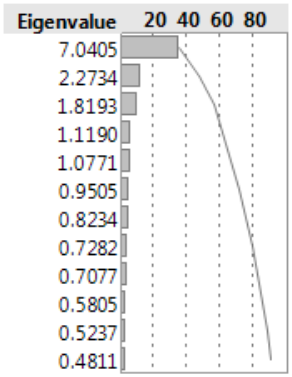


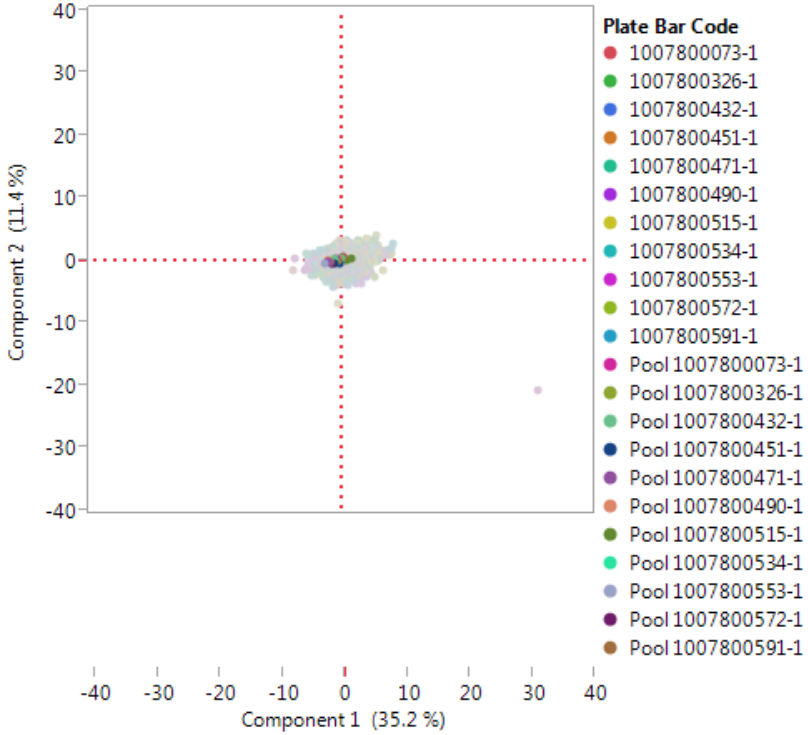


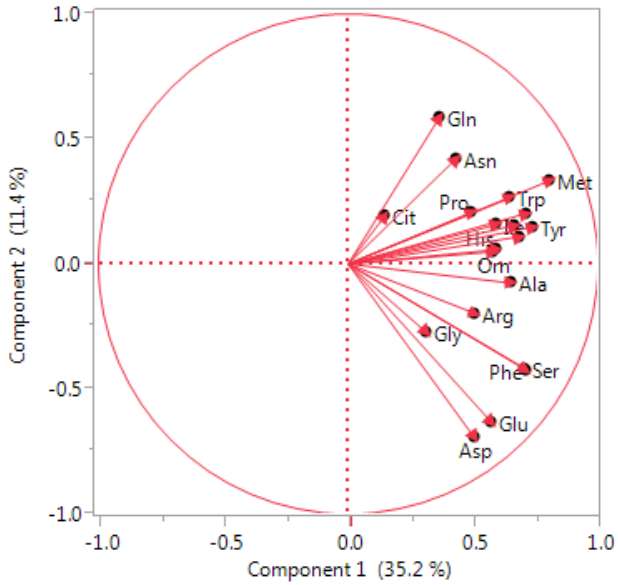


ADNI Alzheimer's Disease Neuroimaging Initiative

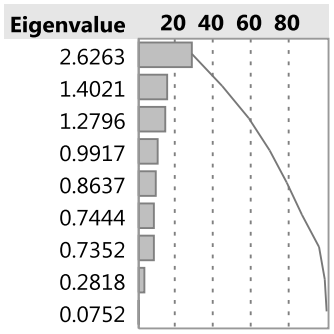
Principal Components Analysis of Amino Acids (Pools Highlighted)

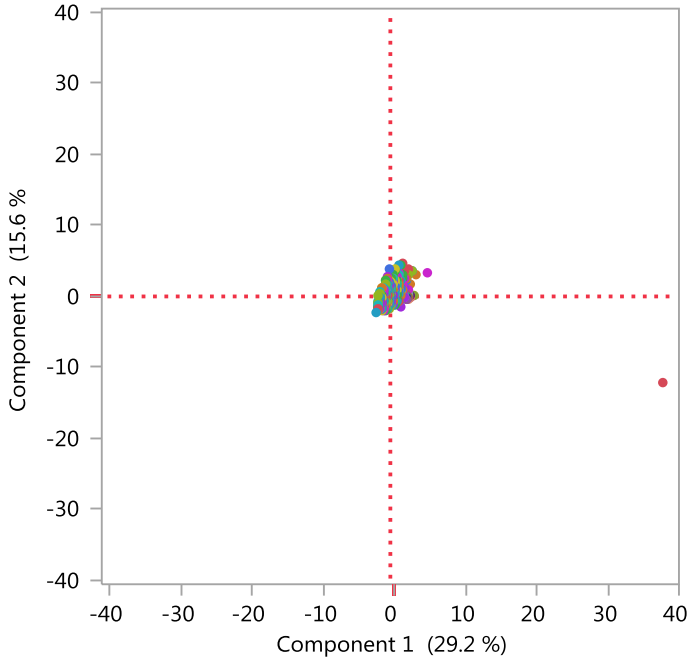


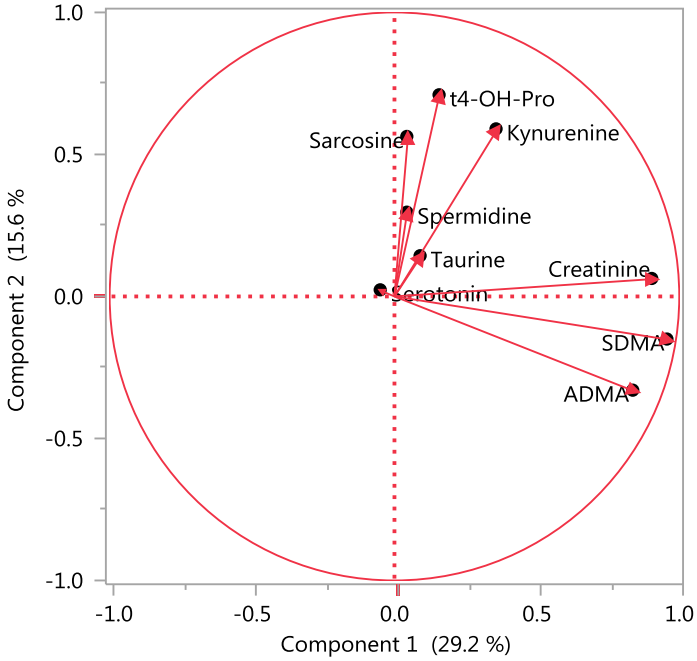




Principal Components Analysis of Biogenic Amines

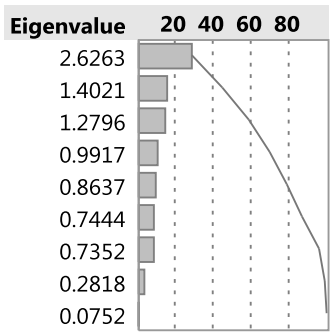


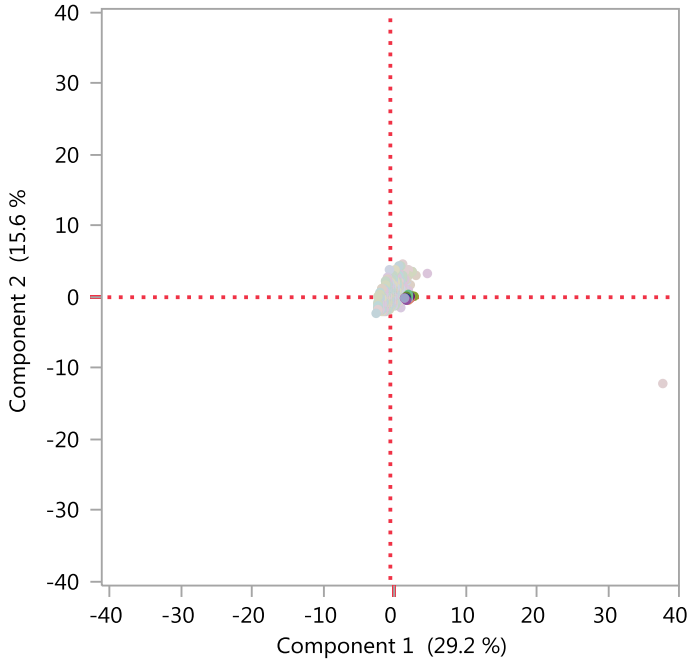


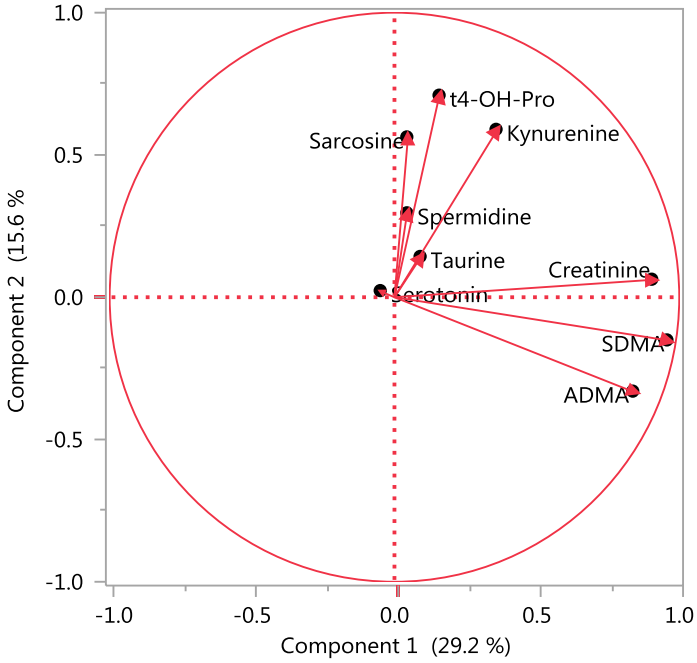


ADNI Alzheimer's Disease Neuroimaging Initiative

Principal Components Analysis of Biogenic Amines (Pools Highlighted)

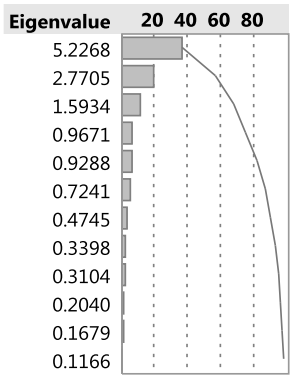


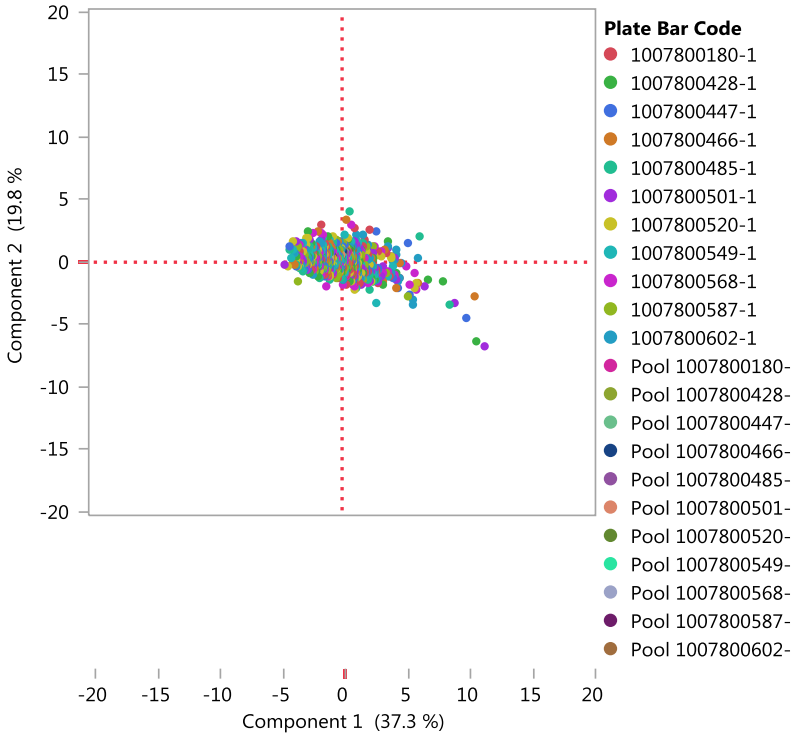


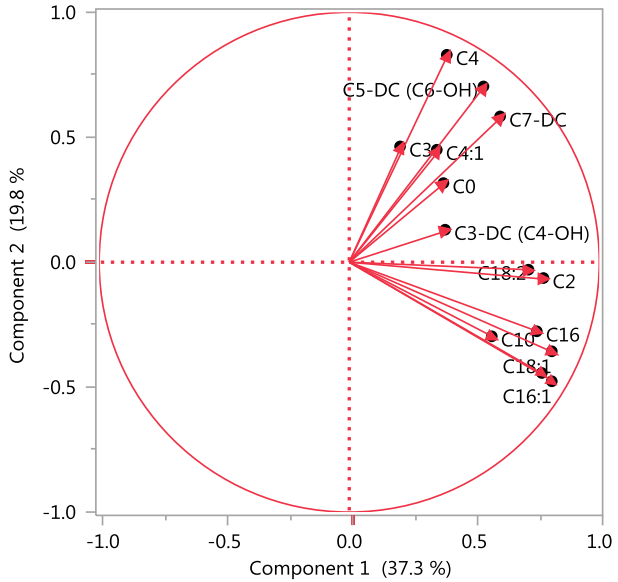


ADNI Alzheimer's Disease Neuroimaging Initiative

Principal Components Analysis of Acylcarnitines

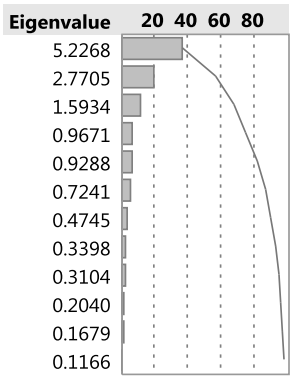


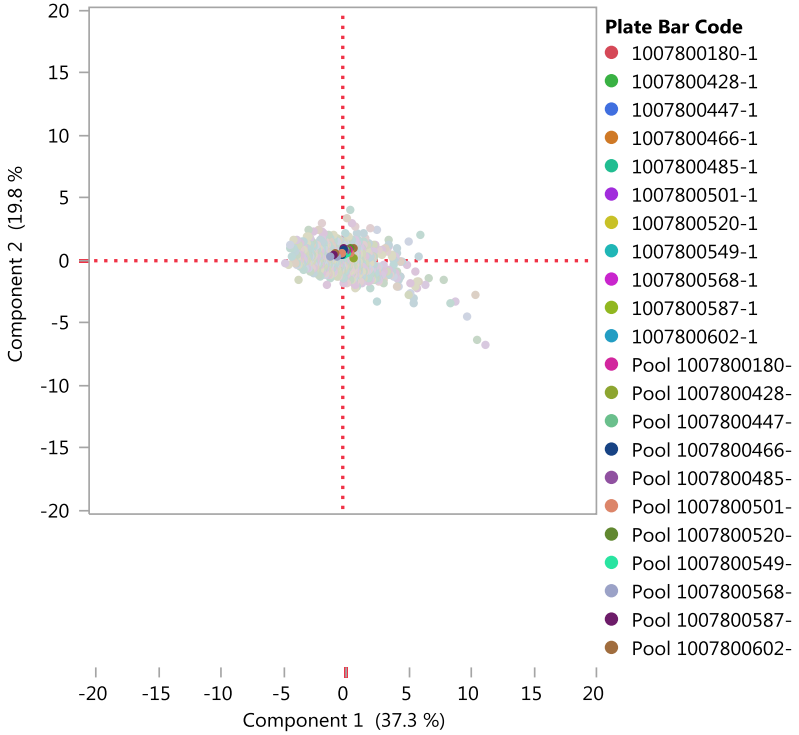


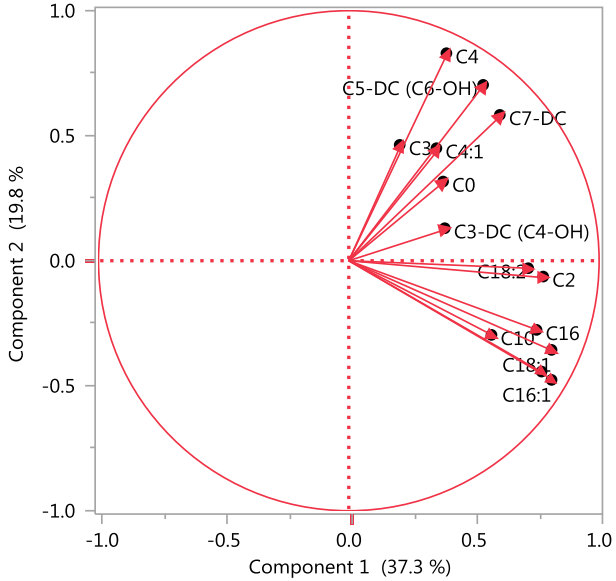


ADNI Alzheimer's Disease Neuroimaging Initiative

Principal Components Analysis of Acylcarnitines (Pool Highlighted)

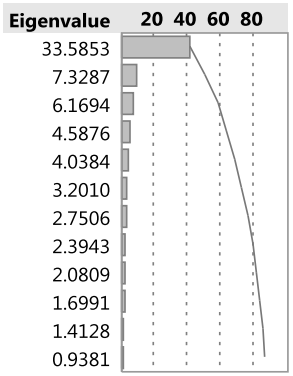


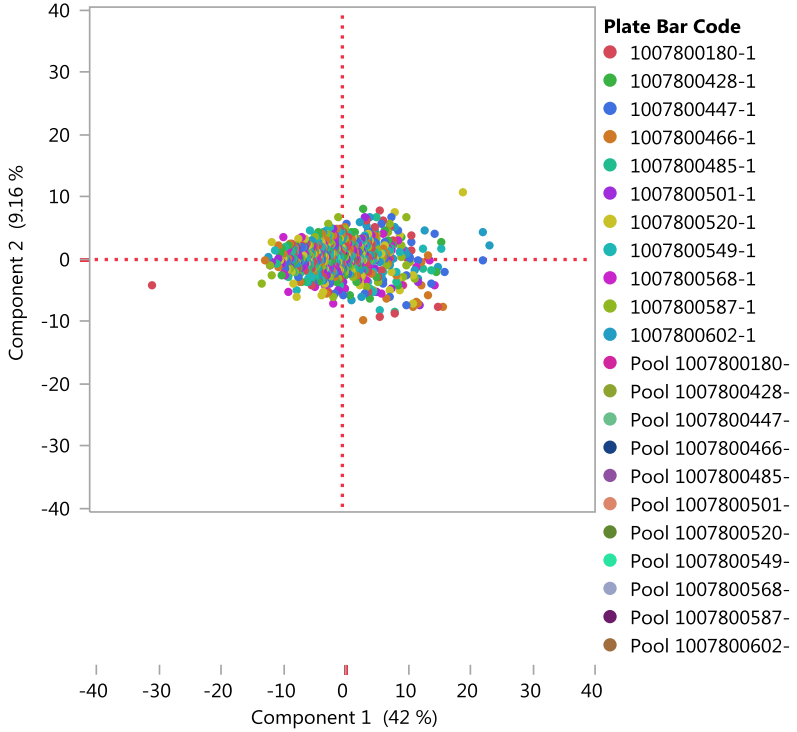


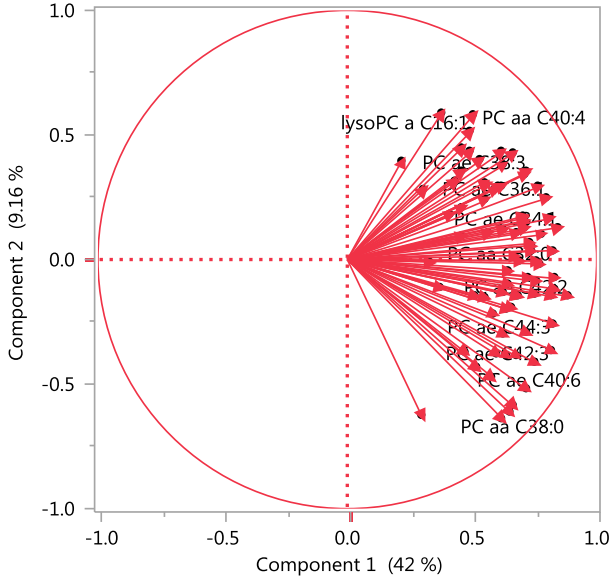


ADNI Alzheimer's Disease Neuroimaging Initiative

Principal Components Analysis of Glycerophospholipids

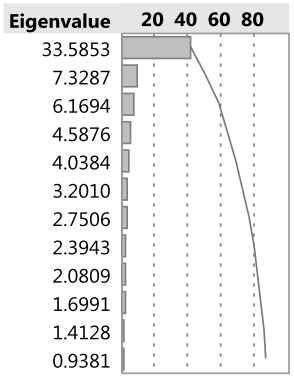


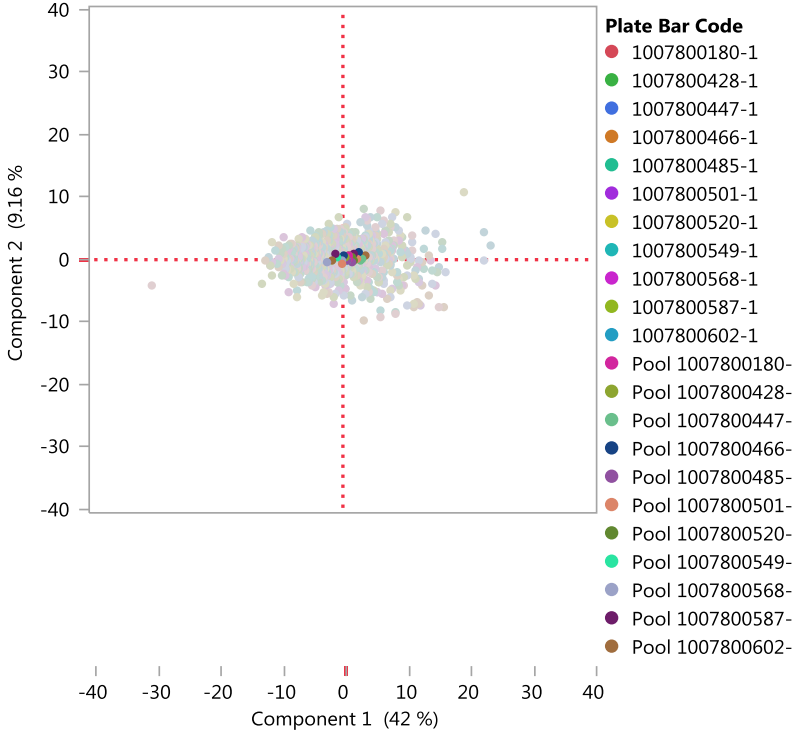


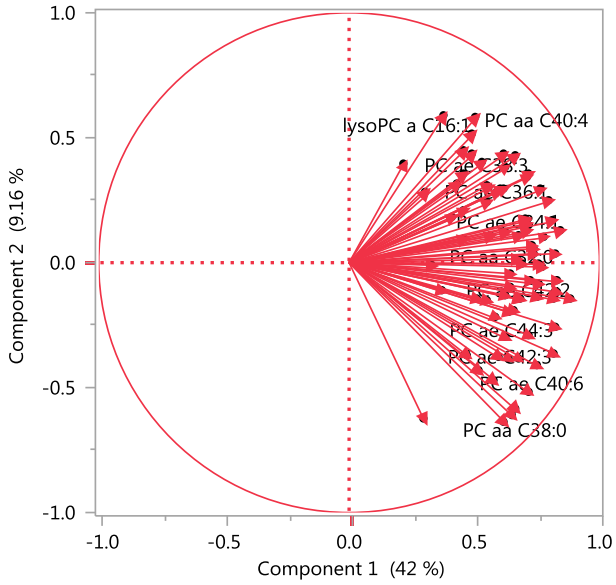


ADNI Alzheimer's Disease Neuroimaging Initiative

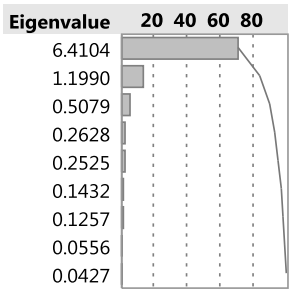
Principal Components Analysis of Glycerophospholipids (Pool Highlighted)

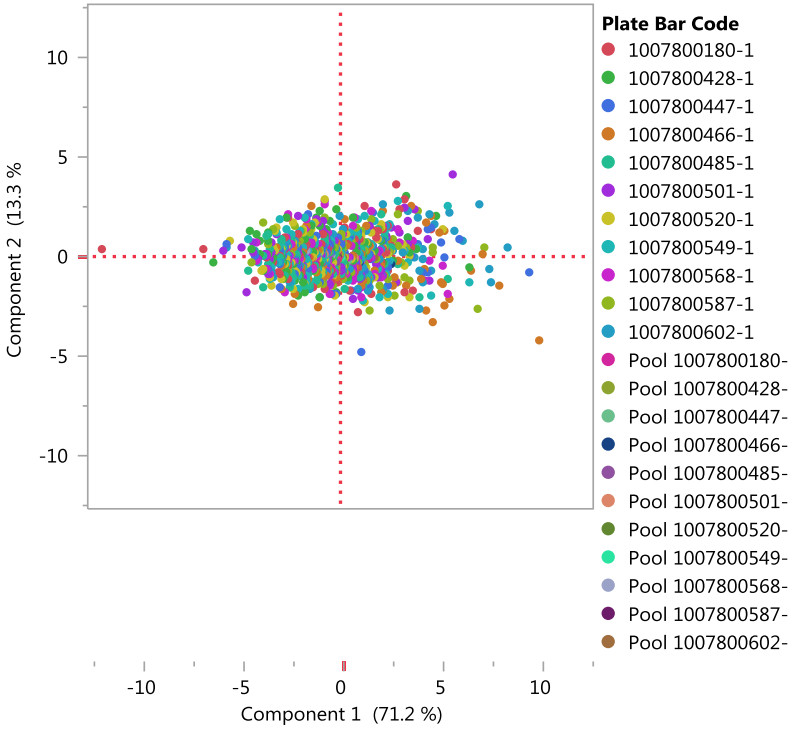


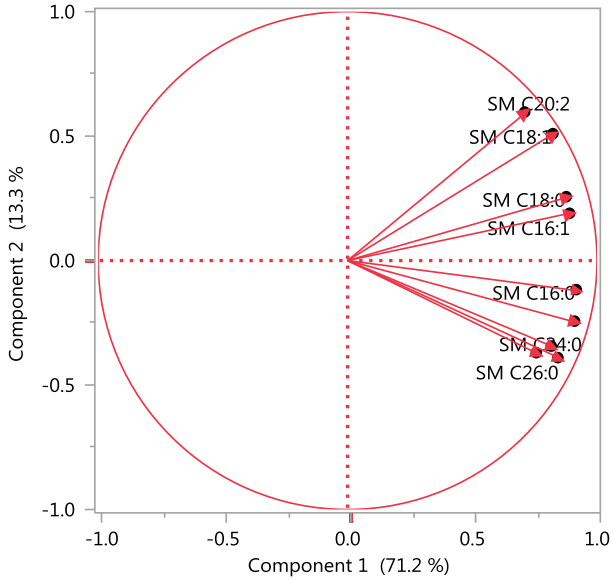




Principal Components Analysis of Sphingolipids







ADNI Alzheimer's Disease Neuroimaging Initiative

Principal Components Analysis of Sphingolipids (Pool Highlighted)