Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Pharmaceutical and Biomedical Analysis
33 (2003) 1103–1115
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

A ^{1}H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects

E.M. Lenz [a,*], J. Bright [b], I.D. Wilson [a], S.R. Morgan [c], A.F.P. Nash [c]

[a] *Department of Drug Metabolism and Pharmacokinetics (DMPK), AstraZeneca Pharmaceuticals, Mereside, Alderley Park, Macclesfield SK10 4TG, UK*
[b] *Department of Enabling Science and Technology, AstraZeneca Pharmaceuticals, Mereside, Alderley Park, Macclesfield SK10 4TG, UK*
[c] *Department of Experimental Medicine, AstraZeneca Pharmaceuticals, Mereside, Alderley Park, Macclesfield SK10 4TG, UK*

Received 27 May 2003; received in revised form 30 June 2003; accepted 5 July 2003

Abstract

The aim of the study was to assess the feasibility of metabonomics in clinical studies. A ^{1}H nuclear magnetic resonance (NMR)-based metabonomic analysis was performed on plasma and urine samples obtained from a group of 12 healthy male subjects on two separate study days 14 days apart. The subjects were fed a standard diet and plasma and urine samples were obtained on both days. The ^{1}H NMR spectra obtained for urine and plasma samples were analysed using principal components analysis (PCA) in order to generate metabonomic data. In plasma there was relatively little variability between subjects and study days. In the case of endogenous urinary metabolite profiles there was considerable inter-subject variability, but less intra-subject variation. In all subjects diurnal variation was seen with urine samples. This suggests the possibility to collect consistent metabonomics data in clinical studies.

Keywords: Metabonomics; ^{1}H NMR spectroscopy; Diurnal variation; Metabolic profiling; Human subjects

1. Introduction

The potential benefits of the so-called ''post-genomic'' technologies for increasing our understanding of human disease processes are widely recognised. One of these new technologies, in addition to ''genomics'' and ''proteomics'' is the rapidly emerging field of ''metabonomics'' [1]. Metabonomics studies complement genomic and proteomic investigations by providing a quantitative description of the low-molecular mass endogenous metabolites present in a biological sample such as urine, plasma or tissue. In such studies analysis is usually performed using high field ^{1}H nuclear magnetic resonance (NMR) spectroscopy, which gives a ''metabolite fingerprint''. These spectra thus provide characteristic patterns that, with suitable chemometric analysis using, e.g. principal components analysis (PCA), can be used to provide information on such diverse areas as gender, strain of animal, diurnal variation, response to toxic insults [2–4] and disease [5].

* Corresponding author. Tel.: +44-1625-51-4653; fax: +44-1625-51-6962.
E-mail address: eva.lenz@astrazeneca.com (E.M. Lenz).

0731-7085/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0731-7085(03)00410-2

Potentially, metabonomics also has a role in monitoring the response to drug therapy. Metabonomics has, to date, mostly been applied to studies in experimental animals [6–9]. However, it is inevitable that this powerful technique will increasingly be used in studies of both healthy and diseased humans and recently Brindle et al. [5] have demonstrated how metabonomics can be used to diagnose the presence and severity of coronary heart disease. One problem faced in clinical investigations is the inherently greater variability in a human population compared with that seen in experimental animals. Thus in animal studies factors such as age, gender, diet and other environmental influences, are much more under the control of the investigator than when human populations are investigated. Clearly, for reliable application of metabonomics (and arguably for proteomics as well), in humans it is necessary to demonstrate that it would be possible to detect the effects of disease progression or drug therapy in subjects under clinical study conditions. Here we describe a study on healthy subjects designed to evaluate the variability in metabonomic data and its clinical utility for incorporation into early-phase drug trials.

2. Experimental

2.1. Subjects

This study was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The local Ethics Committee approved the protocol before commencement of the study, and all subjects gave written informed consent. 12 healthy, male, non-vegetarian subjects, over 21 years of age with a body mass index of between 18 and 30, were recruited for this study from the AstraZeneca Subject Panel, following a qualifying medical questionnaire. Exclusion criteria used in the selection of the subjects included current use of any regular medication or therapy, participation in another study within 3 months before the start of the present study, acute illness within the 2 weeks preceding the start of the study. In addition subjects were excluded from the study if clinically significant abnormalities in clinical chemistry or haematology were present or there was a risk (in the investigator's opinion) of transmitting, through blood or other body fluids, the agents responsible for acquired immune deficiency syndrome, hepatitis B or C. Excessive intake of alcohol, defined as a regular maximum weekly intake of greater than 28 units, treatment in the previous 3 months with any drug known to have a well-defined potential for hepatotoxicity (e.g. halothane), a history of asthma, hayfever or allergies to food or animals, osteoarthritis, rheumatoid arthritis, renal or hepatic disease, or vaccinations during the month prior to study start were also used to exclude subjects from the study.

A pre-study blood sample was taken in order to perform standard clinical chemistry and haematology assessments, drugs of abuse screen, and for hepatitis B surface antigen and hepatitis C antibody determination. Urine was tested pre-study for the drugs of abuse methadone, benzodiazipines, cocaine, amphetamines, tetrahydro-cannabinol (THC), opiates, methamphetamines (ecstasy) and barbiturates. A positive result for drugs of abuse was used as a reason for exclusion.

Subjects were required to abstain from taking any medication (including over-the-counter remedies) from 24 h before sampling until the final sample had been taken on each study day, unless the investigator had given prior consent. Subjects were fasted from midnight on the night before each study day (permitted fluids only were allowed from 06:00 h). A standard breakfast, lunch and dinner were provided on each study day. Subjects were required to abstain from smoking, consuming grapefruit, liquorice or caffeine-containing drinks or foods (e.g. coffee, tea, cocoa, chocolate and cola) from midnight before each study day until the final sample had been taken on that day and abstain from eating any cheese, fish or cherries from 24 h before each study day until the completion of the 24 h urine collection. In addition, subjects were required to abstain from eating any food in between meals or between dinner and the time of the last blood sample and not to drink alcohol from 24 h before each study day until the final sample had been taken on that day. The

subjects were also asked to refrain from all strenuous/hazardous physical activities (e.g. running, swimming, etc.) from 72 h before each study day until the final sample had been taken on that day and not to donate blood during the study.

2.2. Food and fluid intake

Identical standard meals were provided for all subjects on each study day (breakfast at 07:30 h, lunch at 12:00 h, and dinner at 17:00 h; all approximate times). There was free access to permitted fluids from 06:00 h and throughout each study day and there was no restriction on permitted foods between 21:00 and 09:00 h for the 12–24 h collection period. Diary cards were used to record all dietary intake (food and fluids) from 21:00 to 09:00 h.

2.3. Samples

Venous blood (2.7 ml) was taken at approximately 09:00 h into a lithium heparin tube and centrifuged at $1000 \times g$ at 4 °C for 10 min to provide plasma for analysis.

The first void, and 0–12 h urine collections were frozen at −20 °C on collection and stored frozen until analysis. The 12–24 h urine collections were refrigerated until being stored frozen at −20 °C prior to analysis. Urinary pH values were measured and recorded.

2.4. ^1H NMR spectroscopy

Analysis of plasma and urine samples was carried out by ^1H NMR spectroscopy on a Bruker DRX500 spectrometer operating at 500.13 MHz ^1H resonance frequency. The NMR-probe used for that purpose was a SEI-microprobe and the samples were run in NMR-microtubes with a total volume of 250 μl.

For plasma samples, 200 μl aliquots of neat plasma were transferred into the NMR tubes and 50 μl of TSP (3-trimethylsilyl-$^2H_4$-propionic acid) in D_2O (1 mg/ml) were added. With urine samples an aliquot (3 ml) was freeze-dried prior to analysis and reconstituted into 500 μl of D_2O. A 200 μl aliquot of each urine concentrate was transferred into the NMR tubes to which a further 50 μl of TSP in D_2O (1 mg/ml) were added. ^1H NMR spectra were acquired immediately after preparation of each individual sample. The D_2O provided a field frequency lock-solvent for the NMR spectrometer and the TSP served as an internal chemical shift reference (δ_{1H} 0.0).

All spectra were recorded at 30 °C. Typically, ^1H NMR spectra were measured with 64 scans into 65 536 data points over a spectral width of 9980.04 Hz, which resulted in an acquisition time of 3.28 s. A relaxation delay of 1.8 s additionally ensured T_1-relaxation between successive scans. Solvent suppression of the residual water signals (a broad singlet at δ_{1H} 4.8) was achieved via the Noesypresat pulse sequence (Bruker Spectrospin Ltd.) in which the residual water peak is irradiated during the relaxation delay and during the mixing time of 150 ms.

Spectra from plasma samples acquired with the Noesypresat pulse sequence typically result in spectra containing rolling baselines due to broad, unresolved, signals from proteins, hence, the plasma samples were also analysed using the Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence. The CPMG-pulse (Bruker Spectrospin Ltd.) sequence allows the 'spectral editing/filtering' of the broad protein signals and produces spectra with flat baselines. CPMG ^1H NMR spectra were acquired with 16 384 data points collected over a spectral width of 5995.20 Hz, resulting in an acquisition time of 1.37 s. A 3 s relaxation delay ensured T_1-relaxation between successive scans and a 96 ms pulse train ($\tau = 0.5$ ms) allowed spectral editing via faster T_2-relaxation, and hence, attenuation of broad signals. Here, the CPMG pulse sequence was additionally modified to allow for presaturation of the residual water peak ('cpmgpr', involving irradiation of the water frequency during the relaxation delay).

All ^1H NMR spectra were manually corrected for phase and baseline distortions within XWINNMR™ (version 2.6, Bruker Spectrospin Ltd.). Spectra were referenced to TSP (δ_{1H} 0.0) prior to data-reduction into 245 spectral integral regions corresponding to the chemical shift range of δ_{1H} 0.2–10 utilising AMIX (version 2.7.5, Analysis of MIXtures, Bruker Spectrospin Ltd.).

The region of δ_{1H} 4.52–6.0 was set to zero to remove the effects of variations in the presaturation of the water resonance in all NMR spectra, and to alleviate cross-relaxation effects in the urea signal via solvent exchanging protons. Integration into bins (or buckets) across the spectral regions of 0.04 ppm was performed automatically in AMIX. Normalisation, i.e. making the total integrated intensity of each spectrum the same, compensated for the differences in overall concentration between individual urine and plasma samples. The resulting data matrix (peak integral values/bins per sample) was analysed by pattern recognition methods within SIMCA-P (version 8, UMETRICS AB, Box 7960, SE 90719, Umeå, Sweden) and/or visualised using "SPOTFIRE" (Spotfire DecisionSite 6.2 version 6.2.0).

2.5. Statistical methods and software

Following the processing of the spectra by AMIX, data analysis was performed using various techniques including PCA. The SPOTFIRE program was used to visualise both the spectral data (reconstructed data-reduced spectra) and the output from SIMCA-P. Visualising the spectra in this way aided the identification of unusual spectra and individual peaks in the spectra, which increased or diminished over time. SIMCA-P was used to perform the PCA. PCA was performed using centred scaling. Detailed accounts of pattern recognition methods can be found in [10].

3. Results and discussion

3.1. Demography

A group of 12 healthy male subjects was recruited for the trial, consisting of 11 Caucasians and one Afro-Caribbean, with an average age of 37.8 ± 9.4 years (mean $\pm$ S.D.) and a body mass index of between 18 and 30.

3.2. Plasma

As indicated in Section 2.4, ^{1}H NMR spectra of plasma samples were obtained using two different pulse sequences, which provide different types of data. The Noesypresat spectra typically consist of broad peaks (rolling baseline) from high molecular mass species such as proteins, which partially obscure the sharper peaks from the low MW components. The CPMG technique on the other hand edits out the resonances from the proteins and, hence, reveals the smaller molecules in the sample, such as glucose, alanine, etc. [11]. Examples of typical Noesypresat and CPMG spectra obtained in this study are shown in Fig. 1a and b, obtained for subject 5 (on study day 2). The spectra, following data reduction in AMIX, were then reconstructed (representing integral values vs. buckets) and superimposed within SPOTFIRE. This showed that all of the CPMG spectra were comparable between subjects and study days. In the case of the Noesypresat spectra, SPOTFIRE indicated the presence of one subject showing an unusual spectrum on day 1 of the study, and this sample was also subsequently identified as an outlier using PCA. However, close inspection of the spectrum revealed that this was due to a spectral artefact (baseline-dip) rather than a metabolic difference (data not shown). Indeed there was relatively little variation in the spectra between subjects. The differences observed, identified from the appropriate scores plots, mainly resulted from variation in the concentrations of lipids, lactate and glucose. The PCA scores plot (generated in SIMCA-P and displayed using SPOTFIRE) for the plasma data is shown in Fig. 2 (CPMG spectra). The first two principal components (PC1 and PC2) contained 95.5% (Noesypresat, data not shown) and 97.4% (CPMG) of the variation in the original spectra.

As this plot shows, in many cases, the two samples obtained for each subject grouped closely to each other (e.g. see those for volunteer 5). In some cases, such as, that for volunteer 4 the grouping was less tight in both, Noesypresat and CPMG scores plots. Close examination of the ^{1}H NMR spectrum for this subject (see Fig. 3a and b) reveals that the difference between the spectra obtained for the day 1 and 14 samples were due to higher concentrations of lipid on study day 2. Generally, however, intra-subject variation seemed less than inter-subject variation. Subjects 9 and 12,

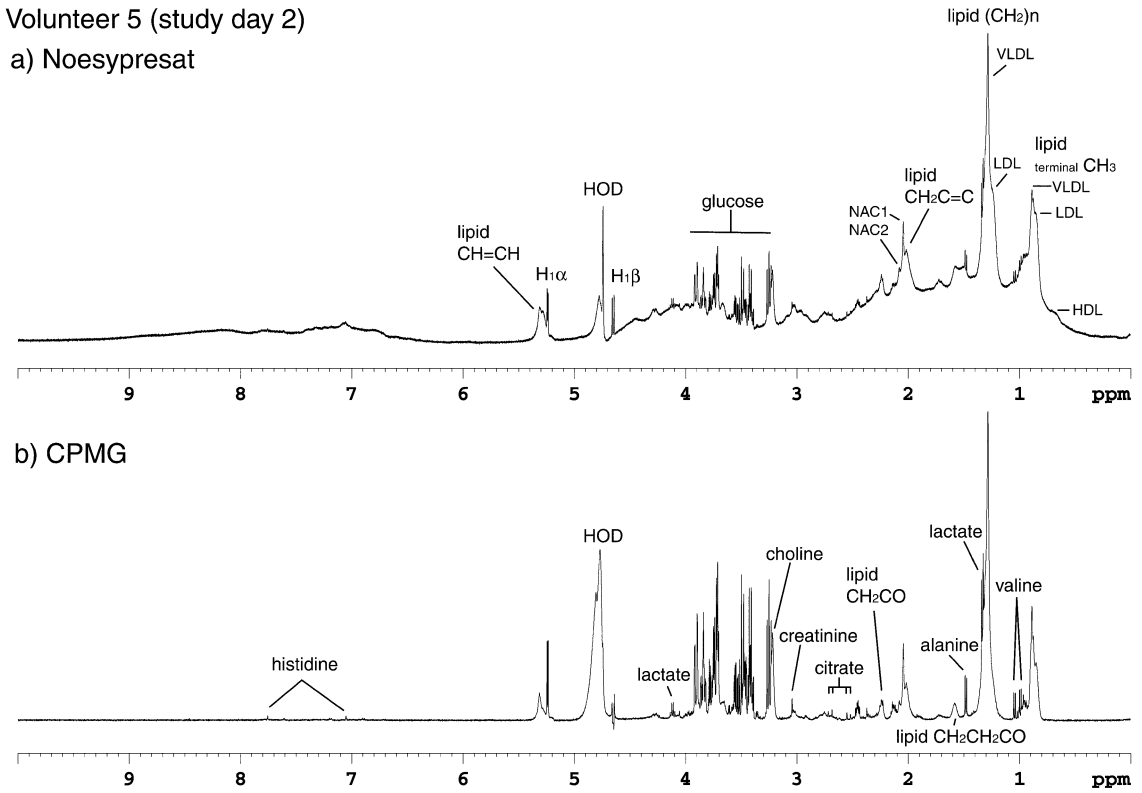


Fig. 1. An example of a Noesypresat (a) and CPMG (b) spectrum of plasma for subject 5 on study day 2. Key: H1α/β, anomeric proton of α/β-glucose; NAC1/2, *N*-acetyl glycoprotein signals; HDL, LDL, VLDL, high, low and very low density lipoprotein, respectively; HOD, water/D_2O.

for example, contained the highest concentrations of glucose and lipid, respectively, on both study days.

3.3. Urine

In comparison to plasma, urine samples present fewer technical problems with regard to obtaining good quality ^{1}H NMR spectra. However, human urines are relatively dilute and, with the probe configuration used here, can require relatively long spectral acquisition times. In this study we, therefore, found it expedient to obtain spectra of urine that had been freeze-dried, and re-dissolved in a smaller volume of D_2O in order to concentrate it. A typical ^{1}H NMR spectrum obtained in this study is shown in Fig. 4, with the major components identified.

As with the plasma samples, each urine spectrum was first data reduced within the AMIX software and then reconstructed within the SPOTFIRE data visualisation package. Investigations within SPOTFIRE highlighted three unusual spectra, clearly shown in Fig. 5a. The same urines were also identified as outliers using PCA (Fig. 5b). The three spectra in question were the first voids for subjects 2 and 3 on day 1, and for subject 11 on study day 2. Further investigation showed that the main differences between these three spectra and the remainder were almost certainly due to pH effects. The overall pH (mean $\pm$ S.D.) of the urines obtained from the subjects was 6.72 $\pm$ 1.05 (Table 1) whilst the pH-values of these urines were 5.28, 5.35 and 5.47, respectively. These, in turn, appeared to have caused a slight downfield shift of the creatinine peaks (from δ_{1H} 3.06 to 3.09) into

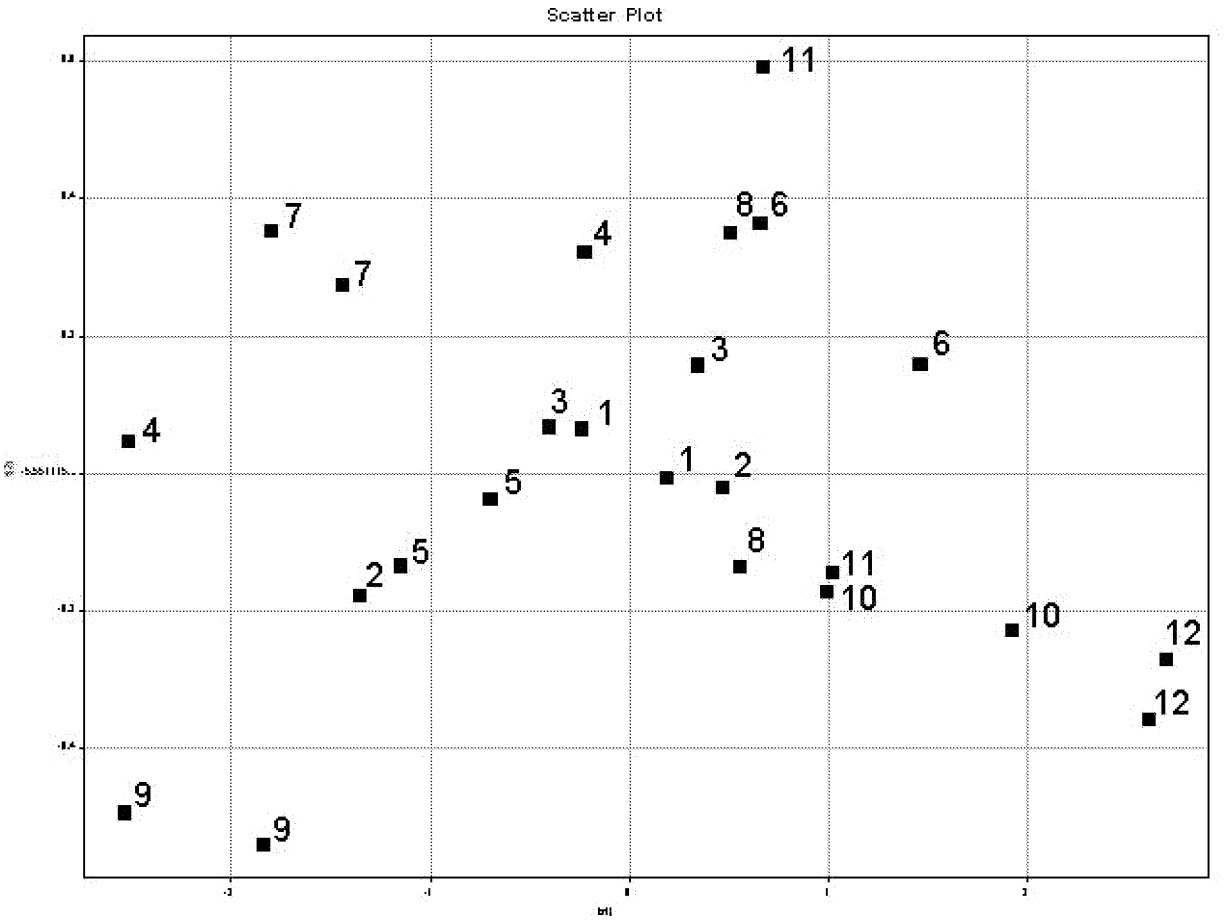


Fig. 2. PCA scores plot (PC1 vs. PC2) generated in SIMCA-P and displayed using SPOTFIRE for ^{1}H NMR CPMG spectra of the plasma samples collected on days 1 and 2.

the adjacent bucket. Hence, the integral regions were pooled (merged) to take account of the pH dependent chemical shifts of the creatinine signal (adjacent bins 3.10–3.06 and 4.14–4.06), which also affected the less dominant citrate resonances (adjacent bins 2.74–2.50), to produce "superbins" labelled 3.1+ and 4.14+ and 2.74+, respectively. The resulting PCA scores plot is shown in Fig. 6a, with the first two principal components containing 39.6% of the variation in the original spectra.

Generally, there was somewhat more variation in first voids for a given person across the two study days than at the subsequent two time-points. This was quite striking for subjects 5 and 11, both of whom produced reasonably consistent samples at 0–12 and 12–24 h.

The features/trends contained within this plot (scores plot) are more easily visualised when the SIMCA-P generated data is displayed in SPOTFIRE (Fig. 6b). Here the red points correspond to the first void, the blue points to 0–12 h and the yellow points to 12–24 h. Generally, it appears that the first voids were relatively spread out compared with the 0–12 and 12–24 h. Further, a slight separation was evident: the red points (first voids) tend to occupy the bottom half of the plot, while the blue (0–12 h) and yellow points (12–24 h) occupy the top half. Examination of the loadings plot (Fig. 7a) indicated that the first voids had generally larger peaks in bins 3.98, 7.82 and 7.54 (corresponding to hippuric acid), and "superbin" 3.1+ (creatinine), than the 0–12 and 12–24 h

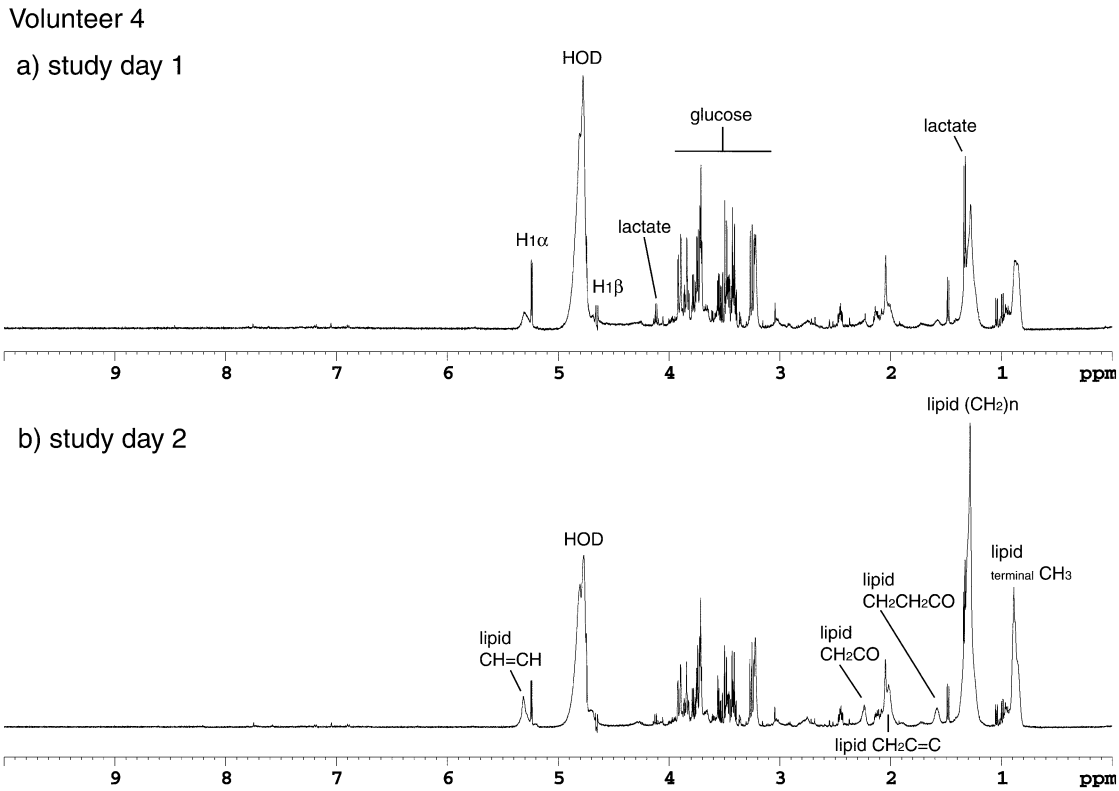


Fig. 3. ^{1}H NMR CPMG spectra of the plasma samples of subject 4 on study day 1 (a) vs. study day 2 (b).

urines. It also indicated that the first voids had generally smaller peaks in "superbin" 2.74+ (citric acid) and bin 3.26 (trimethylamine-*N*-oxide). In general, the first void urines were shown to be dominated by hippurate excretion while the urines at later time points contained additional aromatic amino acid resonances such as histidine and its derivatives, carnosine and anserine. These signals have been associated with a meat rich diet. The reconstructed and superimposed spectra of each urine samples within SPOTFIRE clearly highlighted this trend. The characteristic region of δ_{1H} 8.04–7.32 is shown in Fig. 7b.

We found, that splitting the plot into 12 panes—one for each subject for the two study days—was a more informative approach to examine the data-set for intra-subject variability (Fig. 8, corresponding to the scores plot in Fig. 6b). This representation showed that, despite clear inter-subject variability, the intra-subject variability over the two study days was low. That is to say, whilst the subjects are rather different from one another, a given subject is generally quite similar to himself across the two study days. This is an important observation if pre- and post-treatment samples were to be compared in a clinical study setting. As shown in Fig. 8, the 'trellis' representation of the urine spectra showed that there was little change for subjects 6, 7 and 9, across all three time-points and study days, whereas subject 12 was consistent from study day to study day in that his urine changed markedly after being put on the standard diet. This illustrates the importance of taking into account dietary influences and diurnal variation in sample composition in "omic" studies.

In summary, the metabonomic data was consistent for the 0–12 and 12–24 h urine samples, and the plasma samples (collected at a single time point). The low intra-subject variability in the plasma samples will allow treatment effects within

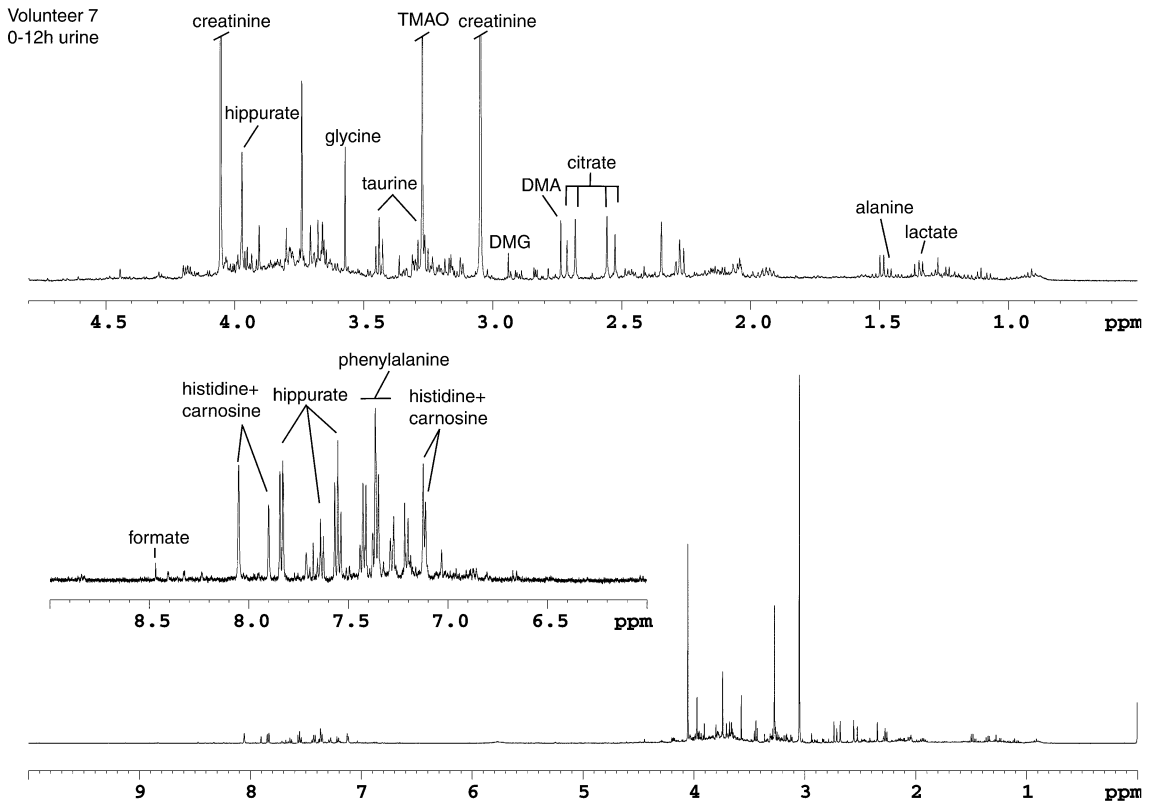


Fig. 4. Representative ^{1}H NMR spectrum of urine (subject 7, 0–12 h). Key: DMA, dimethylamine; DMG, dimethylglycine; TMAO, trimethylamine-*N*-oxide.

individuals to be observed. For urine samples, a standard diet, lifestyle restrictions and time of sample collection should be considered in clinical studies.

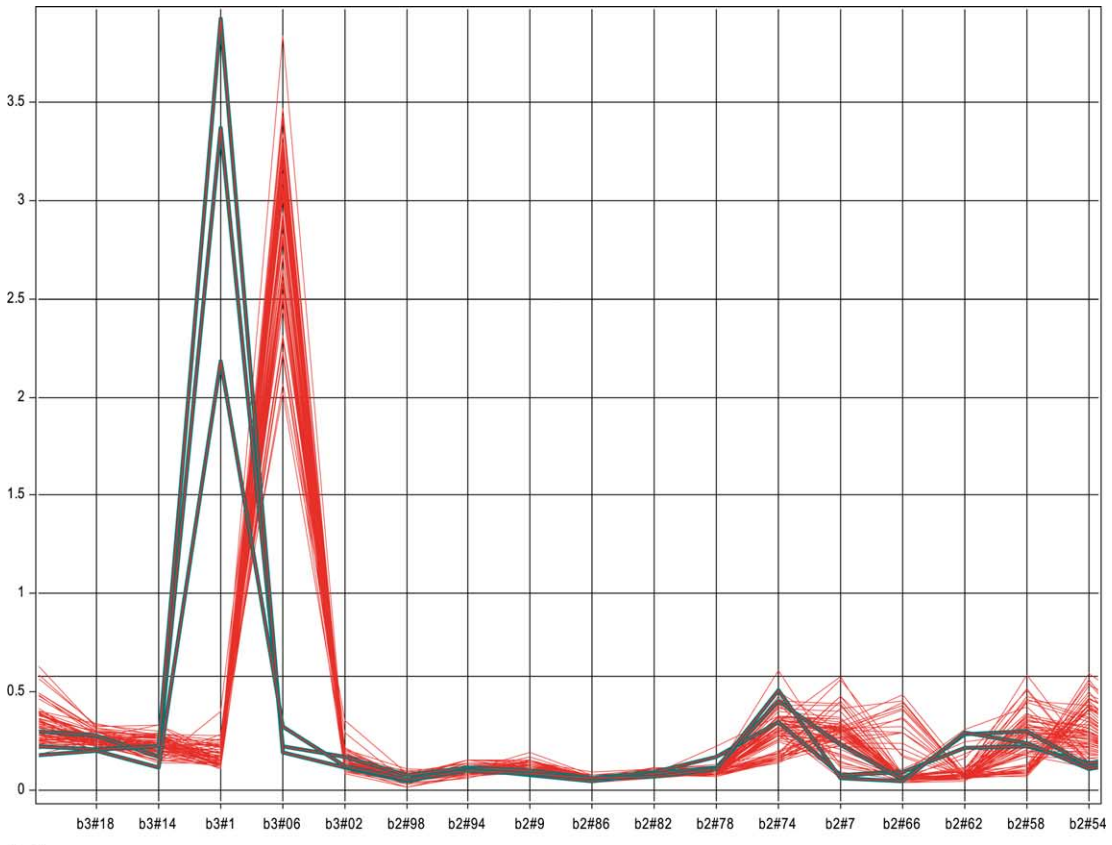
Table 1
Urinary pH values (mean ± S.D.)

Study day	First void	0–12 h	12–24 h
1	5.96 ± 0.51	7.27 ± 0.43	6.94 ± 0.27
2	6.04 ± 0.47	7.40 ± 0.44	6.74 ± 0.42

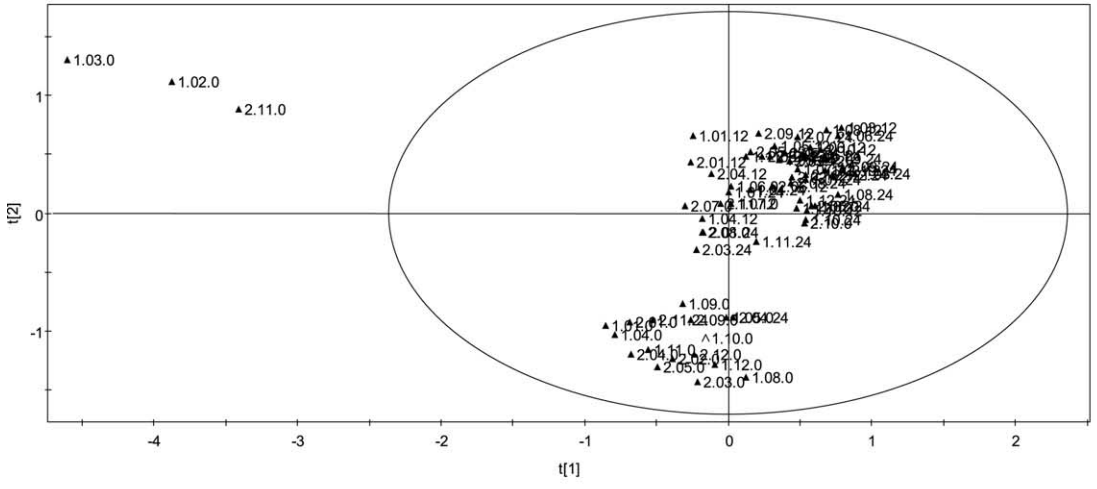
4. Conclusions

This study has shown that the plasma spectra from this group of subjects were subject to relatively low inter- and intra-individual variability across the two study days. Slight fluctuations were observed in the relative ratios of glucose to lipid concentrations. Hence, blood sampling at a 'precise' single time point, following a standard breakfast, gave consistent results.

The urines, collected at three different time points, were expected to be prone to dietary and lifestyle influences, and diurnal variation. However, surprisingly, although we found subjects to display significant inter-individual variability (representing normal genetic variation), intra-individual variability was less pronounced. Interestingly, the first void urines were rather more variable than either of the subsequent two voids (0–12 and 12–24 h), collected in the 'clinic' under standardised conditions. It seems evident that the first void diversity may reflect differences in the subject's lifestyle and diets. Hence, standardisation of diet, lifestyle and time of sample collection appear to



(a)



(b)

Fig. 5. (a) Data reduced and reconstructed spectra of all urine samples superimposed and displayed within SPOTFIRE (expansion of δ_{1H} 3.2–2.54) showing three ''outliers'' and (b) the corresponding PCA scores plot (PC1 vs. PC2) generated in SIMCA-P (label: study day–subject number–time point of collection).

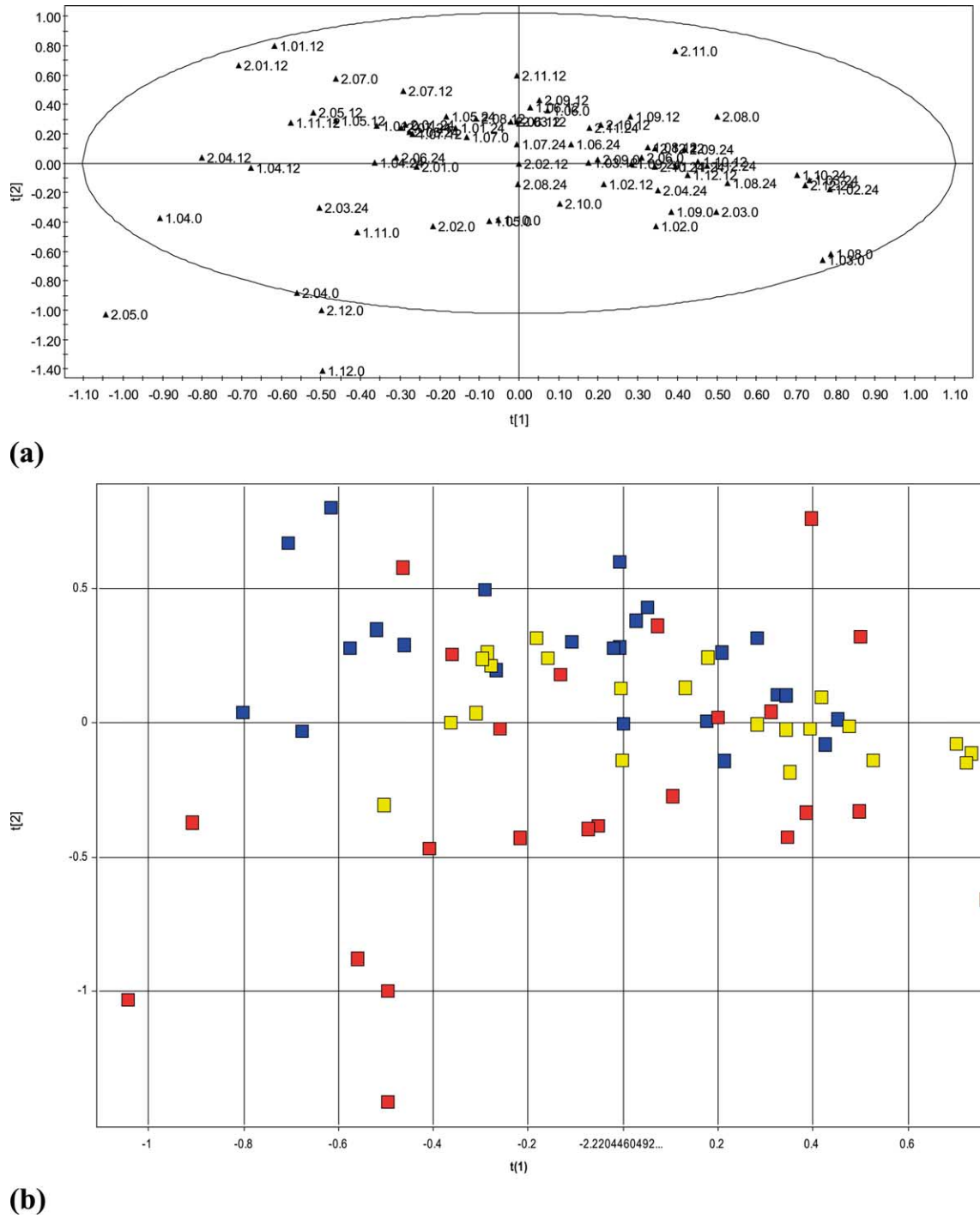
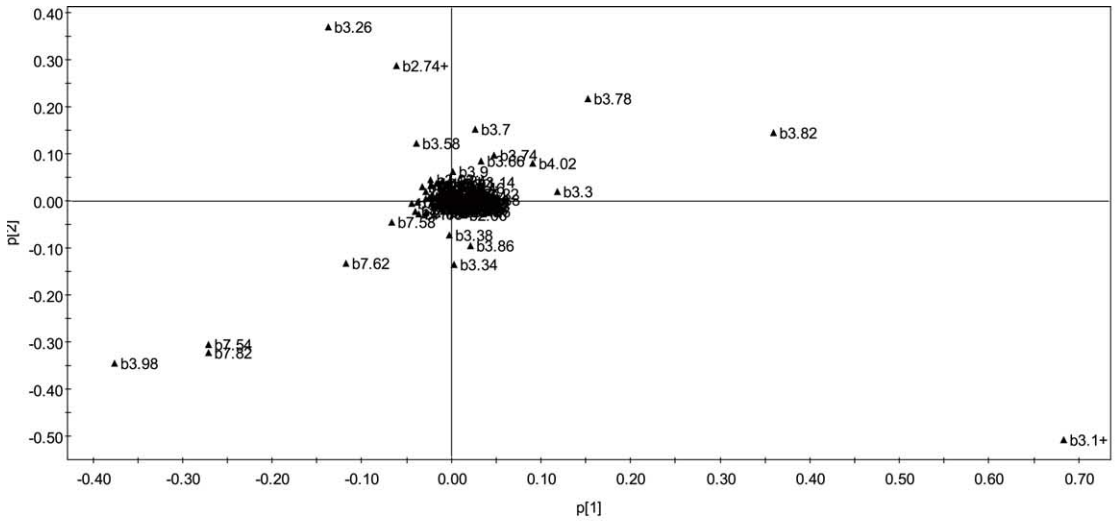
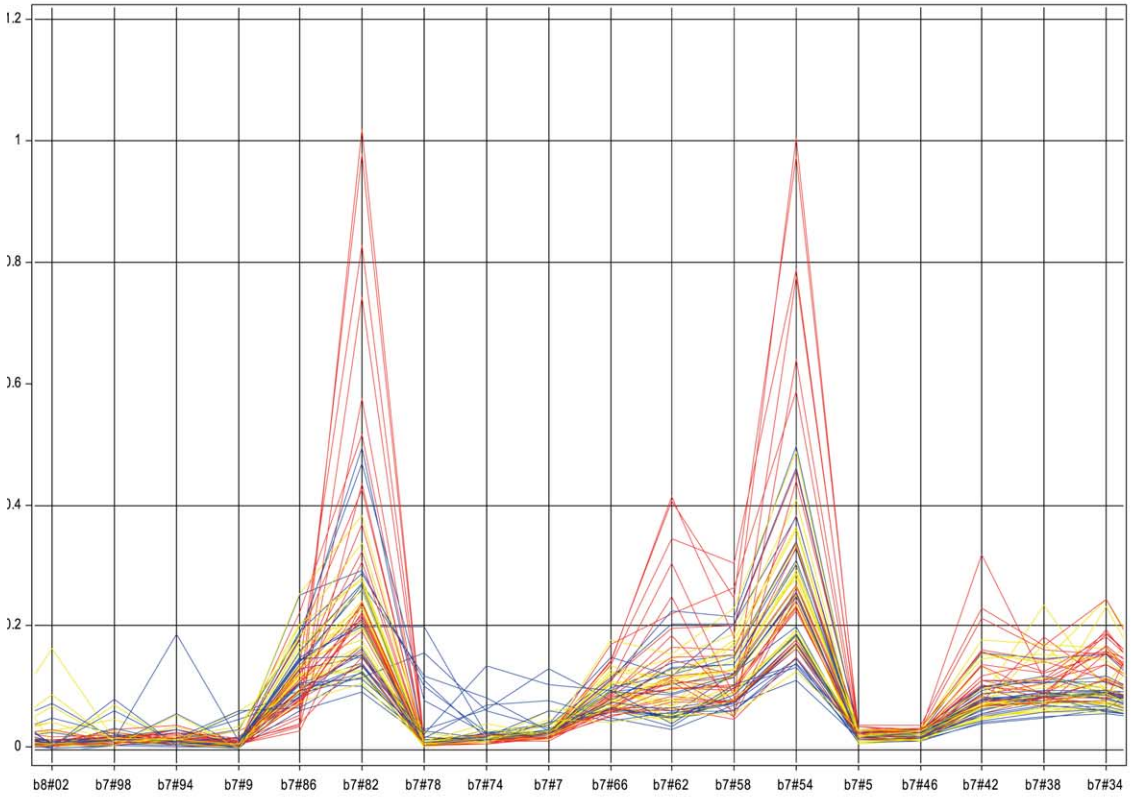


Fig. 6. (a) PCA scores plot (PC1 vs. PC2) of urines with outliers incorporated via ''superbins'' (3.1+ and 4.14+, and 2.74+) displayed in SIMCA-P (label: see Fig. 5) and (b) displayed in SPOTFIRE (red, first void; blue, 0–12 h; yellow, 12–24 h).



(a)



(b)

Fig. 7. (a) The corresponding loadings plot to Fig. 6, and (b) the superimposed, reconstructed urine spectra in SPOTFIRE, displaying the aromatic region of δ_{1H} 8.04–7.32. The decrease in hippurate signals over time is clearly observable. Key: red, first void; blue, 0–12 h; yellow, 12–24 h urines.

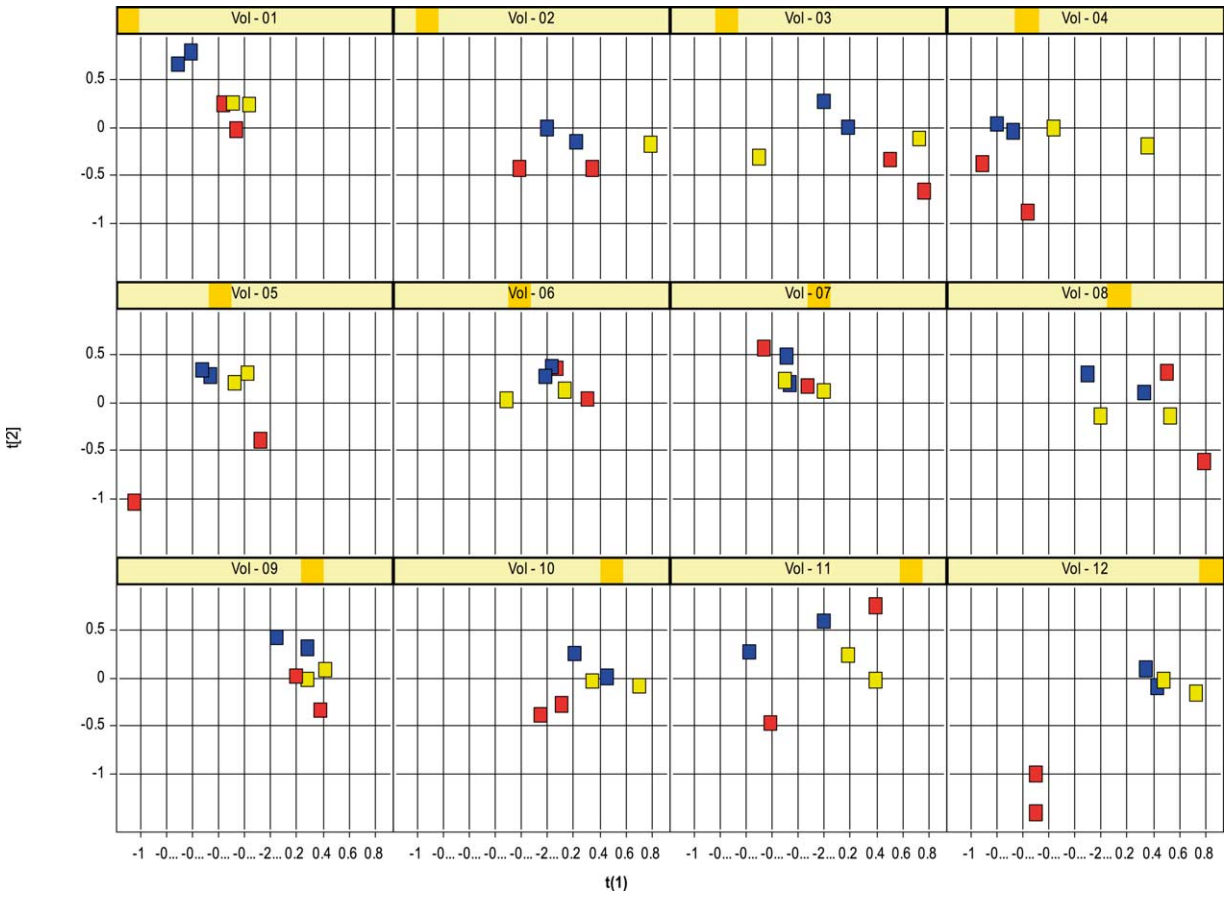


Fig. 8. PCA scores plot (PC1 vs. PC2, generated in SIMCA-P and displayed in SPOTFIRE) for each individual (plots are in consecutive order, subjects 1–12, from top left to bottom right; for colour coding, see Fig. 7).

play significant roles in clinical studies and will be of importance when biomarkers are to be interpreted reliably and correctly. Overall, this study provides reassurance that metabonomic data has acceptable variability and may highlight biomarkers of disease or toxicity that have clinical utility in monitoring the effects of drug therapy in early-phase clinical studies.

References

[1] J.K. Nicholson, J.K. Lindon, E. Holmes, Xenobiotica 29 (1999) 1181–1189.
[2] J.K. Nicholson, J. Connelly, J.C. Lindon, E. Holmes, Nat. Rev. Drug Disc. 1 (2002) 153–161.
[3] J.C. Lindon, J.K. Nicholson, E. Holmes, H. Antti, M.E. Bollard, H. Keun, O. Beckonert, T.M. Ebbels, M.D. Reily, D. Robertson, G.J. Stevens, P. Luke, A.P. Breau, G.H. Cantor, R.H. Bible, U. Niederhauser, H. Denn, G. Schlotterbeck, U.G. Sidelmann, S.M. Laursen, A. Tymiak, B.B. Car, L. Lehman-Mckeeman, J.-M. Colet, A. Loukaci, C. Thomas, Toxicol. Appl. Pharmacol. 187 (2003) 137–146.
[4] J.K. Nicholson, I.D. Wilson, Prog. NMR Spectrosc. 21 (1989) 449–501.
[5] J.T. Brindle, H. Antti, E. Holme, G. Tranter, J.K. Nicholson, H.W.L. Bethell, S. Clarke, P.M. Schofield, E. McKilligan, D.E. Mosedale, D.J. Granger, Nat. Med. 8 (2002) 1439–1443.
[6] C.L. Gavaghan, J.K. Nicholson, S.C. Connor, I.D. Wilson, B. Wright, E. Holmes, Anal. Biochem. 291 (2001) 245–252.
[7] M.E. Bollard, E. Holmes, J.C. Lindon, S.C. Mitchell, D. Branstetter, D. Zhang, J.K. Nicholson, Anal. Biochem. 295 (2001) 194–202.

[8] C.L. Gavaghan, E. Holmes, E. Lenz, I.D. Wilson, J.K. Nicholson, FEBS Lett. 484 (2000) 169–174.
[9] C.L. Gavaghan, I.D. Wilson, J.K. Nicholson, FEBS Lett. 530 (2002) 191–196.
[10] J.C. Lindon, E. Holmes, J.K. Nicholson, Prog. NMR Spectrosc. 39 (2001) 1–40.
[11] S. Deprez, B.C. Sweatman, S.C. Connor, J.N. Haselden, C.J. Waterfield, J. Pharm. Biomed. Anal. 30 (2002) 1297–1310.