PHYTOTHERAPY RESEARCH
Phytother. Res. **26**: 127–135 (2012)
Published online 12 July 2011 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/ptr.3547

Urinary ^{1}H-NMR Metabonomics Study on Intervention Effects of Soya Milk in Africans

Olumuyiwa L. Ogegbo,[1] Wimal Dissanyake,[1] John Crowder,[1] Olayinka Asekun,[2] Oluwole Familoni,[2] Christopher J. Branford-White[1] and S. W. Annie Bligh[1]*

[1]Institute for Health Research and Policy, London Metropolitan University, 166-220 Holloway Road, London N7 8DB, UK
[2]Department of Chemistry, Faculty of Science, University of Lagos, Akoka, Lagos, Nigeria

Metabonomics is an important tool in understanding the toxicological or therapeutic effects of interventions by analysing metabolic profiles and interpreting complex multi-dimensional spectroscopic/spectrometric data using multivariate data analysis. The objectives of this study were to evaluate the metabolic changes following a short-term 5 day soya milk intervention, and to investigate factors that influence soy-phytoestrogen metabolism focused on Africans based in either UK or Nigeria. ^{1}H-NMR metabonomics was applied to analyse urine samples collected at four phases I–IV (pre, days 3 and 5, and post) of the soy-intervention from African volunteers ($n = 40$ in total). Individual proton NMR spectra were visually and statistically assessed using multivariate analyses (MVA): principal component analysis (PCA) and (orthogonal-) partial-least square–discriminant analysis ((O-) PLS-DA). In addition, 22 endogenous metabolites were quantified using a Chenomx NMR suite. The results showed the levels of analysed endogenous metabolites (creatinine adjusted) present ranged from 4 μM to 12 mM with large inter-subject variances in acetate, acetone, lactate and trimethylamine. The MVA results showed high inter-individuality and sampling variances based on PCA score plots, and demonstrated soy metabolism to be significantly influenced by location and gender by both PLS-DA and O-PLS-DA. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: phytoestrogens; quantitation; endogenous metabolites; Nigeria; UK; statistical analysis.

INTRODUCTION

Metabonomics is a systematic interpretation of the metabolic responses of living systems to patho-physiological stimuli using multivariate statistical analysis of biological NMR spectroscopic data (Nicholson *et al.*, 1999). It is applied in analytical fields such as drug toxicology and discovery (Griffin, 2004; Lindon *et al.*, 2004), disease diagnosis (Gowda *et al.*, 2008; Lindon *et al.*, 2004) and nutrigenomics (Muller and Kersten, 2003; Rezzi *et al.*, 2007). Metabonomics seems to be more comprehensive than conventional *in vitro* studies due to its 'individualized' or higher subject-specificity and thus gives better insights into the integrated function of individual complex bio-systems at a system level (Nicholson *et al.*, 1999). Various researches involving induced diet metabonomics are now being used to gain a broader understanding of the biochemical effects of major bioactive components present in diets (Muller and Kersten 2003; Rezzi *et al.*, 2007). In this investigation, the intervened diet was soya milk, which is prepared from soybeans containing mainly phytoestrogens (isoflavones) including daidzein, genistein, glycitein, daidzin, genistin and glycitin. This group of compounds possesses significantly high estrogenic activities compared with estradiol, a human estrogen, due to their structural similarity (Knight and Eden, 1996). Further to this, several studies have reported their health benefits in combating hormone dependent diseases (Knight and Eden, 1996; Shu *et al.*, 2001; Somekawa *et al.*, 2001), bone health diseases such as osteoporosis (Vanharanta *et al.*, 1999), cardiovascular diseases (Somekawa *et al.*, 2001; Vanharanta *et al.*, 1999) and hypercholesterolemia (Somekawa *et al.*, 2001; Vanharanta *et al.*, 1999). These studies were based on epidemiological (Adlercreutz, 1995; Duffy *et al.*, 2007) and *in vitro* simulative studies, for instance, using cell lines (Rice and Whitehead, 2006).

Although there is growing interest in dietary phytoestrogens, particularly with soy products, little work has been studied and reported on the biochemical effects based on 'stimulated *in vivo*' study (metabonomics). The only published work in relation to this was by Solanky *et al.* (2005, 2003), who monitored 15 endogenous metabolites and identified their biochemical changes following a textured vegetable protein or miso intervention over a month on five to six subjects of unspecified ethnic origin and gender. Their reports suggested soy induced some changes in carbohydrate and lipid metabolism.

Phytoestrogen-based dietary studies involving urinary excretion of phytoestrogens, their metabolites and/or their dietary phytoestrogens intake on multi-ethnic populations including African-Americans have been reported (Horn-Ross *et al.*, 1997; Kolonel *et al.*, 2000; Park *et al.*, 2009; Schabath *et al.*, 2005). The general results of these studies suggested variations in the excretion levels in African-Americans compared with the other populations: Caucasian, Latina, Hawaiian and Japanese. However, there is no report of such a study on Africans in the African continent and overseas.

This study aimed to evaluate the differences, if any, in the urinary metabolic profile on the effects of intervention of a high phytoestrogen drink, soya milk, in Africans

* Correspondence to: S. W. Annie Bligh, Institute for Health Research and Policy, London Metropolitan University, 166–220 Holloway Road, London N7 8DB, UK.
E-mail: a.bligh@londonmet.ac.uk

Received 25 February 2011
Accepted 07 May 2011

in UK and in Nigeria. Twenty-two urinary endogenous metabolites were quantitated in two cohorts (i.e. location) of Africans: London (UK) and Lagos (Nigeria). The biochemical changes were interpreted in relation to key metabolisms, and statistical multivariate analysis (MVA) undertaken for pattern recognition study on three main hypothesized factors: location, gender and age, which could influence soy metabolism. None of these factors have been previously reported and/or considered in any soy-phytoestrogen induced metabonomic study in humans.

MATERIALS AND METHODS

Chemicals. Monobasic sodium phosphate, dibasic sodium phosphate, trimethylsilyl-2,2,3,3-tetradeuteropropionic acid (TSP-d_4) and deuterated water (D_2O) were purchased from Fisher Scientific, UK. Distilled water was purified with Millipore Simplicity 185 (Millipore Corp., USA) at 18.2 $m\Omega.cm^{-1}$.

Material. Alpro soya-milk was procured from local supermarkets at their respective locations (London, UK and Lagos, Nigeria) for each cohort. In both procurements of soya-milk, the serving of cholesterol- and lactose-free Alpro soya milk is compliant with the guideline daily allowance (GDA) of 200 mL containing approximately 8 mg daidzein and 10 mg genistein.

Subjects and sample collection. Ethical approval for this research study was reviewed and approved by the London Metropolitan University human research ethics committee and all procedures complied with the National Health and Medical Research Council standards. Forty healthy African subjects were recruited: 20 each from UK and Nigeria; each sub-grouped into 10 males and 10 females. All volunteers had been resident in their respective locations for at least 2 years prior to recruitment. All subjects were interviewed, and signed informed consents were obtained from all participants before commencement of the study. All participant samples and information were confidential and remained completely anonymous. Factors that could influence variations such as age, BMI, lifestyle habits such as smoking, drinking and exercising, were noted; however, no inclusion/exclusion criteria were used prior sampling analysis.

On the induction day and a week after intervention, both subject groups (UK and Nigeria) supplied spot urine samples collected on day 1 (Phase I/pre-) and day 12 (Phase IV/post-). On commencing the intervention, Alpro soya milk at the guideline daily allowance (GDA) of 200 mL was consumed daily for five consecutive 24 h periods, and two different early morning urine samples were collected on day 3 (Phase II) and day 5 (Phase III). All samples were collected in the period 08:00–12:00 of the selected days and stored without preservatives at −80°C prior to NMR analysis, which was performed within a month.

Sample preparation. Phosphate buffer (0.1 M, pH 7.4) was prepared in 30:70 of D_2O: water mixture containing 1.44 mM of TSP-d_4. An aliquot of urine sample (600 μL) was added to 300 μL phosphate buffer solution in a 0.5 mL polypropylene Microfuge tube (Fisher Scientific, UK). Urine was left to stand at room temperature for 10 min before centrifugation at 13000 × **g** for 10 min using an Eppendorf centrifuge 5804R (Germany) at 4°C. The supernatant (850 μL) was transferred to a 5 mm NMR tube for ^{1}H-NMR analysis. This sample preparation procedure was carried out in triplicate for each sample.

^{1}H-NMR spectroscopy. ^{1}H-NMR spectroscopic analyses were measured using a Bruker® Avance AV-500 NMR spectrometer at 500.13 MHz and 303 K. The acquisition type used for the collection of the ^{1}H-NMR spectrum was by standard pulse sequence for water peak suppression obtained using 1D-NOESY pulse sequence with water pre-saturation (Claridge, 1999). The sequence used was [RD-90°-t_1-90°-t_m-90°-FID], where the recycle delay (RD) was 2 s, fixed interval (t_1) was 10 μs and mixing time (t_m) was 100 ms.

All spectra were acquired at 14.0019 ppm (7002.801 Hz) sweep width, 2.34 s acquisition time and 128 transients. The FID spectra were zero filled to 32 k data points, exponential 0.3 Hz line broadened, Fourier transformed and manually fine-tuned (phase and baseline correction) using the pre-installed computer-aided software for data acquisition, and analysis was Bruker Topspin software-version 2.0 (Bruker Biospin, Rheinstetten, Germany).

Data processing: identification/quantitation of endogenous metabolites and spectral data binning. Acquired ^{1}H-NMR spectra were exported into the Chenomx NMR suite 5.1 for identification and quantitation of metabolites. This was supported by the complementary database containing almost 300 metabolite compounds (standards) and, with the aid of a known concentration of TSP reference, the quantification of urinary metabolites was possible. Furthermore, identification and quantitation of individual compounds were established from the chemical shift and splitting patterns (multiplicities) of each standard.

Two main types of binning were considered: target binning and whole spectral binning. Target binning is a new advanced technique where only identified and quantified (in this case by Chenomx Profiler), signals/peaks are data-reduced into customized varied width sized bins prior to multivariate statistical analysis. Alternatively, in the standard spectral binning, ^{1}H-NMR spectra were sequentially data-reduced using an Analysis of Mixtures software package (Bruker AMIX, Version 3.6.8, Bruker Biospin GmbH, Germany) across the ^{1}H-NMR regions δ 0.0–9.0. This resulted into 225 simple sequential rectangular integral segments of 0.04 ppm width-sized simple rectangular buckets (bins). The integration mode used was the sum of intensity with the spectra being scaled to the reference (TSP).

Statistical analysis. Basic statistical analyses: Student's *t*-tests and groups' comparison by means of one-way ANOVA were performed using Minitab (version 15). The resulting binned data processed in AMIX were exported into an Excel spreadsheet (Microsoft 2007) due to compatibility with the SIMCA–P + multivariate data analysis software (version 12.0.0.0, Umetrics, Umeå, Sweden). Data-reduced variables were imported from Excel (Microsoft 2007) into SIMCA-P+. Multivariate analysis using pattern recognition techniques were performed applying standard SIMCA-P + applications. In both generated bin types (target profiled and whole

spectral computer-generated bin), data were mean centred and Pareto scaled for the MVA.

RESULTS AND DISCUSSION

By visual assessment of ^{1}H-NMR spectra, there were obvious metabolic changes in each phase of intervention (inter-sampling) and subject (inter-subject) sampling (Fig. 1A, B respectively). On examining Phase I (pre-soy, day 1) and Phase IV (post-soy, day 12), no significant changes in their overall metabolic profile were noted, however, there was a slight increase in the trimethylamine signal (δ: ~2.2), three unidentified signals at δ: 4.4 (doublet), δ: 5.3 (doublet) and δ: 8.9 (singlet) observed on Phase IV. Additionally, when comparing the consumption of soya milk on days 3 and

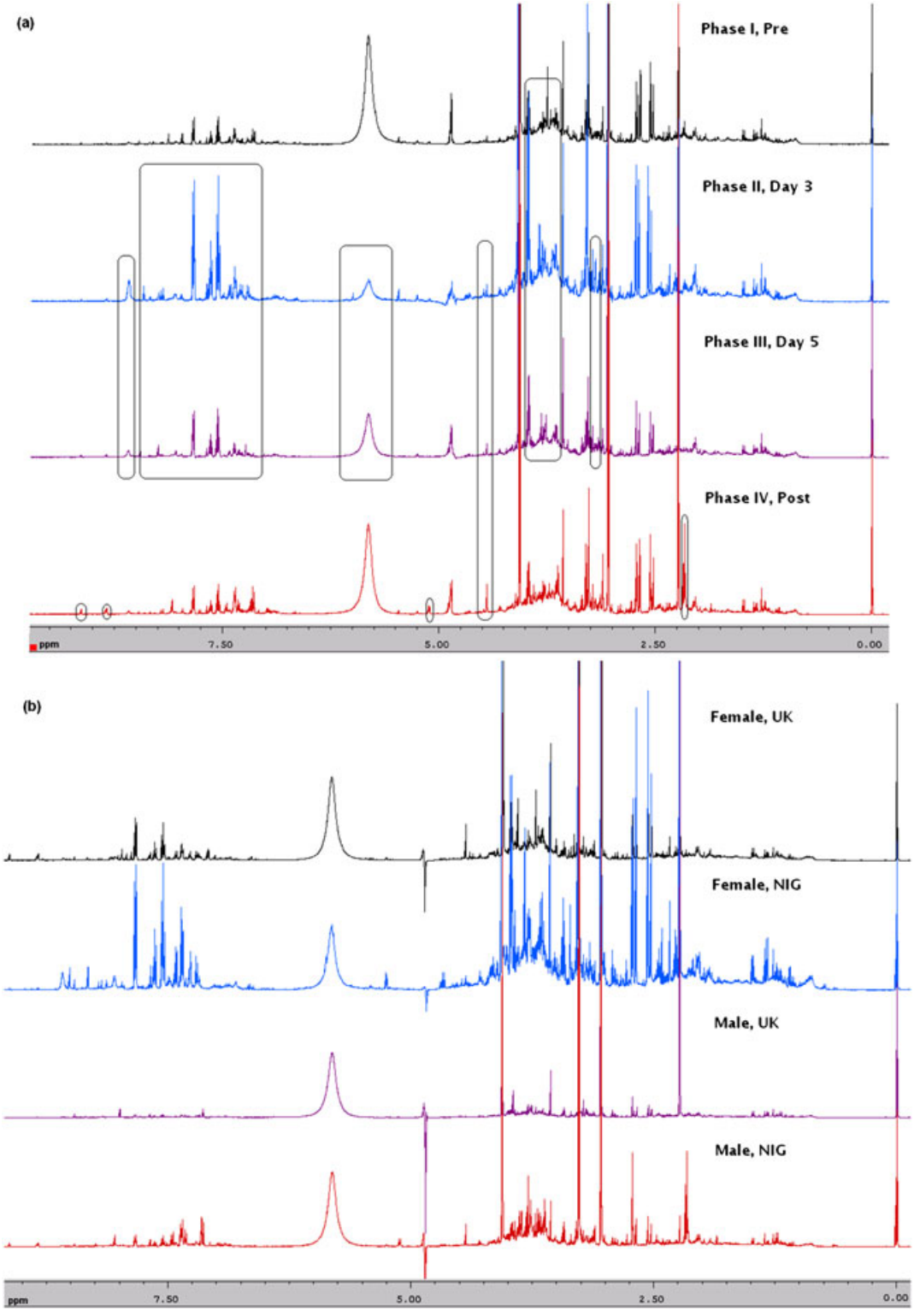

Figure 1. ^{1}H-NMR spectra (a) inter-sampling comparison of a subject's metabolic profile changes; (b) inter-subject comparison of metabolic profile changes.

5 with pre- and post- soy, increased variances in certain regions of the ^{1}H-NMR spectrum were observed (Fig. 1A). The regions covered were aromatic (δ: 7.4 to 7.8), saccharides (δ: 3.4 to 4.0), creatine (δ: 4.0 to 4.1; 3.0 to 3.1), creatinine phosphate (δ: 3.9 to 4.0; 3.0 to 3.1), dimethylglycine (δ: 3.25 to 3.35), formate (δ: 8.4) and an unidentified signal (δ: 4.4). The aromatic region changes could be due to an increase in hippurate level, polyphenol metabolism (i.e. possibly traces of phytoestrogen and their gut-microflora metabolites) and/or β-oxidation of fatty acids group (Solanky *et al.*, 2003, 2005). Also, it was noted that the soy-phytoestrogens influence on endogenous metabolites was greatest on Phase II (day 3) (Fig. 1a), which is the first urine sample collected following the ingestion of soya milk.

In respect to the different subjects' spectra, there were noticeable metabolic profile changes in inter-subjects as represented in Fig. 1b. Lenz and co-workers previously highlighted this phenomenon showing the high degree of variation within the human population based on visual NMR spectra inspection and 'blind' PCA study (Lenz *et al.*, 2004). In this study, comparison of the ^{1}H-NMR spectra of subjects based on their different locations and genders showed noticeable changes in certain regions, especially those linked to the aromatic (δ: 6.8 to 8.5), saccharides (δ: 3.4 to 4.0) and selected aliphatic (δ: 2.4 to 3.0) regions. Conversely, insignificant or no changes were visually spotted in the spectrum δ: 0.8 to 1.6, and the urea peak had no specific pattern in either location or gender groups.

Visual inspection confirmed differentiation of gender, whereby female subjects generally showed more metabolites than males, and the Nigerian (NIG) subjects had more metabolites than the UK counterparts.

Quantitation of endogenous metabolites

Phytoestrogens and their metabolites were not specifically investigated by NMR spectroscopy due to their poor solubility in aqueous solution and hence the concentration was too low to be detected. The main focus of this investigation was solely on the quantitation of endogenous metabolites after the intervention of soya milk (Table 1).

On analysing the entire metabolite levels prior to location or gender grouping, in Phase I (pre-soy, day 1), half of the metabolites had great variances based on their corresponding high standard deviation (SD) values. However, during intervention Phase II and III (day 3 and 5, respectively), it was observed that the variances in terms of SD were slightly less pronounced in all 22 metabolites apart from acetate, acetone, lactate, succinate and TMA. This could be due to the converging of groups (group uniformity) resulting from subjects having similar biochemical effects and response to the soya milk consumption. Finally in Phase IV (post-soy, day 12), similar increase or decrease effects as during intervention phases II and III were observed.

Data treatment by subject grouping was used to improve the statistical analysis. One-sample *t*-tests using Minitab were carried out to evaluate significant differences between groupings based on location and gender. In this study, the baseline (pre-soy) creatinine adjusted levels of metabolites in the UK and NIG cohorts' urine are detailed in Table 2. The unadjusted levels were comparable to the results reported previously in Canada, 0.830 (citrate), 0.188 (creatine), 6.830 (creatinine), 0.809 (glycine), 1.012 (hippurate) and 0.097 (lactate) mM (Solanky *et al.*, 2005; Weljie *et al.*, 2006). The data given here are consistent with our UK cohort. The data

Table 1. Concentration (mM) levels of metabolites at four different phases of intervention

Metabolite	Phase I		Phase II		Phase III		Phase IV	
	Mean	SD (±)	Mean	SD (±)	Mean	SD (±)	Mean	SD (±)
Acetate	0.046	0.107	0.137	0.247	0.106	0.318	0.120	0.255
Aceto-acetate	0.016	0.022	0.012	0.005	0.012	0.006	0.012	0.006
Acetone	1.376	4.307	0.949	1.669	1.123	1.299	1.532	4.266
Alanine	0.038	0.057	0.037	0.018	0.034	0.025	0.032	0.015
Betaine	0.016	0.015	0.013	0.008	0.013	0.009	0.014	0.012
Choline	0.007	0.009	0.007	0.005	0.009	0.008	0.006	0.005
Citrate	0.202	0.117	0.200	0.137	0.174	0.125	0.191	0.121
Creatine	0.050	0.034	0.050	0.034	0.071	0.078	0.041	0.022
Creatine phosphate	0.062	0.023	0.069	0.034	0.077	0.040	0.066	0.031
Dimethylamine	0.039	0.023	0.032	0.009	0.036	0.013	0.032	0.007
Formate	0.020	0.031	0.024	0.021	0.032	0.068	0.025	0.035
Glycine	0.172	0.101	0.216	0.145	0.229	0.170	0.203	0.124
Hippurate	0.125	0.080	0.114	0.092	0.161	0.143	0.123	0.136
Lactate	0.032	0.037	0.035	0.033	0.033	0.062	0.025	0.019
Methylamine	0.005	0.004	0.004	0.001	0.004	0.003	0.004	0.005
N,N-Dimethylglycine	0.007	0.010	0.007	0.004	0.007	0.003	0.006	0.003
Pyruvate	0.011	0.014	0.010	0.003	0.009	0.003	0.009	0.003
Succinate	0.014	0.012	0.021	0.040	0.019	0.040	0.024	0.059
Trimethylamine	0.004	0.005	0.006	0.009	0.008	0.020	0.007	0.011
Trimethylamine-*N*-oxide	0.053	0.056	0.042	0.030	0.049	0.044	0.039	0.031
Urea	11.367	6.938	14.714	8.582	15.681	7.909	11.713	7.447

SD, standard deviation.

Table 2. Creatinine-adjusted concentration (mM) levels of metabolites in location- and gender- grouped control samples (Phase I intervention)

Metabolite	UK						Nigeria					
	Male			Female			Male			Female		
	Mean	SD (±)	*p*-value	Mean	SD (±)	*p*-value	Mean	SD (±)	*p*-value	Mean	SD (±)	*p*-value
Acetate	0.018	0.024	0.0058	0.014	0.008	0.0006	0.116	0.202	0.12[a]	0.053	0.099	0.12[a]
Aceto-acetate	0.024	0.037	0.0091	0.022	0.027	0.028	0.008	0.003	0.0001	0.013	0.006	0.0001
Acetone	2.097	5.485	0.28[a]	2.917	7.459	0.25[a]	0.387	0.315	0.0062	0.762	0.940	0.031
Alanine	0.069	0.124	0.13[a]	0.034	0.013	0	0.029	0.012	0.0001	0.030	0.014	0.0001
Betaine	0.019	0.030	0.1[a]	0.018	0.009	0.0002	0.015	0.006	0.0001	0.017	0.008	0.0001
Choline	0.010	0.011	0.03	0.008	0.005	0.0005	0.009	0.016	0.12[a]	0.005	0.003	0.0003
Citrate	0.205	0.189	0.012	0.229	0.064	0	0.167	0.094	0.0007	0.214	0.0126	0.0005
Creatine	0.039	0.018	0.0002	0.062	0.039	0.0007	0.034	0.010	0	0.068	0.049	0.0018
Creatine phosphate	0.066	0.028	0.0001	0.073	0.014	0	0.051	0.018	0	0.060	0.028	0.0001
Dimethylamine	0.043	0.047	0.024	0.033	0.005	0	0.044	0.019	0.0001	0.042	0.010	0
Formate	0.032	0.055	0.12[a]	0.016	0.015	0.0059	0.020	0.038	0.16[a]	0.014	0.011	0.0029
Glycine	0.121	0.053	0.0001	0.238	0.136	0.0004	0.176	0.122	0.0026	0.171	0.070	0
Hippurate	0.103	0.074	0.0033	0.149	0.053	0	0.109	0.113	0.02	0.114	0.063	0.0003
Lactate	0.045	0.069	0.087[a]	0.026	0.013	0.0001	0.034	0.043	0.043	0.029	0.012	0
Methylamine	0.006	0.008	0.064[a]	0.004	0.002	0	0.003	0.001	0	0.005	0.003	0.0011
N,N-Dimethylglycine	0.011	0.021	0.14[a]	0.005	0.002	0	0.007	0.003	0.0001	0.006	0.003	0.0003
Pyruvate	0.019	0.030	0.089[a]	0.010	0.003	0	0.010	0.004	0.0001	0.007	0.002	0
Succinate	0.012	0.014	0.028[a]	0.015	0.009	0.0006	0.016	0.018	0.027	0.014	0.009	0.0008
Trimethylamine	0.005	0.006	0.046	0.003	0.000	0	0.003	0.001	0	0.006	0.008	0.043
Trimethylamine-*N*-oxide	0.045	0.034	0.0045	0.040	0.021	0.0002	0.064	0.052	0.0061	0.085	0.095	0.02
Urea	11.509	5.749	0.0003	16.628	9.227	0.0003	11.492	6.169	0.0005	7.497	4.429	0.0005

SD, standard deviation.
[a]$p > 0.05$.

similarity could be due to both samplings being collected within populations in Western countries.

In the Phase I (pre-soy, day 1), all except nine metabolites in the UK male group had values of $p > 0.05$, which implied significant differences in terms of gender grouping (Table 2). On the other hand, only acetone was found to show a significant difference within the UK female group ($p = 0.25$, which is > 0.05). In the NIG cohort, only acetate, choline and formate were significantly different (p values > 0.05) in the male group, whilst only acetate was significantly different ($p = 0.12$, which is > 0.05) in the female group. This result indicates that females (regardless of location) have a very similar metabolic profile.

Biochemical effect and relevance of soy-phytoestrogen

Dietary use of soy-protein has been proposed to reduce/prevent bone loss (Arjmandi *et al.*, 2005; Lydeking-Olsen *et al.*, 2004). Collagen as a major component of bone is one of the essential factors for bone and joint function and its increased production prevents osteoarthritis and osteoporosis (Bailey and Knott, 1999; Roudsari *et al.*, 2005). The key amino acids present in collagen, glycine, proline and alanine, could be used as monitors for collagen synthesis. In this study, both glycine and alanine significantly increased, therefore implying an increased collagen synthesis with soya milk intake, and hence, confirming its use as an osteoporosis treatment.

Also following soy consumption, there was an observed increase in the concentrations of acetate (ketone body), alanine, choline; these are responsible for lipid metabolism and biosynthesis. Solanky also reported these results (in the plasma profile) being evidence for the positive contributing effect of soy in cardiac and renal disease groups (Solanky *et al.*, 2005, 2003).

Lastly, the results shown in this study suggest gluconeogenesis or glycogen breakdown and inhibited glycolysis or glycogen synthesis as a result of lower lactate, pyruvate and citrate levels with increased creatine phosphate level (Wiback and Palsson, 2002). This finding contradicts the previous studies of Solanky *et al.* (2003), but who later reported elevated levels of these metabolites suggesting glycolysis inhibition (Solanky *et al.*, 2005).

Statistical multivariate analyses

Of the two different spectral normalization modes, the total area and standardized area, the standardized area was used in this study. All areas of the spectrum were calibrated based on the internal standard concentration and area. This normalization is known to compensate for spectrum-to-spectrum variability, i.e. more reproducible results are obtained even though dilution effects in the urine are not considered.

For evaluation and exploration of the differences among the four phases of intervention, all the ^{1}H-NMR spectral data were selected and digitized for MVA using the correlation method of PCA-X, PLS-DA and O-PLS-DA. Prior to confirmation of the hypothesized factors influencing soy metabolism, random test groups were selected from the whole group and analysed for the first

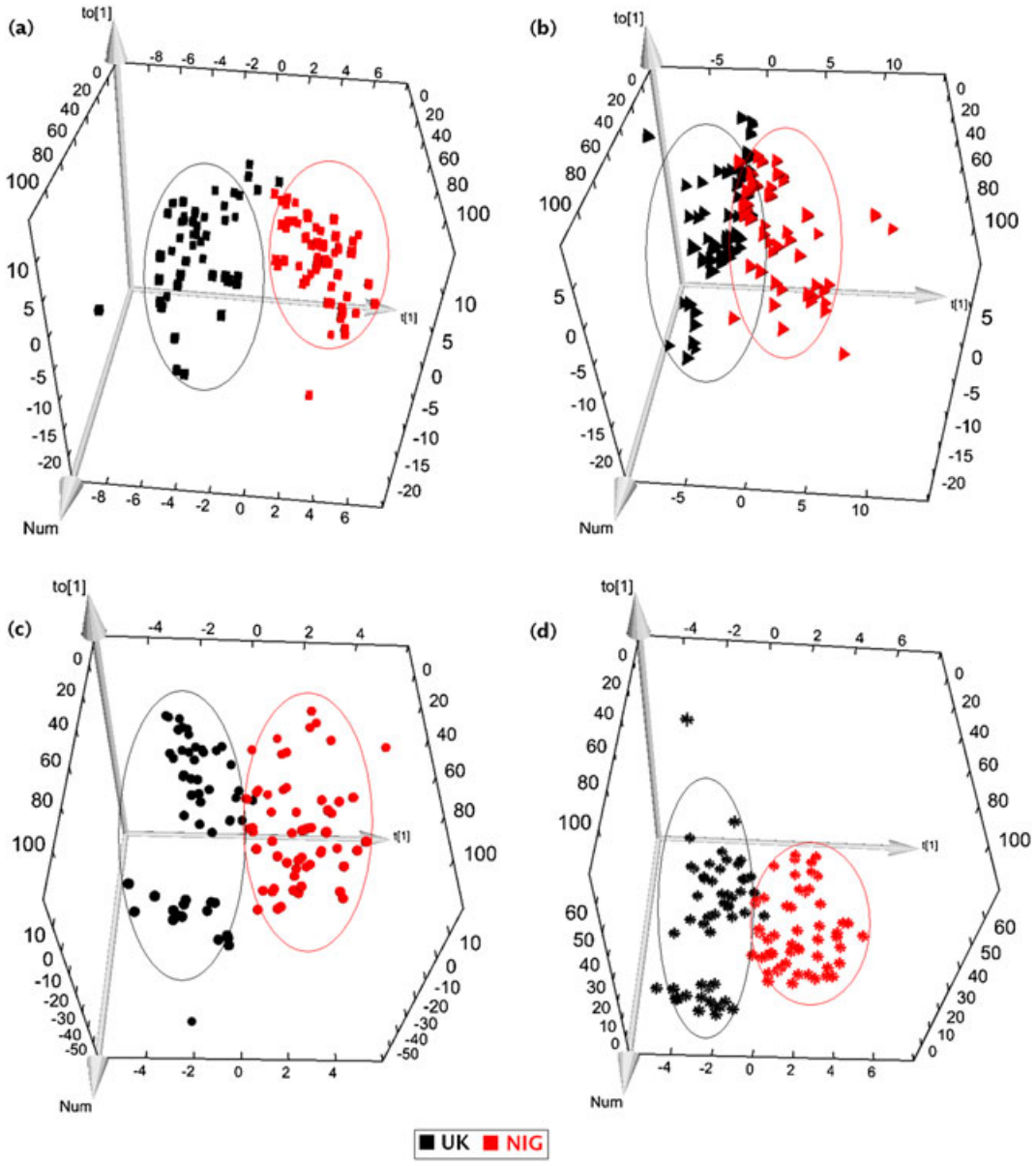

Figure 2. Influence of demography on soy-metabolism using supervised MVA (O-PLS-DA) 3D-score plot of phase I-IV of soy-intervention. (a) Phase I (pre-): day 1; (b) Phase IV (post-): day 12; (c) Phase II: day 3; (d) Phase III: day 5.

instance as the 'training set'. For example in PCA, a random selection of the whole cohorts 'training set' was subjected to 'blind PCA' and these scores were identified and checked that they grouped accordingly based on the grouping criteria provided. The hypothesized factors (grouping criteria) proposed to influence inter-variations in MVA include: demographical location, gender and age.

Based on investigating demographic differences, even though an unsupervised PCA-X showed partial differentiation between the two location groups supported by the training set analysis, supervised MVA, PLS-DA and O-PLS-DA, were applied to validate the discrimination of these two groups. With both MVA analyses, location was found to be discriminatory factor based (Fig. 2), where two obvious group clusters were observed in the score plots when a random selection of the data as the training set was used. The remaining data set 'test set' proved and validated that location was a discriminatory factor. Relatively clear group separations were observed in the PC1, PC2 and PC3 score plots based on demographic variation, i.e. location - Nigeria and UK, in all four phases (days 1–12) independently. However in 3-D, the Phase IV (post-) showed a partial overlap-grouping of the two locations considered (UK and Nigeria cohorts) compared with the other three phases.

Similarly, in order to assess the gender effect on the metabolic responses, supervised and unsupervised MVA models were generated using the data different phases of intervention. Likewise, PLS-DA and O-PLS-DA showed clear gender-discrimination plots with acceptable R^2 and Q^2 cumulative values (> 0.6).

Following this, the combination of location and gender grouping was considered in phases I–IV of soy-intervention using PLS-DA and O-PLS-DA (Fig. 3). This showed the separation of groups based on location in PC 1 and gender in PC 2. Clear group separation was achieved in phase I (pre-soy, day 1) and phase II (day 3), and thereafter, merging and overlapping of some groups were observed in phase III (day 5) and phase IV (post-soy, day 12), especially the female subjects with slight convergence. This could possibly be

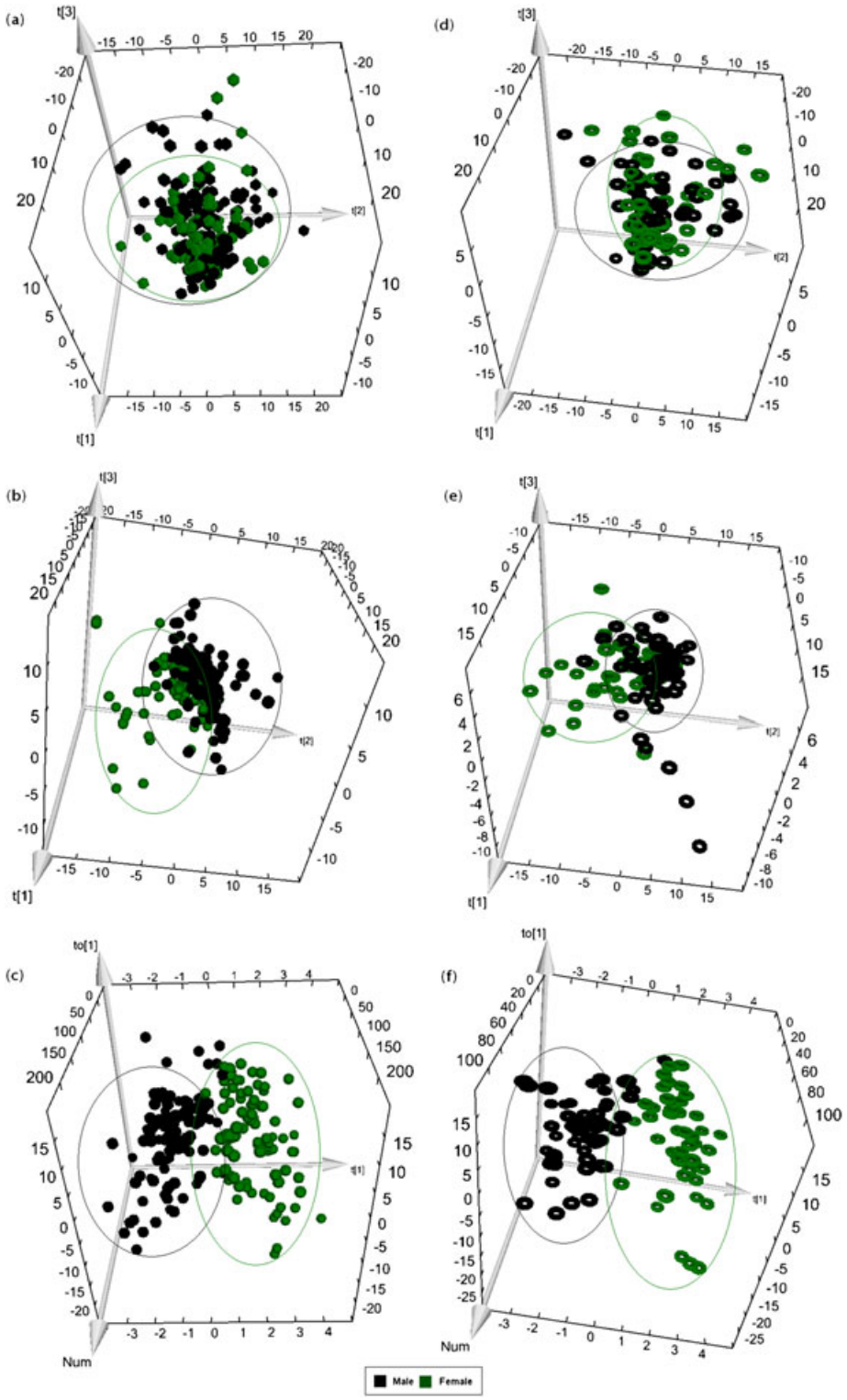

Figure 3. Influence of gender on soy-metabolism using three MVA 3D- score plots. Phase II and III combined dataset: (a) PCA-X score plot; (b) PLS-DA score plot; (c) O-PLS-DA score plot. Phase I only dataset: (d) PCA-X score plot; (e) PLS-DA score plot; (f) O-PLS-DA score plot.

due to similarities in the beneficial effects of soy-phytoestrogens, particularly within female subjects regardless of location.

Lastly, based on the MVA exploration of the effects of age on soy-metabolism, there are no clear group clusters as hypothesized. This is due to the significant overlapping of clusters with no definitive pattern observed from the 40 volunteers (approximate average: 28 ± 7 years; range: 21–45 years).

Target binning

Identified and quantified data of the 24 metabolites from Phase I–IV intervention samples using the Chenomx suite profiler were exported into SIMCA-P + for statistical analysis. By means of the unsupervised MVA tool, PCA-X was performed on the dataset and insufficient discrimination of these groups was observed. In support of the score plots in PLS-DA and O-PLS-DA (Fig. 4)

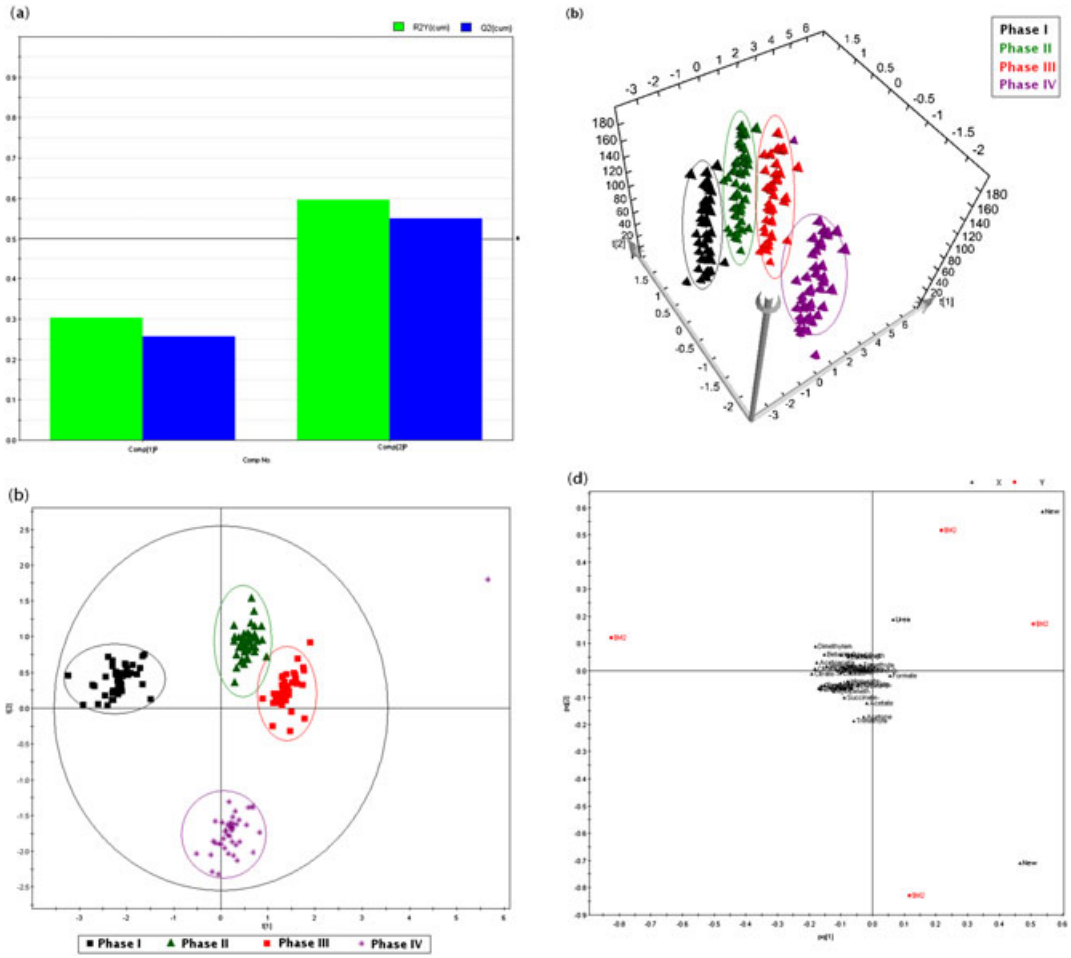

Figure 4. Supervised MVA of Chenomx profiled data. (a) Model overview plot; (b) 2D-O-PLS-DA score plot; (c) 3D-O-PLS-DA scoreplot; (d) loading scatter plot. * indicates R^2 and $Q^2 > 0.5$, which is the critical expected for good model fit and predictivity, respectively.

analysis, it has been shown that target binning multivariate statistical analysis is an equally good binning technique. Additionally, with the R^2 and Q^2 cumulative values having ≥ 0.5 each in both PLS- DA models further proves the effectiveness of target binning.

CONCLUSION

Twenty-two metabolites in African subjects' urine were simultaneously assessed with the aid of the Chenomx suite after the intervention of soya milk. The levels of these metabolites at the baseline (before soy-intervention) were found to be comparable to other reported studies ranging from 4 μM to 12 mM (creatinine adjusted). Most of the metabolites showed significant variances in the general cohort. Prior to further analysis, subject taxonomy was performed on the dataset based on proposed factors influencing soy metabolism. Of the proposed factors, two main factors were shown to influence differences in individual metabolism following soy-intervention. These two factors were location (Nigeria and UK cohorts) and gender (male and female groups). These results were confirmed by MVA studies using partial least square-discriminatory analysis (PLS-DA), which were shown to be the better expressive technique to differentiate the proposed theories of influencing factors. The observations made in this study with regard to influencing factors suggest that the use of location and gender have potential to be useful for a more vital understanding of diet/drug metabolism by *in vivo* relationship of the compound or the drug of interest's metabolites on endogenous metabolites.

Acknowledgement

This research was supported by a grant from the Institute for Health Research and Policy (O.L.O.), London Metropolitan University. We would like to thank all the volunteers who participated in this study.

Conflict of Interest

The authors have declared that there is no conflict of interest.

REFERENCES

Adlercreutz H. 1995. Phytoestrogens: epidemiology and a possible role in cancer protection. *Environ Health Perspect* **103**: 103–112.

Arjmandi BH, Lucas EA, Khalil DA *et al*. 2005. One year soy protein supplementation has positive effects on bone formation markers but not bone density in postmenopausal women. *Nutr J* **4**: 8.

Bailey AJ, Knott L. 1999. Molecular changes in bone collagen in osteoporosis and osteoarthritis in the elderly. *Exp Gerontol* **34**: 337–351.

Claridge TDW. 1999. High resolution NMR techniques in organic chemistry (1st edn). Elsevier: New York.

Duffy C, Perez K, Partridge A. 2007. Implications of phytoestrogen intake for breast cancer. *CA Cancer J Clin* **57**: 260–277.

Gowda GN, Zhang S, Gu H, Asiago V, Shanaiah N, Raftery D. 2008. Metabolomics-based methods for the early disease diagnostics. *Expert Rev Mol Diagn* **8**: 617–633.

Griffin JL. 2004. The potentials of metabonomics in drug safety and toxicology. *Drug Discov Today Tech* **1**: 285–293.

Horn-Ross PL, Barnes S, Kirk M, Coward L, Parsonnet J, Hiatt RA. 1997. Urinary phytoestrogen levels in young women from a multiethnic population. *Cancer Epidemiol Biomarkers Prev* **6**: 339–345.

Knight DC, Eden JA. 1996. A review of the clinical effects of phytoestrogens. *Obstet Gynecol* **87**: 897–904.

Kolonel LN, Hankin JH, Whittemore AS, Wu AH, Gallagher RP, Wilkens LR. 2000. Vegetables, fruits, legumes and prostate cancer: a multiethnic case–control study. *Cancer Epidemiol Biomarkers Prev* **9**: 795–804.

Lenz EM, Bright J, Wilson ID *et al*. 2004. Metobonomics, dietary influences and cultural differences: a 1H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J Pharm Biomed Anal* **36**: 841–849.

Lindon JC, Holmes E, Nicholson JK. 2004. Metabonomics and its role in drug development and disease diagnosis. *Expert Rev Mol Diagn* **4**: 189–199.

Lydeking-Olsen E, Beck-Jensen JE, Setchell KDR, Holm-Jensen T. 2004. Soymilk or progesterone for the prevention of bone loss: A 2 year randomized, placebo-controlled trial. *Eur J Nutr* **43**: 246–257.

Muller M, Kersten S. 2003. Nutrigenomics: goals and strategies. *Nature Rev Genet* **4**: 315–322.

Nicholson JK, Lindon JC, Holmes E. 1999. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **29**: 1181–1189.

Park S-Y, Wilkens LR, Franke AA *et al*. 2009. Urinary phytoestrogen excretion and prostate cancer risk: a nested case–control study in the Multiethnic Cohort. *Br J Cancer* **101**: 185–191.

Rezzi S, Ramadan Z, Fay LB, Kochhar S. 2007. Nutritional metabonomics: applications and perspectives. *J Proteome Res* **6**: 513–525.

Rice S, Whitehead SA. 2006. Phytoestrogens and breast cancer – promoters or protectors? *Endocr Relat Cancer* **13**: 995–1015.

Roudsari AH, Tahbaz F, Hossein-Nezhad A, Arjmandi B, Larijani B, Kimiagar SM. 2005. Assessment of soy phytoestrogens' effects on bone turnover indicators in menopausal women with osteopenia in Iran: a before and after clinical trial. *Nutr J* **4**: 30.

Schabath MB, Hernandez LM, Wu XF, Pillow PC, Spitz MR. 2005. Dietary phytoestrogens and lung cancer risk. *J Am Med Assoc* **294**: 1493–1504.

Shu XO, Jin F, Dai Q et al. 2001. Soyfood intake during adolescence and subsequent risk of breast cancer among Chinese women. *Cancer Epidemiol Biomarkers Prev* **10**: 483–488.

Solanky KS, Bailey NJC, Beckwith-Hall BM *et al*. 2003. Application of biofluid H-1 nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile. *Anal Biochem* **323**: 197–204.

Solanky KS, Bailey NJ, Beckwith-Hall BM *et al*. 2005. Biofluid H-1 NMR-based metabonomic techniques in nutrition research metabolic effects of dietary isoflavones in humans. *J Nutr Biochem* **16**: 236–244.

Somekawa Y, Chiguchi M, Ishibashi T, Aso T. 2001. Soy intake related to menopausal symptoms, serum lipids, and bone mineral density in postmenopausal Japanese women. *Obstet Gynecol* **97**: 109–115.

Vanharanta M, Voutilainen S, Lakka TA, van der Lee M, Adlercreutz H, Salonen JT. 1999. Risk of acute coronary events according to serum concentrations of enterolactone: a prospective population-based case–control study. *Lancet* **354**: 2112–2115.

Weljie AM, Newton J, Mercier P, Carlson E, Slupsky CM. 2006. Targeted profiling: Quantitative analysis of H-1 NMR metabolomics data. *Anal Chem* **78**: 4430–4442.

Wiback SJ, Palsson BO. 2002. Extreme pathway analysis of human red blood cell metabolism. *Biophys J* **83**: 808–818.