Metabolome in progression to Alzheimer's disease

Matej Orešič, *et al.*

Supplementary material

Contents

Supplementary Methods

Lipidomic analysis using UPLC-MS

The serum samples (10 µl) were mixed with 10 µl of 0.9% sodium chloride in Eppendorf tubes, spiked with a standard mixture consisting of 10 lipids (0.2 µg/sample; PC(17:0/0:0), PC(17:0/17:0), PE(17:0/17:0), PG(17:0/17:0), Cer(d18:1/17:0), PS(17:0/17:0), PA(17:0/17:0), MG(17:0/0:0/0:0), DG(17:0/17:0/0:0), TG(17:0/17:0/17:0)) and extracted with 100 µl of chloroform/methanol (2:1). After vortexing (2 min) and standing (1 h) the tubes were centrifuged at 10 000 rpm for 3 min. and 60 µl of the lower organic phase was separated and spiked with a standard mixture containing 3 labelled lipids (0.1 µg/sample; LPC(16:1/0:0-D_3), PC(16:1/16:1-D_6), TG(16:0/16:0/16:0-^{13}C3)).

Lipid extracts were analysed in a randomized order on a Waters Q-Tof Premier mass spectrometer combined with an Acquity Ultra Performance LC™ (UPLC; Waters, Milford, MA). The column (at 50°C) was an Acquity UPLC™ BEH C18 1 × 50 mm with 1.7 µm particles. The solvent system included 1) ultrapure water (1% 1M NH_4Ac, 0.1% HCOOH) and 2) LC-MS grade acetonitrile/isopropanol (5:2, 1% 1M NH_4Ac, 0.1% HCOOH). The gradient started from 65% A / 35% B, reached 100% B in 6 min and remained there for the next 7 min. There was a 5 min re-equilibration step before next run. The flow rate was 0.200 ml/min and the injected amount 1.0 µl (Acquity Sample Organizer; Waters, Milford, MA). Reserpine was used as the lock spray reference compound. The lipid profiling was carried out using ESI+ mode and the data was collected at mass range of m/z 300-1200 with scan duration of 0.2 sec. The data was processed by using MZmine 2 software[1] and the lipid identification was based on an internal spectral library.

Metabolomic analysis using GCxGC-TOFMS

Each serum sample (30 µl) was spiked with internal standard (20 µl labeled palmitic acid, c=258 mg/L) and the mixture was then extracted with 400 µl of methanol. After centrifugation the supernatant was evaporated to dryness and the original metabolites were then converted into their methoxime (MEOX) and trimethylsilyl (TMS) derivative(s) by two-step derivatization. First, 25 µl MOX reagent was added to the residue and the mixture was incubated for 60 min at 45 °C. Next, 25 µl MSTFA was added and the mixture was incubated for 60 min at 45 ºC. Finally, retention index standard mixture (n-alkanes) in hexane was added to the mixture.

For the analysis, a Leco Pegasus 4D GC×GC-TOFMS instrument (Leco Corp., St. Joseph, MI) equipped with a cryogenic modulator was used. The GC part of the instrument was an Agilent 6890 gas chromatograph (Agilent Technologies, Palo Alto, CA), equipped with split/splitless injector. The first-dimension chromatographic column was a 10 m RTX-5 capillary column with an internal diameter of 0.18 mm and a stationary-phase film thickness of 0.20 µm, and the second-dimension chromatographic column was a 1.5 m BPX-50 capillary column with an internal diameter of 100 µm and a film thickness of 0.1 µm. A DPTMS deactivated retention gap (3 m x 0.53 mm i.d.) was used in the front of the first column. High-purity helium was used as the carrier gas at a constant pressure mode (39.6 psig). A 5 s separation time was used in the second dimension The mass spectra was recorded by EI at 70 eV in the m/z range of 45 – 700 amu with 100 spectra/sec. Split injection (1 µl, split ratio 1:20) at 260 °C was used. The temperature program was as follows: the first-dimension column oven ramp began at 50 °C with a 1 min hold after which the temperature was programmed to 295 °C at a rate of 10 °C/min and then held at this temperature for 3 min. The second-dimension column temperature was maintained 20 °C higher than the corresponding first-dimension column. The programming rate and hold times were the same for the two columns.

The original GC×GC-TOFMS data included retention times, retention indices (RI), spectral information, possible identification, spectral similarity value (S=0-1000) and peak response data. The linear retention indexes were calculated based on the retention times of the compounds and the retention times of the retention index standards (*n*-alkanes). The data were transferred into an in-house developed program Guineu[2]. The alignment of the data was done based on the two retention times and spectra. After alignment of the GC×GC-TOFMS data, two filtration criteria were used for positive identification: 1) spectral match < 850 and 2) the RI_{exp}-RI_{lit} < 25 or or RI_{exp}-$RI_{std,\ exp}$ < 25, in which RI_{exp} is the experimental RI for a compound and RI_{lit} is the literature value for the identified compound and $RI_{std,\ exp}$ is the experimental RI value for a standard compound. Compounds not fulfilling the criteria were renamed to unknowns and were subjected to further identification.

The identification was based on the spectral search from NIST library or the in-house collected library and the retention indices. The compounds for which no spectral match was found by the two libraries, GOLM database (http://gmd.mpimp-golm.mpg.de) was utilized. The database was also used for functional group prediction of the metabolites.

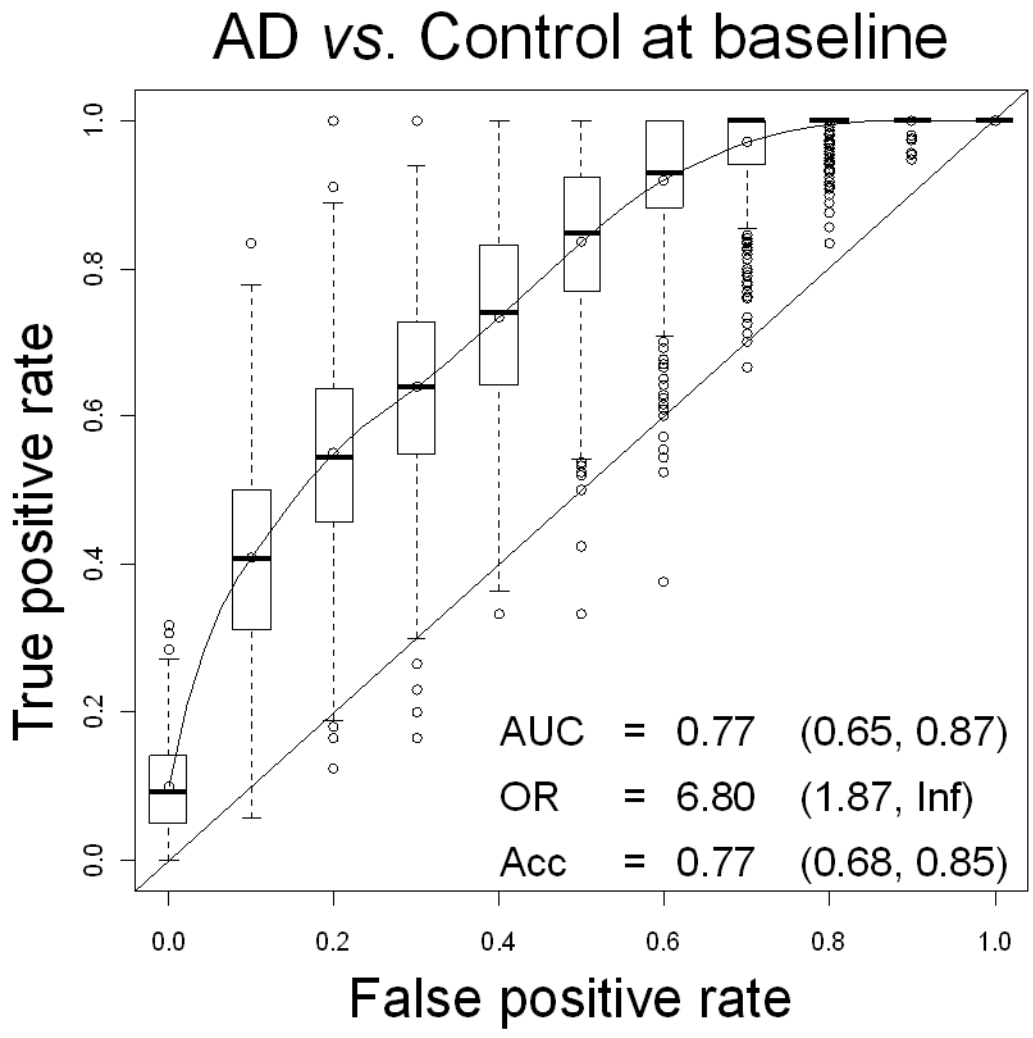
Calibration curves (six points, c= 15 -2000 ng, cholesterol c= 170-6000 ng) were constructed for the following compounds: pyruvic acid, alanine, 3-hydroxybutyric acid, valine, leucine, isoleucine, proline, glycine, succinic acid, fumaric acid, serine, threonine, malic acid, methionine, aspartic acid, alpha-ketoglutaric acid, phenylalanine, glutamic acid, ornithine, citric acid, tyrosine, palmitic acid, linoleic acid, oleic acid, stearic acid, arachidonic acid, and cholesterol.

Quality control of the method showed that the day-to-day repeatability of control serum samples, and the relative standard deviation (RSD) values for quantified compounds (amino acids and carboxylic acids) was on average below 10% , ranging from 3 to 17%. The repeatability was also studied for the same control serum samples using peak areas of the semiquantified compounds, and including all major peaks in the investigation (220 peaks). The average RSD for peak areas, utilizing TIC trace, was below 24 %. The internal standards added to all serum samples in the study had an average RSD under 10%.

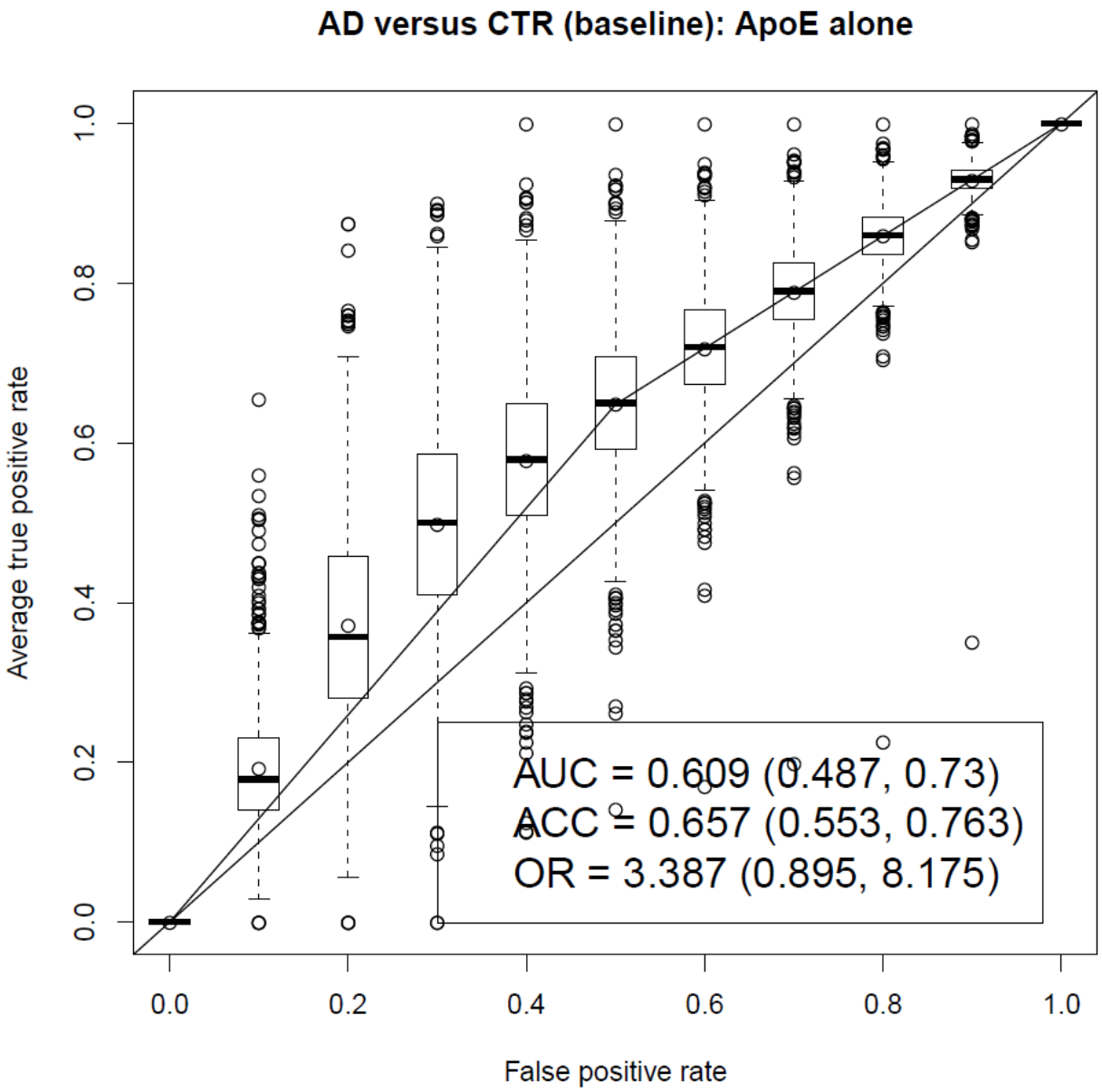
Supplementary Figure S1 Feasibility of diagnosing AD, based on concentrations of four metabolites (lactic acid, ketovaline, PC(18:0/18:2), PC(16:0/20:4)) in subjects at baseline diagnosed with AD or healthy controls.

The characteristics of the model (AUC, OR, Acc) independently tested in 1/3 of the sample are shown as mean values (5^{th}, 95^{th} percentiles), based on 2,000 cross-validation runs.

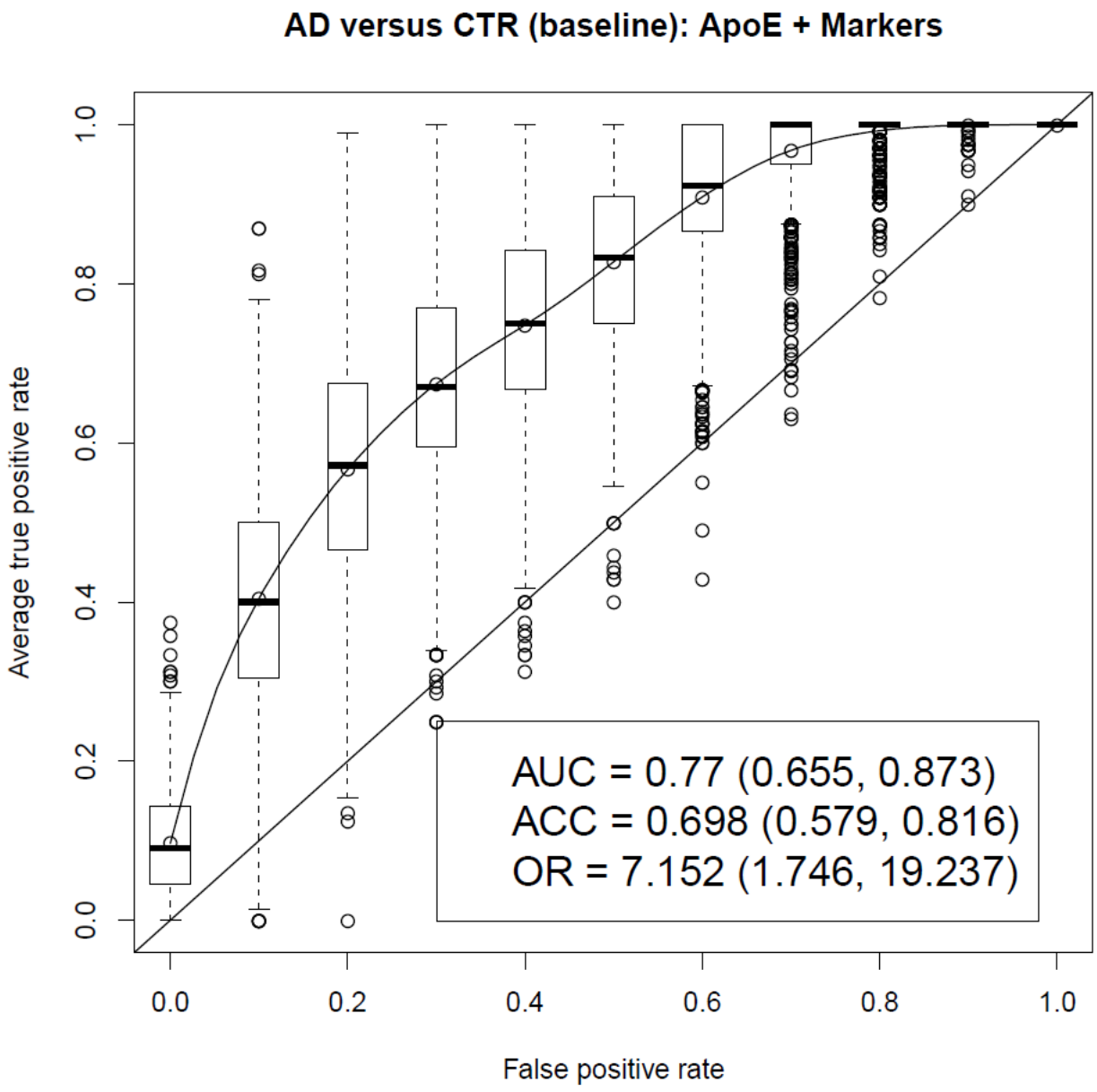
Acc = classification accuracy; AUC = area under the Receiver Operating characteristic (ROC) curve; OR = odds ratio.



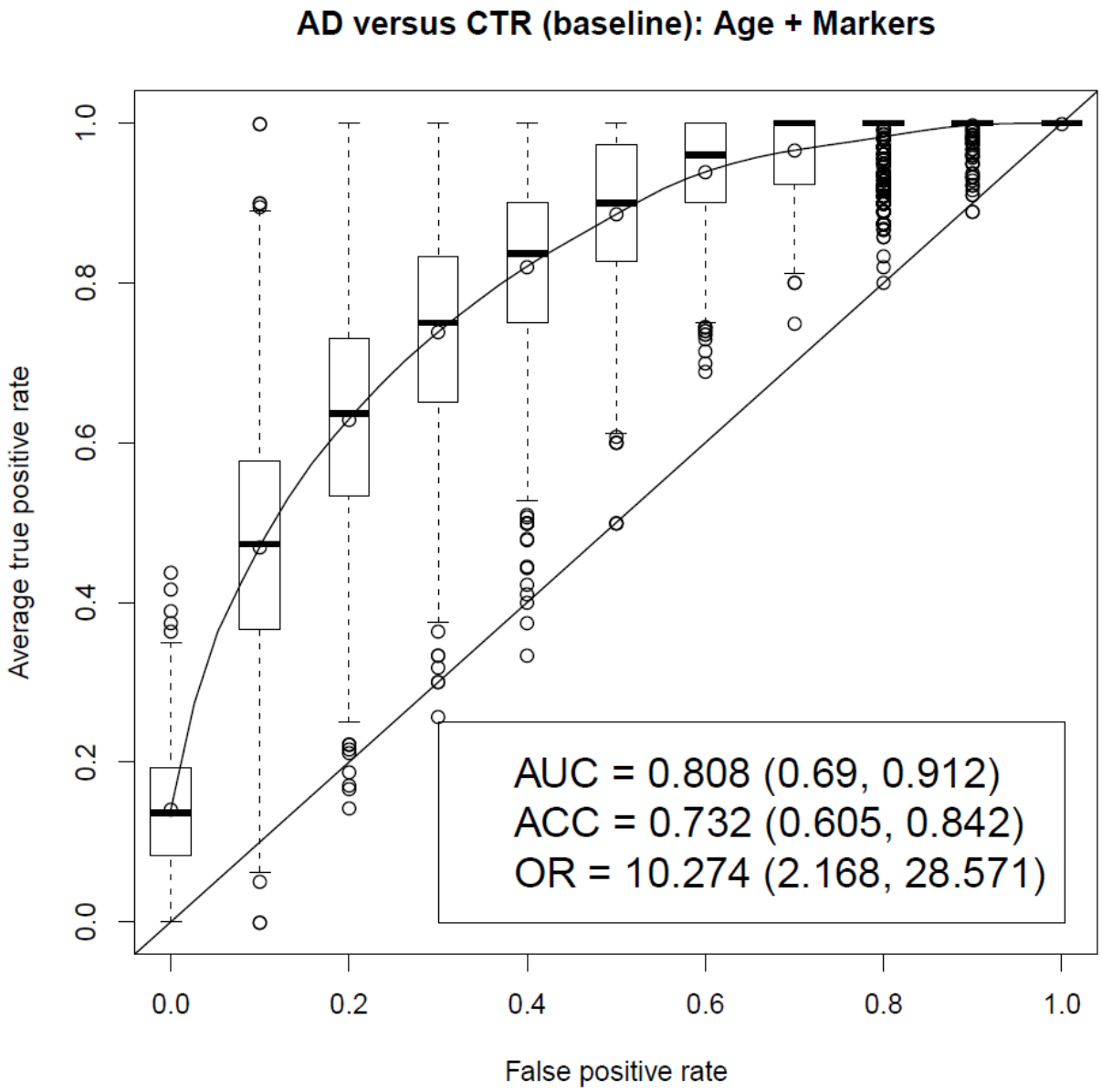
Supplementary Figure S2 Feasibility of diagnosing AD. ROC curve corresponding to the model based on ApoE genotype.



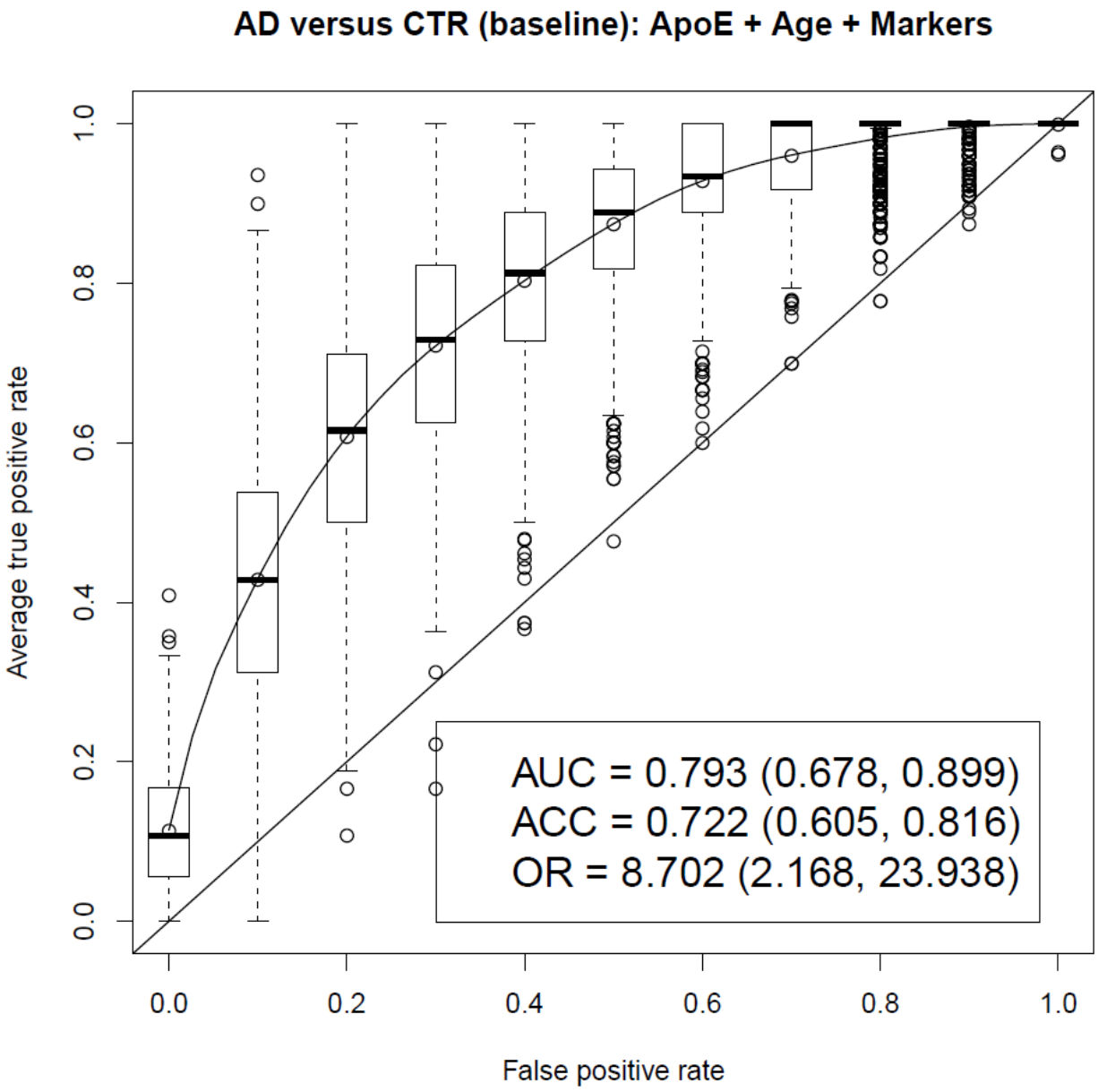
Supplementary Figure S3 Feasibility of diagnosing AD. ROC curve corresponding to the model based on concentrations of four metabolites (lactic acid, ketovaline, PC(18:0/18:2), PC(16:0/20:4)), and apoE genotype in subjects at baseline diagnosed with AD or healthy controls.



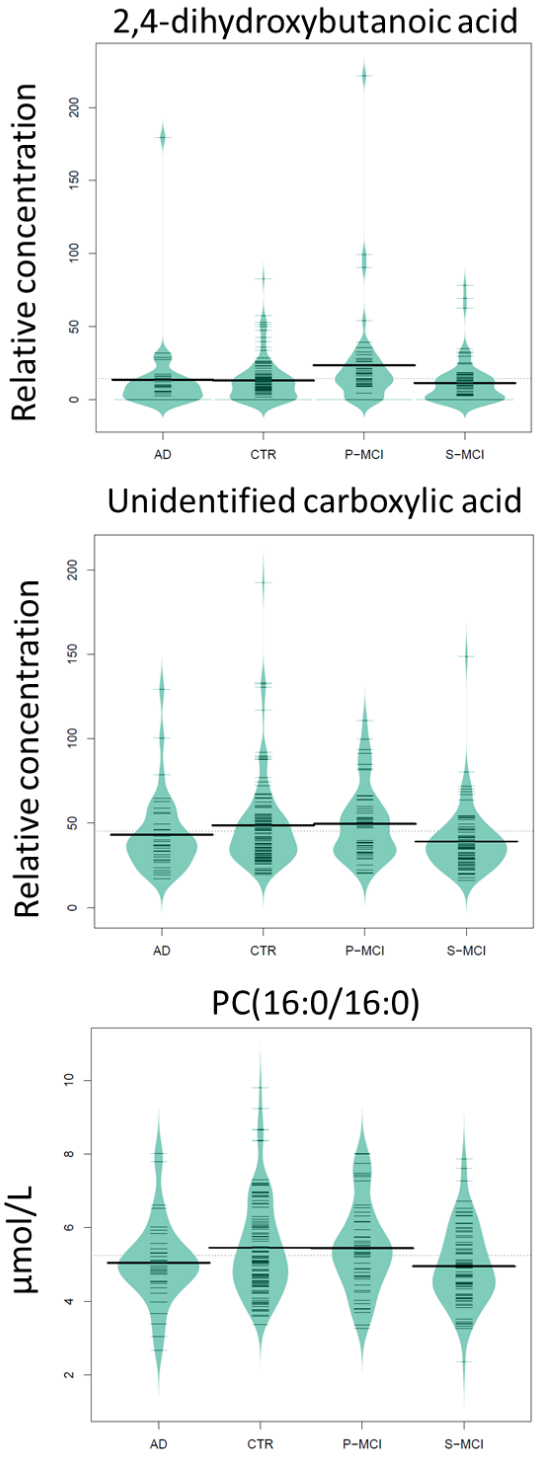
Supplementary Figure S4 Feasibility of diagnosing AD. ROC curve corresponding to the model based on concentrations of four metabolites (lactic acid, ketovaline, PC(18:0/18:2), PC(16:0/20:4)) and age, in subjects at baseline diagnosed with AD or healthy controls.



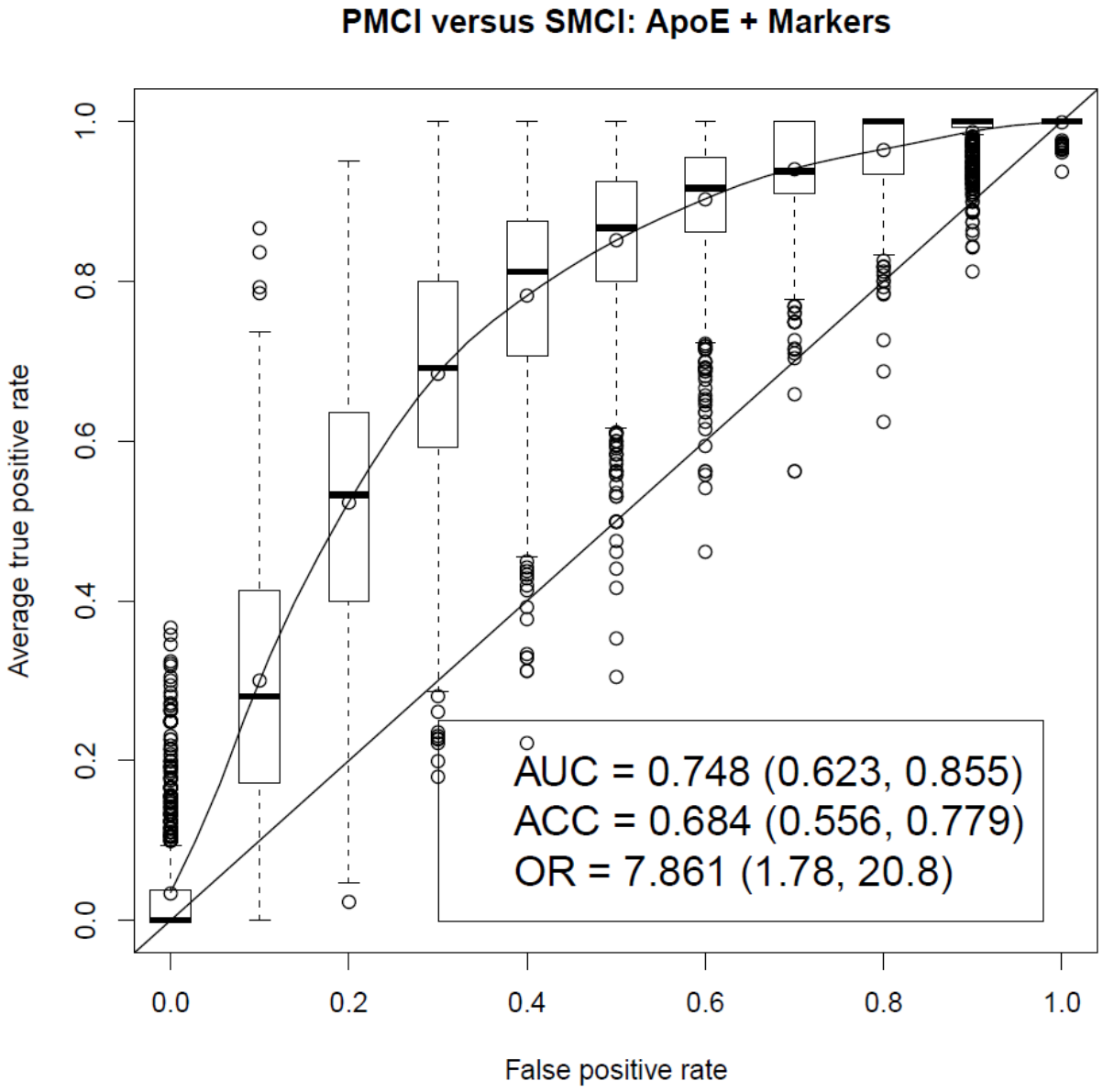
Supplementary Figure S5 Feasibility of diagnosing AD. ROC curve corresponding to the model based on concentrations of four metabolites (lactic acid, ketovaline, PC(18:0/18:2), PC(16:0/20:4)), apoE genotype and age, in subjects at baseline diagnosed with AD or healthy controls.



Supplementary Figure S6. Beanplots of the three metabolites included in the predictive model across four of the patient groups included in the study (Table 1).



Supplementary Figure S7 Feasibility of predicting AD. ROC curve corresponding to the model based on concentrations of three metabolites (2,4-dihydroxybutanoic acid, unidentified carboxylic acid, PC(16:0/16:0)) and ApoE genotype in subjects at baseline who were diagnosed with MCI.



References

1. Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 2010; **11**(1)**:** 395.

2. Castillo S, Mattila I, Miettinen J, Orešič M, Hyötyläinen T. Data analysis tool for comprehensive two-dimensional gas chromatography-time of flight mass spectrometry *Anal Chem* 2011; **83**(8)**:** 3058–3067.