

Multilevel Data Analysis of a Crossover Designed Human Nutritional Intervention Study

Ewoud J. J. van Velzen,^{†,§} Johan A. Westerhuis,^{*,†} John P. M. van Duynhoven,[§]
 Ferdi A. van Dorsten,[§] Huub C. J. Hoefsloot,[†] Doris M. Jacobs,[§] Suzanne Smit,[†]
 Richard Draijer,[§] Christine I. Kroner,[§] and Age K. Smilde[†]

Biosystems Data Analysis, Swammerdam Institute for Life Sciences, Universiteit van Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands, Unilever Food and Health Research; Institute, Olivier van Noortlaan 120, 3133 AT Vlaardingen, The Netherlands

Received February 21, 2008

A new method is introduced for the analysis of 'omics' data derived from crossover designed drug or nutritional intervention studies. The method aims at finding systematic variations in metabolic profiles after a drug or nutritional challenge and takes advantage of the crossover design in the data. The method, which can be considered as a multivariate extension of a paired *t* test, generates different multivariate submodels for the between- and the within-subject variation in the data. A major advantage of this variation splitting is that each submodel can be analyzed separately without being confounded with the other variation sources. The power of the multilevel approach is demonstrated in a human nutritional intervention study which used NMR-based metabolomics to assess the metabolic impact of grape/wine extract consumption. The variations in the urine metabolic profiles are studied between and within the human subjects using the multilevel analysis. After variation splitting, multilevel PCA is used to investigate the experimental and biological differences between the subjects, whereas a multilevel PLS-DA model is used to reveal the net treatment effect within the subjects. The observed treatment effect is validated with cross model validation and permutations. It is shown that the statistical significance of the multilevel classification model ($p \ll 0.0002$) is a major improvement compared to an ordinary PLS-DA model ($p = 0.058$) without variation splitting. Finally, rank products are used to determine which NMR signals are most important in the multilevel classification model.

Keywords: ANOVA • ASCA • cross model validation • metabolomics • MSCA • NMR • nutrition • paired data analysis • PCA • permutation test • PLS-DA • rank product • urine

Introduction

One of the major challenges in nutritional metabolomic studies is the detection and identification of metabolites in different biofluids that can be linked to the human nutrition metabolome.^{1–3} In nutritional intervention studies, these metabolic responses are often small and subtle since the volunteers are generally healthy and in metabolic homeostasis.^{4,5} Moreover, the effects of the nutritional treatment tend to be much smaller than the biological variation that exists between the individuals.^{3,6–8}

Taking advantage of the design of an experiment and the underlying data structure can be useful to uncover minor treatment effects.^{9,10} Since most traditional multivariate data analysis techniques such as Principal Component Analysis (PCA), Partial Least Squares (PLS) and Partial-Least Squares-Discriminant Analysis (PLS-DA) do not take the experimental design into account, the power of these multivariate methods is not fully employed. Particularly, multivariate methods that

can optimally exploit the paired data structure in crossover designed studies are lacking, even though this experimental design is rather common in nutritional intervention trials.^{11,12} The particular strength of the crossover design is that treatments (interventions) are evaluated on the same subjects, allowing comparison at the individual level rather than on the group level.¹³

A specific limitation of using PCA and PLS (-DA) in crossover designed experiments is that the net treatment effect is not separated from the biological variation between the subjects. As a result, subtle treatment effects within the subjects are often largely overwhelmed by the strong biological variation between subjects.^{14,15} Recently, a combination of Analysis-of-Variance (ANOVA) and Simultaneous Component Analysis (SCA) was introduced that enables the analysis of metabolomic studies with an experimental design.^{9,16,17} The basic principle of ANOVA-SCA (ASCA) is the variation splitting property of ANOVA which allows a separate analysis and interpretation of the variation sources induced by the different factors in the experimental design. A special case of ASCA is Multilevel Simultaneous Analysis (MSCA).^{10,16} MSCA takes into account the multilevel structure in the data, and is particularly suitable

* To whom correspondence should be addressed. Phone: +31 20 525 6546. Fax: +31 20 525 6971. E-mail: j.a.westerhuis@uva.nl.

[†] Universiteit van Amsterdam.

[§] Unilever Vlaardingen.

for the analysis of temporal and longitudinal crossover studies. When only a single factor is regarded in the experimental design, MSCA decomposes the data into an offset term, a between-subject part and a within-subject part.^{9,18}

In the present study, the variation splitting property of MSCA is applied to a crossover designed human nutritional intervention study in which the metabolic impact of grape/wine extract consumption on the urinary ¹H NMR profiles is evaluated. To find systematic differences among the intervention groups, multilevel PLS-DA is performed on the within-subject data. To investigate the underlying variation in the between-subject data, multilevel PCA or MLCA¹⁸ is used. The combination of multilevel data analysis and PLS-DA is introduced as a new multivariate approach to investigate treatment effects in crossover designed experiments.

To examine whether the use of multilevel PLS-DA indeed results in an improved multivariate solution, its performance is benchmarked against the ordinary PLS-DA approach. In this assessment, the prediction error of the PLS-DA models are determined by means of cross model validation,¹⁹ and compared with the prediction errors of permuted data.²⁰ The combination of multilevel PLS-DA, cross model validation, and permutation testing finally allows the selection and interpretation of candidate biomarkers that can be linked with the intended treatment effect.

Experimental Section

Study Protocol. The study was carried out in the Consumer Centre at Unilever R&D, Vlaardingen, The Netherlands. The study had a double-blind, placebo-controlled crossover design with two treatments, a 2-week run-in period, and two 4-week consecutive intervention periods. The two treatments consist of a placebo (Avicel PH101cellulose; FMC Biopolymer, PA) and a mix of wine extract (Provinols, Seppic, France) and grape juice extract (MegaNatural, Polyphenolics, CA). The mixture of wine and grape juice extract contained 800 mg of polyphenols based on gallic acid equivalents,²¹ and corresponds to approximately 80% of the normal daily intake of polyphenols (~1 g/day).²² The extracts and the placebo were given daily, and administered as nontransparent capsules.

In total, 29 male and female human subjects in the age range of 35–75 years and mildly hypertensive (systolic blood pressure, 130–179 mmHg; diastolic blood pressure, < 100 mmHg) participated in the study. Urine was collected in 2-L urine collection vessels over a time span of 24 h after the interventions, and stored at 193 K before use. The vessels were supplied with a gelatinized layer of metaphosphoric acid (MFA, 15.5 g) to stabilize the phenolic compounds and to inhibit bacterial growth.

The subjects were asked to refrain from vitamin-, mineral-, and other supplementation, as well as plant sterol or plant stanol containing food. Furthermore, each of the volunteers was requested to follow a similar dietary and lifestyle pattern for the duration of the study. The protocol was approved by the medical ethical committee of the University of Wageningen and conducted in accordance with the ICH-GCP guidelines for Good Clinical Practise (ICH GCP, 1996).

Sample Pretreatment. Urine samples were allowed to thaw at room temperature. To 450 μ L of each urine, 200 μ L of phosphate buffer solution (0.1 M Na₂HPO₄/0.1 M NaH₂PO₄) and 50 μ L of deuterium oxide (D₂O) were added to adjust the pH to 3.0 \pm 0.2. The phosphate buffer solution furthermore contains 0.01 mg/mL 3-(trimethylsilyl)propionic acid-*d*₄ so-

dium salt (TSP) as a chemical shift reference. After homogenization, the sample was centrifuged at 10 000 rpm for 5 min. Then, 650 μ L of the clear urine supernatant was transferred into a 2 mL amber vial and closed with a screw cap.

¹H NMR Data Acquisition. One-dimensional ¹H NMR spectra were acquired on a 600 MHz Bruker Avance NMR spectrometer, equipped with a 60 μ L flow probe and a Gilson 215 auto sampler. After transferring 60 μ L of each sample into the flow cell, 1D ¹H NMR spectra were acquired with presaturation of the water resonance using a noesy1dpr pulse sequence RD-90°-*t*₁-90°-*t*_{mix}-90°-FID (Bruker Biospin, Germany). Here, *t*₁ is a 4 μ s delay time, and *t*_{mix} is the mixing time (150 ms). The FIDs were collected into 32K points (128 scans) with a spectral width of 9000 Hz, an acquisition time of 1.82 s, and a relaxation delay of 3 s. During acquisition, the temperature was kept constant at 300 K. The measurements were carried out in random order in two separate NMR runs, consisting of 46 and 12 measurements, respectively. The urine samples from each subject were kept together within the same NMR run, and measured within 12 h after preparation.

¹H NMR Data Processing. An exponential window function was applied to the free induction decay (FID) with a line-broadening factor of 0.5 Hz prior to the Fourier transformation. The Fourier transformed NMR spectra were manually phase and baseline-corrected and calibrated against the TSP methyl resonance at δ 0.0 ppm. The NMR spectra were subdivided into 550 discrete regions ('buckets') of equal width ($\Delta\delta = 0.02$ ppm), from which the integrated areas were determined using AMIX (Analysis of Mixtures, Bruker GmbH, Germany). The spectral region between δ 4.3–5.2 ppm was excluded from the data set to avoid spectral interference of residual water. To compensate for dilution effects, the urine profiles were normalized to the integral of the methyl resonance of creatinine (δ 3.06–3.18 ppm).^{23,24} In this approach, creatinine clearance is considered constant in the group of participating volunteers. As a result, the excreted amount for each metabolite can be expressed as mol/mol creatinine present in the urine samples. Metabolites were identified using a database including reference spectra of metabolites at different pH values (bbiorefcode-2-0-0 implemented in Amix 3.7.3., Bruker Biospin GmbH).

Data Analysis. Data pretreatment, MSCA, PLS-DA, cross model validation and permutation tests were performed using Matlab (version 2008a, The MathWorks) and in-house written Matlab routines. These routines (together with a tutorial) are available via the Internet at <http://pubs.acs.org> or <http://www.bdagroup.nl/>.

Mathematical Methods

Structure of the NMR Data Set. The structure of the data set is shown in Figure 1. The following indices are used: *j* = 1, ..., *J* for the number of spectral variables, *k* = 1, ..., *K* for the number of intervention occasions, and *i* = 1, ..., *I* for the number of subjects. The total number of NMR spectra in the data set is represented by *L* ($= I \times K$).

Normalization. In the ¹H NMR data-matrix **D**, with dimensions *L* \times *J*, *L* is the number of spectra and *J* is the number of variables (chemical shifts). Each of the observed signal-intensities *d*_{*ij*} is linearly related to the concentration level of a urine metabolite in spectrum *l*. Since the excreted volumes of the 24 h urine are different for each subject, direct comparison of the urinary metabolite compositions is not possible. A commonly used procedure to compensate for differences in the excreted urine volumes is creatinine normalization.^{23–25} In this

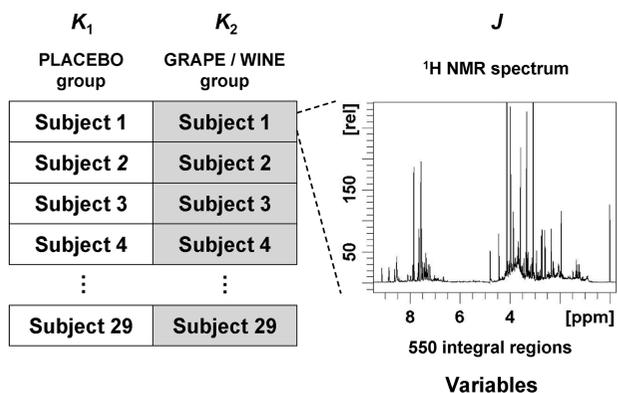


Figure 1. Structure of the data set. K is the number of interventions in the study for each subject. Each subject consist of a paired data set ($K = 2$) comprising the urinary NMR spectra of the placebo group (K_1) and the grape/wine treatment group (K_2). At the right-hand side, a typical urinary NMR spectrum is presented after placebo intervention and consists out of J variables (550 signal intensities along the NMR chemical shift axis).

procedure, the peak area of the creatinine methyl resonance at δ 3.12 ppm (A_l) is used to normalize the raw NMR signals intensities (d_{ij}) in each spectrum l according to eq 1:

$$x_{ij} = \frac{d_{ij}}{A_l} \quad (1)$$

where x_{ij} is the normalized NMR signal-intensity d_{ij} in spectrum l and A_l is the peak area of the creatinine methyl resonance in spectrum l .

Split-Up of Variation. In the normalized data matrix \mathbf{X} , each NMR variable is denoted by the scalar x_{ikj} . When separating the variation in x_{ikj} into a between-subject term (biological variation) and a within-subject term (nutritional variation), an analysis of variance (ANOVA) model can be defined according eq 2:

$$x_{ikj} = \mu_j + \beta_{ij} + \gamma_{ikj} + \varepsilon_{ikj} \quad (2)$$

where μ_j is the population mean (offset term); β_{ij} is the between-subject variation; γ_{ikj} is the within-subject variation and ε_{ikj} is the variation due to experimental noise.^{26,27}

As the currently investigated data set does not consist of repeated measurements, the contribution of the variation term ε_{ikj} cannot be estimated. Decomposition of the remaining variation terms is performed according to two consecutive centring steps, and is similar to the variation splitting procedures in ASCA^{9,16} and (Multilevel) Simultaneous Component Analysis.^{10,18} As given in eq 3, the first centring step is applied on the entire data set \mathbf{X} and results in a mean-centered data block (\mathbf{X}_c) and an offset term ($\mathbf{1}_L \mathbf{x}_m^T$). The offset \mathbf{x}_m is an estimation of μ , and contains the mean signal intensities of all spectral variables j in \mathbf{X} .

$$\mathbf{X} = \mathbf{1}_L \mathbf{x}_m^T + \mathbf{X}_c \quad (3)$$

In eq 3, \mathbf{X}_c ($L \times J$) contains the mean-centered data; $\mathbf{1}_L$ ($L \times 1$) contains ones, and \mathbf{x}_m^T ($1 \times J$) contains the mean values for each column in \mathbf{X} . Concatenating the L row vectors, \mathbf{x}_m^T results in a matrix $\mathbf{1}_L \mathbf{x}_m^T$ ($L \times J$). This matrix can be defined as the offset term \mathbf{X}_m .

On the mean-centered data (\mathbf{X}_c), a second centring step is performed according to eq 4. At this point, mean-centring is applied per subject i over the K intervention occasions. The

mean for subject i (\mathbf{x}_{bi}) is referred to the between-subject term and is different for each subject. The net variation around the mean is described by the within-subject term \mathbf{X}_{wi} and is different for all observations in the data set. The variation between and within the subjects are estimations of β and γ , respectively.

$$\mathbf{X}_{ci} = \mathbf{1}_K \mathbf{x}_{bi}^T + \mathbf{X}_{wi} \quad (4)$$

Here, \mathbf{X}_{ci} ($K \times J$) contains the mean-centered data for subject i and is part of matrix \mathbf{X}_c ; $\mathbf{1}_K$ ($K \times 1$) contains ones; \mathbf{x}_{bi}^T ($1 \times J$) contains the mean values of each column in \mathbf{X}_{ci} and \mathbf{X}_{wi} ($K \times J$) contains the mean-centered data per subject i .

When concatenating $\mathbf{1}_K \mathbf{x}_{bi}^T$ for each individual into \mathbf{X}_b ($L \times J$) and concatenating the I matrices \mathbf{X}_{wi} into \mathbf{X}_w ($L \times J$), eqs 3 and 4 can alternatively be written in the matrix form (eq 5):

$$\mathbf{X} = \mathbf{X}_m + \mathbf{X}_b + \mathbf{X}_w \quad (5)$$

where \mathbf{X}_m contains the offset; \mathbf{X}_b contains the between-subject variations, and \mathbf{X}_w contains the within-subject variations.

On the basis of the usual constraints for the ANOVA model,¹⁶ it has been proven that the column spaces of the independent submodels are orthogonal to each other.^{9,10,16,17} Consequentially, the magnitudes of the different sources of variation can be calculated, and estimated in sum of squares according to eq 6:

$$\|\mathbf{X}\|^2 = \|\mathbf{X}_m\|^2 + \|\mathbf{X}_b\|^2 + \|\mathbf{X}_w\|^2 \quad (6)$$

where $\|\mathbf{X}_b\|^2$ and $\|\mathbf{X}_w\|^2$ can be used to determine the percentages of the between-subject variation and the within-subject variation in the data.

Scaling. Since PCA and PLS-DA are scale-dependent methods, the use of an appropriate scaling technique is essential to consider medium and small features in the spectral data as important as the large features. Among the different scaling techniques, Pareto-scaling and Unit Variance (UV) scaling are the most frequently used ones to NMR data.²⁸⁻³¹

An important disadvantage with UV-scaling has been concerned with the increase of spectral noise, which may lead to severe overfitting problems in PCA and PLS-DA. Pareto scaling compromises between the extremes of no-scaling and UV-scaling because the square root of the standard deviation is used as the scaling factor. In PCA, this will lead to an improved interpretability of the spectral loadings as it keeps the data structure partially intact.²⁸

Multivariate Data Analysis. Prior to the multivariate data analysis, the between-subject data \mathbf{X}_b is Pareto-scaled to facilitate analysis of the major effects in the data.²⁸ The Pareto-scaling procedure for a particular NMR signal j with intensity x_{bij} in (between-subject) spectrum l is given in eq 7:

$$\tilde{x}_{bij} = \frac{x_{bij} - \bar{x}_{bj}}{\sqrt{s_{bj}}} \quad (7)$$

where, \tilde{x}_{bij} is the intensity of the Pareto-scaled NMR signal j ; \bar{x}_{bj} is the mean signal intensity of signal j of \mathbf{x}_b , and s_{bj} is the standard deviation of signal j of \mathbf{x}_b . Concatenation of the Pareto-scaled NMR signals \tilde{x}_{bij} result in $\tilde{\mathbf{X}}_b$ ($L \times J$) and represents the Pareto-scaled between-subject data.

Multilevel Principal Component Analysis (PCA)³² is then used to explore the underlying systematic variations in $\tilde{\mathbf{X}}_b$. The multilevel model for analyzing the data matrix $\tilde{\mathbf{X}}_b$ is given in eq 8,

$$\tilde{\mathbf{X}}_{\mathbf{b}} = \mathbf{T}_{\mathbf{b}}\mathbf{P}_{\mathbf{b}}^T + \mathbf{E}_{\mathbf{b}} \quad (8)$$

where $\mathbf{T}_{\mathbf{b}}$ ($JK \times R_{\mathbf{b}}$) contains the between-subject scores. Here $R_{\mathbf{b}}$ is the number of principal components of the between-subject PCA model. $\mathbf{P}_{\mathbf{b}}^T$ ($R_{\mathbf{b}} \times J$) contains the between-subject loadings, and $\mathbf{E}_{\mathbf{b}}$ ($L \times J$) contains the residuals of the between-individual model.

Similar to multilevel PCA, multilevel PLS-DA can be used to decompose the scaled within-subject data ($\tilde{\mathbf{X}}_{\mathbf{w}}$). However, to consider small peaks as important as the largest peaks in the data, UV-scaling is applied instead of Pareto-scaling. The UV-scaling procedure for a particular NMR signal j with intensity x_{wjl} in (within-subject) spectrum l is given in eq 9:

$$\tilde{x}_{wjl} = \frac{x_{wjl} - \bar{x}_{wj}}{s_{wj}} \quad (9)$$

where \tilde{x}_{wjl} is the intensity of the UV-scaled NMR signal j ; \bar{x}_{wj} is the mean signal intensity of signal j of $\mathbf{x}_{\mathbf{w}}$, and s_{wj} is the standard deviation of signal j of $\mathbf{x}_{\mathbf{w}}$. Concatenation of the UV-scaled NMR signals \tilde{x}_{wjl} result in $\tilde{\mathbf{X}}_{\mathbf{w}}$ ($L \times J$) and represents the UV-scaled within-subject data.

In a multivariate regression method like PLS-DA,³³ the response variable \mathbf{y} (class labels) is used to guide the projections into meaningful directions, and provide information on the relationship between $\tilde{\mathbf{X}}_{\mathbf{w}}$ and \mathbf{y} according to eq 10:

$$\mathbf{y} = \tilde{\mathbf{X}}_{\mathbf{w}}\mathbf{b} + \mathbf{e} \quad (10)$$

where \mathbf{b} ($J \times 1$) is the regression coefficient vector, \mathbf{e} ($L \times 1$) contains the \mathbf{y} model residuals, and \mathbf{y} ($L \times 1$) contains the class labels, which is equal for all subjects at intervention occasion k . The \mathbf{y} values of the subjects in the placebo group (k_1) are assigned to class -1 , while the treatment group (k_2) is assigned to class $+1$. The PLS NIPALS algorithm³³ is used to calculate the regression coefficient vector \mathbf{b} .

Cross Model Validation. Validation of PLS-based classification models is essential as it has been identified that PLS (-DA) in metabolomic applications are prone to serious modeling and validation problems.^{4,34,35} Particularly in metabolomic data where the number of subjects is usually much smaller than the number of variables, the obtained classification models are extremely susceptible to overfitting and chance classifications.^{34,36} To test the original and the multilevel PLS-DA model against overfitting, the classification error of these models is estimated according to a cross model validation (CMV).^{19,37,38}

CMV is a resampling scheme in which the data set is randomly split into a test set, a validation set and a training set. The training set and the validation set are used to establish the optimal model parameters (i.e., the number of PLS components and variable selection), whereas the test set is used to determine the true prediction error of the model. The test samples are left out of the model optimization and are therefore representative for the prediction of new, unseen subjects.

In the model optimization, repeated calibrations are made with different training and validation sets using a single cross validation (1CV) procedure.^{19,20,36} On the basis of the prediction results of the validation set samples in the 1CV, the number of PLS components as well as variable selection is optimized. When a multilevel PLS-DA model is built from the training set, the entire variation splitting procedure is performed. The CMV should therefore be constrained to an adapted resampling scheme, keeping the paired data structure in the training set, the validation set and the test set intact for each individual. As a result, complete individuals are left out of the training set.

Except for variation splitting, the CMV for the multilevel PLS-DA model is similar to the CMV for the original PLS-DA model.

For class prediction of a new individual i (in the test set), $\mathbf{X}_{\mathbf{ci}}^{\text{new}}$ ($K \times J$) with placebo and treatment spectra, is first corrected for the offset $\mathbf{x}_{\mathbf{m}}$ of the training set, according to eq 11:

$$\mathbf{X}_{\mathbf{ci}}^{\text{new}} = \mathbf{X}_{\mathbf{ci}}^{\text{new}} - \mathbf{1}_K\mathbf{x}_{\mathbf{m}}^T \quad (11)$$

where $\mathbf{X}_{\mathbf{ci}}^{\text{new}}$ ($K \times J$) contains the mean-centered data for a new individual i ; $\mathbf{1}_K$ ($K \times 1$) contains ones, and $\mathbf{x}_{\mathbf{m}}^T$ ($1 \times J$) contains the mean values for each column in the training set \mathbf{X} .

Then between-subject data ($\mathbf{x}_{\mathbf{bi}}^{\text{new}})^T$ ($1 \times J$) for a new individual i is calculated as the mean of $\mathbf{X}_{\mathbf{ci}}^{\text{new}}$ over the K intervention occasions. The net variation around the mean is described by the within-subject term $\mathbf{X}_{\mathbf{wi}}^{\text{new}}$ ($K \times J$) for new data. Thus, the variation in the spectral data obtained from a new individual i can be split according to eq 12:

$$\mathbf{X}_{\mathbf{ci}}^{\text{new}} = \mathbf{1}_K\mathbf{x}_{\mathbf{m}}^T + \mathbf{1}_K(\mathbf{x}_{\mathbf{bi}}^{\text{new}})^T + \mathbf{X}_{\mathbf{wi}}^{\text{new}} \quad (12)$$

Before class prediction, $\mathbf{X}_{\mathbf{wi}}^{\text{new}}$ needs to be scaled according to the mean (\bar{x}_{wj}) and standard deviation (s_{wj}) of the training set $\mathbf{X}_{\mathbf{w}}$. The scaling procedure for a NMR signal j in (within-subject) spectrum l is given in eq 13:

$$\tilde{x}_{wjl}^{\text{new}} = \frac{x_{wjl}^{\text{new}} - \bar{x}_{wj}}{s_{wj}} \quad (13)$$

where $\tilde{x}_{wjl}^{\text{new}}$ is the UV-scaled NMR signal j with intensity x_{wjl}^{new} in a new (within-subject) spectrum l ; \bar{x}_{wj} is the mean signal intensity of signal j in the training set, and s_{wj} is the standard deviation of signal j in the training set.

Together with the estimated regression coefficient \mathbf{b} (eq 10), the class predictions ($\hat{\mathbf{y}}_{\text{new}}$) of new samples can be calculated according to eq 14:

$$\hat{\mathbf{y}}_{\text{new}} = \tilde{\mathbf{X}}_{\mathbf{wi}}^{\text{new}}\mathbf{b} \quad (14)$$

Rank Product. To be able to select the most discriminative spectral variables in the classification model, use was made of the rank product.^{20,39} In a rank product (RP), all J variables in the NMR spectrum (chemical shifts) are ranked according to their PLS regression coefficients \mathbf{b} . The largest absolute value is assigned to rank 1, the second largest value to rank 2, and so forth.

In the currently applied validation procedure, CMV was repeated 20 times with a different selection of validation-, training-, and test samples. As a result, also 20 regression coefficient vectors (\mathbf{b}) were obtained. The variables were ranked according their absolute size in regression coefficient and multiplied to obtain a final RP. Since the data set consist of 550 spectral variables, the average rank product for a given variable is approximately $(550/2)^{20} = 6.10^{48}$. The spectral variables with the lowest rank products are the ones with the largest discriminative power.

Variable Selection. To select the most discriminative variables for a PLS-DA classification model, variable selection is integrated in the 1CV procedure.¹⁹ In 5 sequentially performed variable selection steps, respectively, 550, 275, 138, 69, and 35 spectral variables with the lowest ranks are selected for model optimization. In each of these steps, a new RP is calculated, from which only half of the variables is again used in the consecutive variable selection step. On the basis of the 1CV prediction results, obtained in the 5 variable selection steps, the optimum number of variables as well as the optimum

number of PLS components are determined. Finally, the optimized model parameters are used to predict the class membership of the independent samples in the test set.

Validation of the Treatment Effect. To evaluate the statistical significance of the treatment effect, a permutation test was performed.^{20,36,40} In a permutation test, the class labels (y) are randomly permuted, which implies that the order of the placebo and treatment intervention for each subject will be randomly assigned. The basic principle of the permutation test is that randomization of the intervention class labels will lead to poor classification models that should not be able to distinguish between the placebo group and the treatment group. Permutation of the class labels should in theory lead to an average number of 29 misclassifications (NMC), which is exactly 50% of data set. To verify this hypothetical prediction error (as well as the distribution width), the CMV prediction errors from 5000 different permutations were collected. Determination of the CMV prediction error in each permutation is performed in exactly the same way as the original classification model. Thus, each permuted data set is subjected to 20 CMV's from which a mean CMV prediction error is estimated. The distribution of CMV prediction errors obtained from the 5000 permutations is considered as the H_0 distribution of no-effect.

To test whether the observed treatment effect is indeed statistically different from the H_0 distribution of no-effect, use is made of the p -value. The standard probability level of significance ($\alpha = 0.05$) is used to rationalize a statistically significant effect. If the original classification is not significantly better than the permutations, the treatment is not considered significant.

Comparison with the Ordinary PLS-DA Model. To establish whether variation splitting in crossover designed studies indeed improves the predictive ability¹⁹ of the classification model, the CMV prediction error of the multilevel PLS-DA model is compared with the ordinary PLS-DA model. To obtain the prediction error of the ordinary PLS-DA model, a similar data validation procedure is pursued as compared to the multilevel PLS-DA model (including CMV with variable selection, permutation testing and UV-scaling). The statistical significance of the observed treatment effect is again determined by means of the p -value.

Results and Discussion

¹H NMR Spectra of Urine. In the urinary ¹H NMR spectra of the 29 human subjects, a wide range of low-molecular weight metabolites can be identified which have previously been described in literature.^{25,41,42} Some representative NMR spectra from four different subjects after placebo and grape/wine extract intervention are illustrated in Figure 2. In the pairwise comparison of the NMR spectra, two types of variations can be distinguished, that is, the biological variation (between the subjects) and the variation induced by the nutritional treatment (within the subjects). As shown in Figure 2a, biological variations are markedly revealed in the spectral region between δ 1.0 ppm and δ 2.5 ppm, and include resonances of various endogenous metabolites. Among them, alanine (δ 1.48 ppm, d, CH₃), β -hydroxybutyric acid (δ 1.25 ppm, d, CH₃), lactic acid (δ 1.40 ppm, d, CH₃), acetone (δ 2.22 ppm, s, CH₃) and the *N*-acetyl groups (δ 2.02–2.06 ppm, CH₃) can be recognized. Here, the metabolite profiles differ considerably between the subjects, whereas the differences within the subjects are relative small.

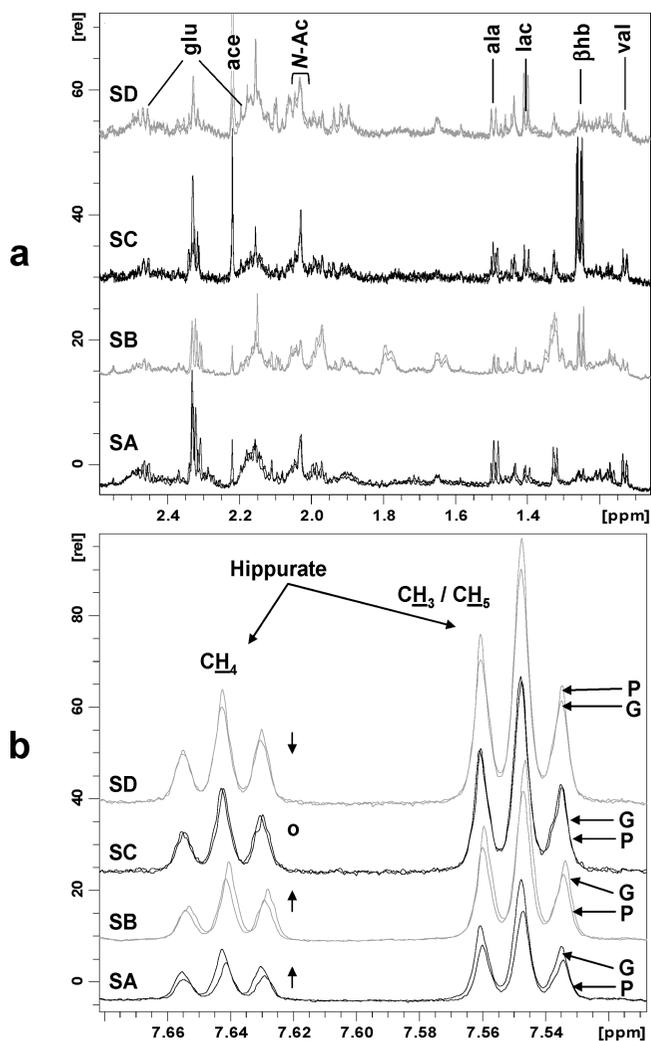


Figure 2. Creatinine normalized urinary 600 MHz ¹H NMR spectra (pH 3.0 ± 0.2), obtained from four subjects (SA, SB, SC and SD) after placebo intervention (P) and grape extract intervention (G). In panel a, the spectral region δ 1.0–2.5 ppm, the biological variations between the data-pairs are illustrated. Among them, the resonances of valine (val), alanine (ala), lactic acid (lac), acetone (ace), *N*-acetyl groups (*N*-Ac), β -hydroxybutyric acid (β hb) and glutamine/glutamate (glu) can be assigned. In panel b, the region between δ 7.5 ppm and δ 7.7 ppm, a part of the intense aromatic resonance pattern of hippuric acid is recognized. The hippuric acid signals increases (↑), decreases (↓) or does not change (o) within the subject pairs after grape/wine extract intervention.

In Figure 2b, the chemical shift region between δ 7.5 ppm and δ 7.7 ppm is highlighted and mainly includes two intense triplets of hippuric acid (δ 7.55 ppm, t, CH₃/CH₅; δ 7.64 ppm, t, CH₄). Although hippuric acid is present in all urinary profiles, distinctive variations were observed within the data-pairs. After intervention of the grape extract, increased signal intensities were observed in the urine samples of subject SA and subject SB. In subject SC, the hippuric acid levels were not affected, whereas subject SD exhibits an opposite treatment effect.

Analysis of Variation. According to eq 5, the variation in the ¹H NMR data set (X) can be partitioned into an offset term (X_m), a between-subject variation term (X_b) and a within subject variation term (X_w). Analysis of the sums of squares of these different subsets according to eq 6 shows that X is primarily described by the offset term. As determined in the sum of

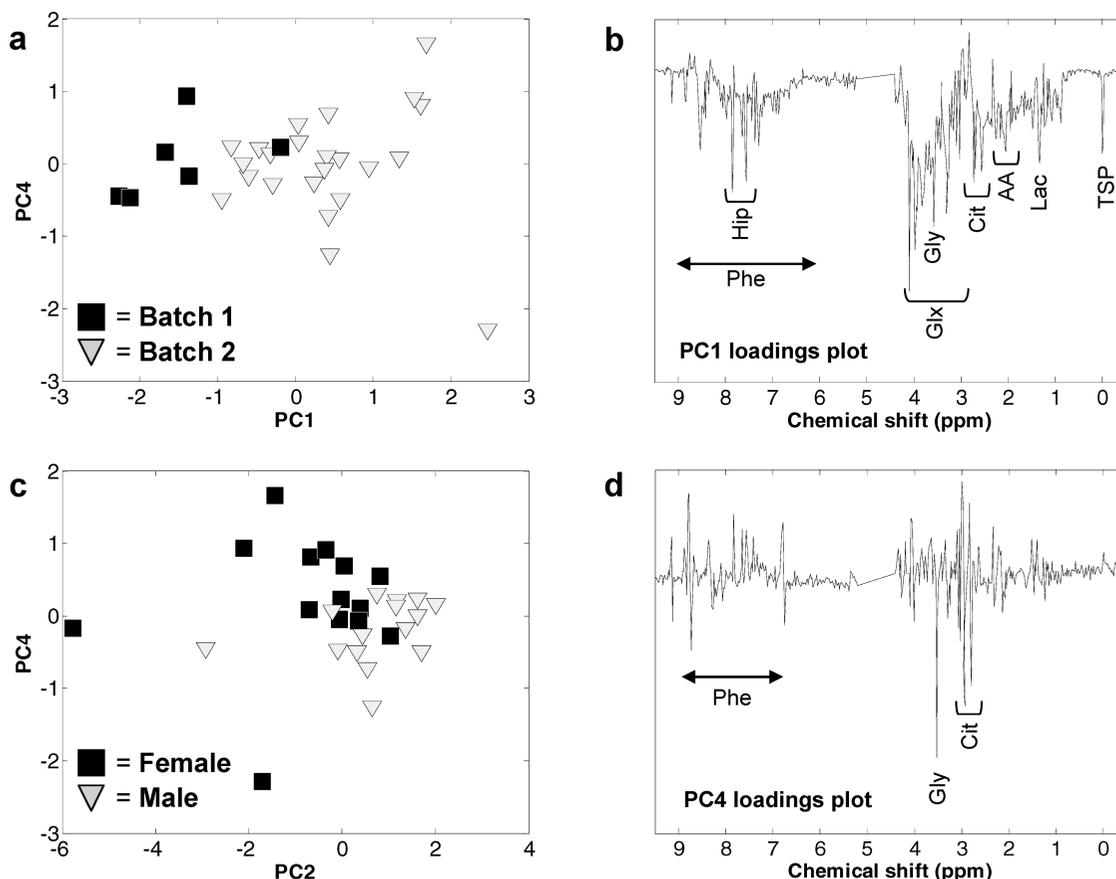


Figure 3. PCA between-individual scores of the first and the fourth PC. The differences in the score plots can mainly be explained by (a) experimental variations and (c) biological variations (gender). The experimental variations are mainly described by the first PC and (b) show in large extent the average urinary ^1H NMR profile. The most abundant metabolites can be assigned to phenolic compounds (Phe), glucose (Glx), glycine (Gly), citric acid (Cit), amino acids (AA), lactic acid (Lac) and hippuric acid (Hip). The gender-based differences are largely described by the (d) fourth PC, in which the most abundant signals can be assigned to glycine and citric acid.

squares ($\|\mathbf{X}_m\|^2$), 75.1% of \mathbf{X} is explained by \mathbf{X}_m . From the remaining 24.9%, 19.5% can be attributed to the between-subject variation ($\|\mathbf{X}_b\|^2$) and only 5.4% to the within-subject variation ($\|\mathbf{X}_w\|^2$). Thus, of the relevant variation between the subjects only one-fifth is due to the treatment and four-fifths is due to biological variation. As in the current data set, the variation in \mathbf{X}_b is much larger than in \mathbf{X}_w , the use of variation splitting is therefore all the more justified. In an ordinary PLS-DA approach, these confounding variation sources typically complicates the interpretation of the multivariate solution.^{9,10,16}

Analysis of the between-Subject Variation. To examine whether the between-subject variation contains discernible underlying data structures, a four component PCA model was fitted to \mathbf{X}_b . The spectral variables in \mathbf{X}_b were Pareto-scaled prior to the PCA decomposition. This approach of scaling after variation splitting is considered as an important benefit of the multilevel approach, as the preferred scaling technique can explicitly be adapted to the part of the data that is examined and the data analysis technique used.

The resulting PCA model explains 71.9% of the total between-subject variation. The scores on the first and the fourth components are shown in Figure 3a and explain, respectively, 37.1% and 7.5% of the between-subject variation. In the score plot, two subclasses can be distinguished along the first PC axis, consisting of 6 and 23 subject-scores. Inspection of these subclasses reveals that the NMR profiles originate from two different NMR runs. The loadings of the first PC in Figure 3b

describe variations in a large number of signals which are generally abundant in a urinary NMR spectrum. This suggests that spectral differences between the NMR runs may be explained by subtle variations in spectral line shapes, positional shifts and data processing parameters (e.g. phase- and baseline corrections), which are common sources of analytical noise in many other analytical and spectroscopic applications.^{43,44}

Besides analytical variation, also a gender-based effect could be recognized in the between-subject score distribution. As observed in Figure 3c, the scores that originate from the male and female subjects tend to cluster in two groups. The second PC and the fourth PC contribute to this gender-based difference and explain, respectively, 19.0% and 7.5% of the between-subject variation in the data. In Figure 3d, the fourth PC is mainly described by high loadings of citric acid (δ 2.81 ppm, d, $1/2\text{CH}_2$; δ 2.95 ppm, d, $1/2\text{CH}_2$) and glycine (δ 3.55 ppm, s, CH_2). These low-molecular weight metabolites have already been identified in previously described studies as related to gender differences in humans.^{45,46}

Analysis of the within-Subject Model. The within-subject data block (\mathbf{X}_w) is particularly examined on systematic variations in the NMR metabolic profiles as a result of the grape/wine extract consumption. A multilevel PLS-DA classification model is used to discriminate between the treatment group and the placebo group, in order to signify the intended treatment effect. After variation splitting, the spectral variables

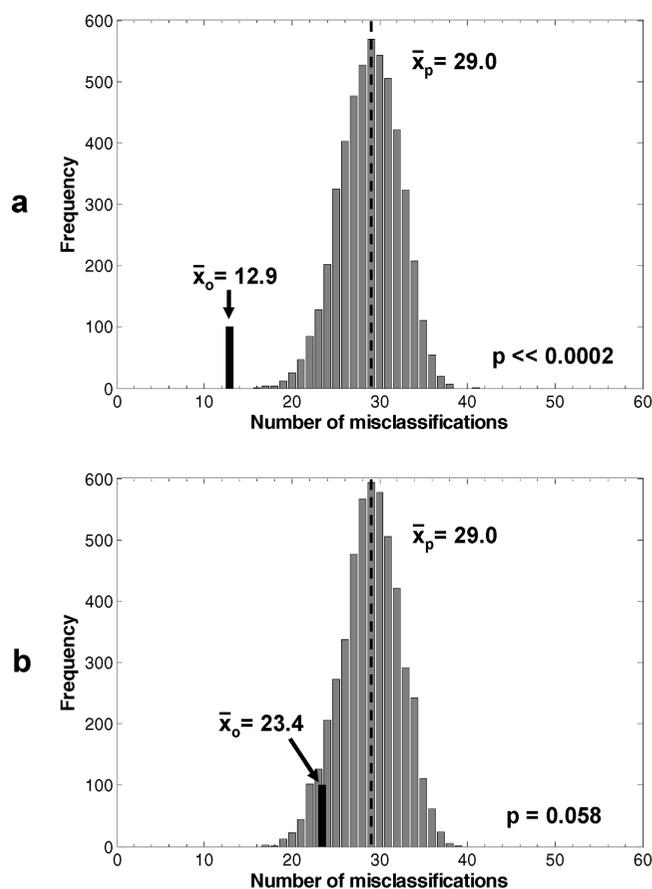


Figure 4. Mean cross model validation (CMV) prediction error (\bar{x}_o), estimated in terms of number of misclassifications and based on 20 CMV rounds. The results obtained from (a) the multilevel PLS-DA model ($\bar{x}_o = 12.9$) and (b) the ordinary PLS-DA model ($\bar{x}_o = 23.4$) are compared with the H_0 -distribution of no effect (\bar{x}_p). The H_0 -distribution of no-effect is based on the CMV prediction error of 5000 permutations.

in X_w were UV-scaled to make variations in the data set independent of the signal strength.

By means of CMV the intervention class labels of the test samples were predicted. To obtain stable class prediction results and a stable RP, the average result of 20 CMVs was calculated. As shown in Figure 4a, on average, 12.9 test samples were predicted wrongly, which is about 22% of all prediction results.

To validate whether the original classification model is significantly better than the classification models from randomly permuted data and not obtained by chance, a comparative permutation test was performed. In this test, the class predictions of permuted data were collected in 5000 permutations and represented as a H_0 -distribution of no effect. In case that no difference exists between the intervention classes, an average number of 29.0 misclassifications would be expected. As shown in Figure 4a, the experimentally obtained H_0 -distribution exactly matches this requirement. Since the permutations suggest that the analysis method is not overfitting the data, consequently the obtained CMV prediction error of the original model (= 12.9) is representative for the prediction of new subjects. Comparing the CMV prediction error of the original model against the permutations under the H_0 -distribution, result in a p -value which is much smaller than 0.0002. It appears that none of the 5000 permutations resulted in a

prediction error lower than 12.9 misclassifications. On the basis of the obtained p -value, the observed treatment effect was considered as statistically significant ($\alpha = 0.05$).

Comparison with the Ordinary PLS-DA Model. To assess the power of the multilevel PLS-DA model, the CMV prediction error was compared with the prediction error obtained from the ordinary PLS-DA model. Without variation splitting, this ordinary PLS-DA classification model is derived from the original data X . With the use of an identical CMV scheme and scaling technique as has been performed in the validation of the multilevel PLS-DA model, on average, 23.4 test samples were misclassified. As shown in Figure 4b, the permuted data again leads to an average number of 29.0 misclassifications. On the basis of the p -value of 0.058, this classification result is however not significant on the 5% significance level. The permutation test showed that 291 out of 5000 permuted models predict the class labels similarly or better than the original model.

The CMV prediction error indicates that the observed treatment effect in the ordinary PLS-DA is not significant. This is different for the multilevel PLS-DA model which indeed shows a significant treatment effect. This key observation supports the idea that within-subject variations in crossover designed metabolomic data should be analyzed by means of a multilevel based method.

Identification and Validation of Biomarkers. To investigate which ^1H resonances are the most important variables in the multilevel PLS-DA model, use was made of the rank product (RP).^{20,39} In Figure 5a, the rank product for each NMR variable ($\text{RP}^{1/20}$) is shown and compared with a representative urinary NMR spectrum for peak identification (Figure 5b,c). In each of the 20 CMVs, only 45 out of 550 variables were consistently selected in the model optimization (ICV). Among these variables, hippuric acid is the strongest NMR biomarker for the intake of the grape/wine extract. As illustrated in Figure 5a, hippuric acid is represented by 10 NMR spectral variables from aliphatic and aromatic protons (δ 7.83 ppm, d, CH₂/CH₆; δ 7.64 ppm, t, CH₄; δ 7.55 ppm, t, CH₃/CH₅; δ 4.13 ppm, d, CH₂), all of which are present within the lowest 10% quantile of the RP. The presence of hippuric acid is in agreement with previously reported studies where the metabolic impact of polyphenolic-rich food consumption (tea) in humans was studied.^{47,48} Hippuric acid, or *N*-benzoylglycine, is basically the glycine conjugated form of benzoic acid, which is thought to be the metabolic end-product of flavonoid degradation by the gut microbiota. However, besides hippuric acid, also other phenolic compounds were important in the classification between the intervention groups, especially those who have resonance frequencies between δ 6.8 ppm and δ 7.2 ppm. On the basis of the chemical shift positions, the multiplicity of the signals, the peak intensities of the resonances and the relative abundance of the metabolites, at least two other phenolic compounds could be assigned, that is, 4-hydroxyhippuric acid and 4-hydroxyphenylacetic acid (Table 1). Like hippuric acid, these phenolic acids are gut microbial fermentation products of flavanoids.^{49–53}

To our knowledge, this is the first time that NMR is able to detect changes in these phenolic compounds in urine after grape/wine extract consumption. This is remarkable since their observed urinary levels (50–300 μM) are close to the lower detection limits of NMR-based metabolic profiling. An important aspect in the CMV procedure that allows the determination of such minor metabolites in the RP is the use of UV-scaling.

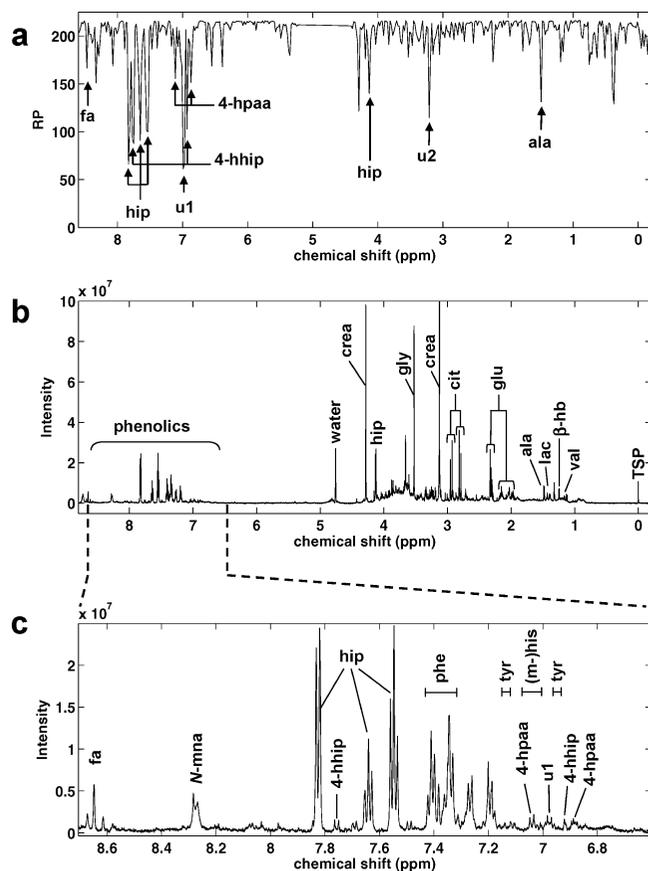


Figure 5. Comparison of (a) the variable ranks ($RP^{1/20}$) with (b) a representative urinary 600 MHz 1D 1H NMR spectrum, whereby (c) the aromatic region between δ 6–9 ppm is expanded. The rank product is the result of 20 cross model validations. The variables with the lowest ranks can be assigned to alanine (ala), hippuric acid (hip), 4-hydroxyphenylacetic acid (4-hppa), 4-hydroxyhippuric acid (4-hhip), formic acid (fa), and two unknown compounds (u1 and u2). Other urinary metabolites in the NMR spectrum could be identified as creatinine (crea), glycine (gly), citric acid (cit), glutamine (glu), lactic acid (lac), β -hydroxybutyric acid (β -hb), valine (val) and *N*-methylnicotine amide (*N*-mna).

Table 1. Assignment of 1H NMR Signals in the Rank Product

compound	chemical shift (multiplicity) and assignment	relative abundance ^a
Hippuric acid	4.127 ppm (d) CH_2 ; 7.547 ppm (d) CH; 7.641 ppm (t) CH; 7.826 ppm (d) CH	↑
4-Hydroxyhippuric acid	4.285 ppm (s) CH_2 (ov); ^b 6.918 ppm (d) CH; 7.756 ppm (d)	↑
4-Hydroxyphenylacetic acid	6.879 ppm (d) CH; 7.112 ppm (d) CH	↑
u1	6.980 ppm (d)	↑
Formic acid	8.664 ppm (s) CH	↑
Alanine	1.480 ppm (d) CH_3	↓
u2	3.210 ppm (s) CH_3	↓
Total phenolics	Various signals: 6.5–8.6 ppm aromatic ring CH; 2.0–7.0 ppm heterocyclic ring CH/ CH_2	n/a

^a Relatively higher (↑) or lower (↓) amount of the metabolites found in urine after grape/wine extract consumption. ^b Ov = overlapping signals.

Although this scaling technique is prone to overestimating noise, a suitable signal-to-noise ratio can be obtained by using

the average result of 20 CMV rounds. The presence and identity of these compounds was confirmed by additional GC-MS experiments, by spectral comparison with commercially available reference standards, and by spiking experiments. Another discriminating signal with a low rank product was observed at δ 6.98 ppm (d). However, further identification of this unknown aromatic signal (u1) was hampered by its low signal intensity and the absence of other resonance patterns.

Also some endogenous metabolites contributed to the differentiation between treatment and placebo groups, including alanine⁵⁴ (δ 1.48 ppm), formic acid (δ 8.66 ppm) and an unknown metabolite (u2) at chemical shift position δ 3.21 ppm. According to the RP analysis they were less discriminative than the phenolic acid metabolites.

Conclusion

Taking advantage of the multilevel structure in a crossover designed metabolomics study has shown major benefits in comparison with the traditionally PLS-DA approach. Not only the interpretability of the different variation sources in the data was improved, but also the predictive strength of the classification model has significantly increased. The between-subject model showed that the major effects in the data can be associated with analytical variations and gender differences. The within-subject model contains the nutritional effect which is only a minor fraction of the total variation. The multilevel PLS-DA model showed that the observed treatment effect was statistically significant as the p-value is much smaller than 0.0002. This is an enormous improvement compared to the original PLS-DA model which did not show a significant treatment effect ($p = 0.058$). The combination of cross model validation, permutation testing and rank products provided several candidate biomarkers that can be associated with the consumption of the grape/wine extract. Among these metabolites, hippuric acid is the most important and well-identified biomarker. This observation fits well in current understanding of bioconversion of polyphenols in humans by the gut microflora. In the rank product, also fermentation products could be assigned that occur at levels close to the lower detection range of NMR-based metabolic profiling, that is, 4-hydroxyhippuric acid, 4-hydroxyphenylacetic acid and formic acid.

Acknowledgment. We are grateful to the European Commission for their financial support of the GutSystem project (MTKI-CT-2006-042786) under the Framework 6 Marie-Curie Transfer of Knowledge Industry-Academia Strategic Partnership scheme. Afranina Leika is acknowledged for her assistance in the sample preparation and sample measurements and Daniël Vis and Maikel Verouden for their helpful contributions in the development and optimization of the data analysis methods.

Supporting Information Available: This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Gibney, M. J.; Walsh, M.; Brennan, L.; Roche, H. M.; German, B.; van Ommen, B. *Am. J. Clin. Nutr.* **2005**, *82*, 497–503.
- (2) Solanky, K. S.; Bailey, N. J.; Beckwith-Hall, B. M.; Bingham, S.; Davis, A.; Holmes, E.; Nicholson, J. K.; Cassidy, A. *J. Nutr. Biochem.* **2005**, *16*, 236–44.
- (3) Rezzi, S.; Ramadan, Z.; Fay, L. B.; Kochhar, S. *J. Proteome Res.* **2007**, *6*, 513–25.
- (4) van der Greef, J.; Smilde, A. K. *J. Chemom.* **2005**, *19*, 376–86.

- (5) van der, G. J.; Martin, S.; Juhasz, P.; Adourian, A.; Plasterer, T.; Verheij, E. R.; McBurney, R. N. *J. Proteome Res.* **2007**, *6*, 1540–59.
- (6) Nicholson, J. K.; Holmes, E.; Lindon, J. C.; Wilson, I. D. *Nat. Biotechnol.* **2004**, *22*, 1268–74.
- (7) Stella, C.; Beckwith-Hall, B.; Cloarec, O.; Holmes, E.; Lindon, J. C.; Powell, J.; van der Ouderaa, F.; Bingham, S.; Cross, A. J.; Nicholson, J. K. *J. Proteome Res.* **2006**, *5*, 2780–88.
- (8) Maher, A. D.; Zirah, S. F. M.; Holmes, E.; Nicholson, J. K. *Anal. Chem.* **2007**, *79*, 5204–11.
- (9) Smilde, A. K.; Jansen, J. J.; Hoefsloot, H. C.; Lamers, R. J.; van der, G. J.; Timmerman, M. E. *Bioinformatics* **2005**, *21*, 3043–48.
- (10) Jansen, J. J.; Hoefsloot, H. C. J.; van der Greef, J.; Timmerman, M. E.; Smilde, A. K. *Anal. Chim. Acta* **2005**, *530*, 173–83.
- (11) Putt, M.; Chinchilli, V. M. *Stat. Med.* **1999**, *18*, 3037–58.
- (12) Jacobs, D. M.; Deltimple, N.; Van Velzen, E. J. J.; Van Dorsten, F. A.; Bingham, M.; Vaughan, E. E.; van Duynhoven, J. P. M. *NMR Biomed.* **2008**, *21*, 615–26.
- (13) Elbourne, D. R.; Altman, D. G.; Higgins, J. P.; Curtin, F.; Worthington, H. V.; Vail, A. *Int. J. Epidemiol.* **2002**, *31*, 140–49.
- (14) Saude, E. J.; Adamko, D.; Rowe, B. H.; Marrie, T.; Sykes, B. D. *Metabolomics* **2006**, *2*, 439–51.
- (15) Lenz, E. M.; Bright, J.; Wilson, I. D.; Hughes, A.; Morrisson, J.; Lindberg, H.; Lockton, A. *J. Pharm. Biomed. Anal.* **2004**, *36*, 841–49.
- (16) Jansen, J. J.; Hoefsloot, H. C.; van der Greef, J.; Timmerman, M. E.; Westerhuis, J. A.; Smilde, A. K. *J. Chemom.* **2005**, *19*, 469–81.
- (17) Vis, D. J.; Westerhuis, J. A.; Smilde, A. K.; van der, G. J. *BMC Bioinf.* **2007**, *8*, 322–34.
- (18) Timmerman, M. E. *Br. J. Math. Stat. Psychol.* **2006**, *59*, 301–20.
- (19) Anderssen, E.; Dyrstad, K.; Westad, F.; Martens, H. *Chemom. Intell. Lab. Syst.* **2006**, *84*, 69–74.
- (20) Smit, S.; van Breemen, M. J.; Hoefsloot, H. C.; Smilde, A. K.; Aerts, J. M.; de Koster, C. G. *Anal. Chim. Acta* **2007**, *592*, 210–17.
- (21) Stratil, P.; Klejdus, B.; Kuban, V. *J. Agric. Food Chem.* **2006**, *54*, 607–16.
- (22) Manach, C.; Scalbert, A.; Morand, C.; Remesy, C.; Jimenez, L. *Am. J. Clin. Nutr.* **2004**, *79*, 727–47.
- (23) Craig, A.; Cloarec, O.; Holmes, E.; Nicholson, J. K.; Lindon, J. C. *Anal. Chem.* **2006**, *78*, 2262–67.
- (24) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. *Anal. Chem.* **2006**, *78*, 4281–90.
- (25) Constantinou, M. A.; Papakonstantinou, E.; Spraul, M.; Sevastiadou, S.; Costalos, C.; Koupparis, M. A.; Shulpis, K.; Tsantili-Kakoulidou, A.; Mikros, E. *Anal. Chim. Acta* **2005**, *542*, 169–77.
- (26) Halouska, S.; Powers, R. *J. Magn. Reson.* **2006**, *178*, 88–95.
- (27) Holmes, E.; Tang, H.; Wang, Y.; Seger, C. *Planta Med.* **2006**, *72*, 771–85.
- (28) van den Berg, R. A.; Hoefsloot, H. C.; Westerhuis, J. A.; Smilde, A. K.; van der Werf, M. J. *BMC Genomics* **2006**, *7*, 142–57.
- (29) Keun, H. C.; Ebbels, T. M. D.; Antti, H.; Bollard, M. E.; Beckonert, O.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Anal. Chim. Acta* **2003**, *490*, 265–76.
- (30) Alam, T. M.; Alam, M. K. *Annu. Rep. NMR Spectrosc.* **2005**, *54*, 41–80.
- (31) Whelehan, O. P.; Earll, M. E.; Johansson, E.; Toft, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2006**, *84*, 82–87.
- (32) Jolliffe, I. T. *Principal Component Analysis*, 2nd ed.; Springer-Verlag: New York, 2002.
- (33) Wold, S.; Sjostrom, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–30.
- (34) Faber, N. M.; Rajko, R. *Anal. Chim. Acta* **2007**, *595*, 98–106.
- (35) Broadhurst, D. I.; Kell, D. B. *Metabolomics* **2006**, *2*, 171–96.
- (36) Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; Van Velzen, E. J. J.; Van Duynhoven, J. P. M.; Van Dorsten, F. A. *Metabolomics* **2008**, *4*, 81–9.
- (37) Stone, M. *J. R. Stat. Soc. B* **1974**, *36*, 111–33.
- (38) Gidskehaug, L.; Anderssen, E.; Alsberg, B. K. *Chemom. Intell. Lab. Syst.* **2006**, *84*, 172–76.
- (39) Breitling, R.; Armengaud, P.; Amtmann, A.; Herzyk, P. *FEBS Lett.* **2004**, *573*, 83–92.
- (40) Bijlsma, S.; Bobeldijk, L.; Verheij, E. R.; Ramaker, R.; Kochhar, S.; Macdonald, I. A.; van Ommen, B.; Smilde, A. K. *Anal. Chem.* **2006**, *78*, 567–74.
- (41) Martin, F. P.; Dumas, M. E.; Wang, Y.; Legido-Quigley, C.; Yap, I. K.; Tang, H.; Zirah, S.; Murphy, G. M.; Cloarec, O.; Lindon, J. C.; Sprenger, N.; Fay, L. B.; Kochhar, S.; van Bladeren, P.; Holmes, E.; Nicholson, J. K. *Mol. Syst. Biol.* **2007**, *3*, 112–28.
- (42) Nicholson, J. K.; Wilson, I. D. *Prog. Nucl. Magn. Reson. Spec.* **1989**, *21*, 449–501.
- (43) Vogt, F.; Booksh, K. *Appl. Spectrosc.* **2004**, *58*, 624–35.
- (44) Teahan, O.; Gamble, S.; Holmes, E.; Waxman, J.; Nicholson, J. K.; Bevan, C.; Keun, H. C. *Anal. Chem.* **2006**, *78*, 4307–18.
- (45) Kochhar, S.; Jacobs, D. M.; Ramadan, Z.; Berruex, F.; Fuerholz, A.; Fay, L. B. *Anal. Biochem.* **2006**, *352*.
- (46) Psihogios, N. G.; Gazi, I. F.; Elisaf, M. S.; Seferiadis, K. I.; Bairaktari, E. T. *NMR Biomed.* **2008**, *21*, 195–207.
- (47) Daykin, C. A.; Van Duynhoven, J. P. M.; Groenewegen, A.; Dachtler, M.; Van Amelsvoort, J. M. M.; Mulder, T. P. J. *J. Agric. Food Chem.* **2005**, *53*, 1428–34.
- (48) Van Dorsten, F. A.; Daykin, C. A.; Mulder, T. P. J.; Van Duynhoven, J. P. M. *J. Agric. Food Chem.* **2006**, *54*, 6929–38.
- (49) Gonthier, M. P.; Cheyner, V.; Donovan, J. L.; Manach, C.; Morand, C.; Mila, I.; Lapiere, C.; Remesy, C.; Scalbert, A. *J. Nutr.* **2003**, *133*, 461–67.
- (50) Hollman, P. C. H. *Methods Enzymol.* **2001**, *335*, 97–103.
- (51) Scalbert, A.; Manach, C.; Morand, C.; Remesy, C.; Jimenez, L. *Crit. Rev. Food Sci. Nutr.* **2005**, *45*, 287–306.
- (52) Rechner, A. R.; Kuhnle, G.; Bremner, P.; Hubbard, G. P.; Moore, K. P.; Rice-Evans, C. A. *Free Radical Biol. Med.* **2002**, *33*, 220–35.
- (53) Ito, H.; Gonthier, M. P.; Manach, C.; Morand, C.; Mennen, L.; Remesy, C.; Scalbert, A. *Br. J. Nutr.* **2005**, *94*, 500–9.
- (54) Holmes, E.; Loo, R. L.; Stamler, J.; Bictash, M.; Yap, I. K.; Chan, Q.; Ebbels, T.; De Iorio, M.; Brown, I. J.; Veselkov, K. A.; Daviglus, M. L.; Kesteloot, H.; Ueshima, H.; Zhao, L.; Nicholson, J. K.; Elliott, P. *Nature* **2008**, *453*, 396–400.

PR800145J