Development of nanoelectrospray high resolution isotope dilution mass spectrometry for targeted quantitative analysis of urinary metabolites: Application to population profiling and clinical studies

*Elena Chekmeneva [a,b], Gonçalo Correia [a], Júlia Dénes [a], María Gómez-Romero [a], Anisha Wijeyesekera [a], Dora R. Perenyi [a], Yvonne Koot [d], Carolien Boomsma [d], Elisabeth. J. Want [a], Peter H. Dixon [f], Nicholas S. Macklon [d,e], Queenie Chan [b], Zoltán Takáts [a,c], Jeremy K. Nicholson [a,c] and Elaine Holmes [a,c]**

[a] Biomolecular Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, South Kensington, London SW7 2AZ, UK

[b] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG, UK,

[c] MRC-NIHR National Phenome Centre, London SW7 2AZ, UK

[d] Utrecht University, The Netherlands

[e] University of Southampton, UK

[f] Institute of Reproductive and Developmental Biology, Imperial College London, UK

*** To whom correspondence should be addressed, *e-mail*:**
elaine.holmes@imperial.ac.uk

ABSTRACT

An automated chip-based electrospray platform was used to develop a high-throughput nanoelectrospray high resolution mass spectrometry (nESI-HRMS) method for multiplexed parallel untargeted and targeted quantitative metabolic analysis of the urine samples. The method was demonstrated to be suitable for metabolic analysis of large sample numbers and can be applied to large-scale epidemiological and stratified medicine studies. The method requires a small amount of sample (5 μL of injectable volume containing 250 nL of original sample), and the analysis time for each sample is three minutes per sample to acquire data in both negative and positive ion modes. Identification of metabolites was based on the high resolution accurate mass and tandem mass spectrometry using authentic standards.

The method was validated for 8 targeted metabolites and was shown to be precise and accurate. The mean accuracy of individual measurements being of 106% and the intra- and inter-day precision (expressed as relative standard deviations) were 9% and 14%, respectively. Selected metabolites were quantified by standard addition calibration using the stable isotope labelled internal standards in a pooled urine sample, to account for any matrix effect. The multiple point standard addition calibration curves yielded correlation coefficients greater than 0.99, and the linear dynamic range was more than three orders of magnitude. As a proof-of-concept the developed method was applied for targeted quantitative analysis of a set of 101 urine samples obtained from female participants with different pregnancy outcomes. In addition to the specifically targeted metabolites, several other metabolites were quantified relative to the internal standards. Based on the calculated concentrations, some metabolites showed significant differences according to different pregnancy outcomes. The acquired high resolution full-scan data were used for further untargeted fingerprinting and improved the differentiation of urine samples based on pregnancy outcome.

INTRODUCTION

Large scale metabolic phenotyping and metabolome-wide association studies are increasingly being applied to analysis of large scale sample cohorts in epidemiological and clinical settings. [1, 2]

Metabolic profiling analysis of different biological samples (biofluids and tissues) presents multiple analytical challenges due to the large variation and huge chemical diversity of their

constituents and significant influence of several external factors such as diet and drugs. In order to maximize the coverage of diverse structural entities over several orders of magnitude of concentration range, several analytical platforms are used nowadays, most notably liquid and solid ^{1}H NMR,[3, 4] LC- and GC-MS (MS/MS).[5-8] The concept of metabolic profiles was introduced by Horning and Horning in 70s[9, 10] who used GC-MS for qualitative and quantitative analyses of steroids, sugars and sugar alcohols, acids, Krebs cycle intermediates and and drug metabolites in blood and/or urine. Modern instrumentation advances allow analysis of tens to hundreds of endogenous or drug metabolites simultaneously using GC-MS or LC-MS/MS in many laboratories. Recently, an extensive quantitative multi-platform study of the urine metabolome was published[11] providing a database containing numerous identified metabolites and their corresponding normal and disease associated concentration ranges.

Among the available analytical techniques, mass spectrometry offers an attractive combination of high sensitivity and selectivity of analysis, along with the possibility of structural elucidation for metabolite identification and quantification at low levels of concentration.[12-15] At present, metabonomic studies benefit from various available MS technologies including different ionization techniques and mass analyzers with different resolving power and mass accuracy. LC separation followed by triple quadrupole-based multiple reaction monitoring (MRM) is the leading choice for targeted metabolite quantification.[14] This approach has been considered as a "gold standard" for absolute quantification of small molecules due to its sensitivity and specificity. However, the information provided by this approach is limited to the targeted metabolites and the other metabolites in the biofluid remain undetected. This limitation can be overcome by using mass spectrometry in full-scan mode. However, in order to compensate for the loss of specificity and to be able to differentiate between species with identical nominal mass, use of high resolution mass analysers (Fourier-transform- ion cyclotron resonance (FT-ICR) or advanced time-of-flight instruments) is desirable. In some cases it has been shown that MS/MS quantitation can be more sensitive than HRMS, if only a few compounds are monitored.[10] An important advantage of using mass spectrometry in full scan mode is that the acquired full-scan data contain information not only about the targeted compounds, but also other components of the samples undergoing ionization; these data can be queried any time post-acquisition for the presence or absence of any well-defined molecular species. The

use of HRMS instruments in full scan mode for drug, biological fluids and environmental analysis have also been reported widely.[16-22]

Sustained analysis of epidemiological scale sample numbers (typically thousands) requires tools capable of high-throughput analysis such as direct infusion mass spectrometry (DIMS).[23-25] In contrast, the use of hyphenated techniques such as LC-MS (MS/MS) can be time consuming due to instrument conditioning and maintenance required on a daily basis during LC experiments, particularly where analysis is performed separately in positive and negative ion modes. The associated cost of analysis can be extremely high for large sample numbers due to the increased solvent requirements. Problems of reproducibility and system stability (especially regarding chromatographic columns, LC pump performance and subtle variations in mobile phase composition) may also arise during the long-term analyses as well as the problem of detector saturation and source contamination with components of mobile phase. For global metabolic profiling, reversed-phase (RP) UPLC gradients of 10-12 min (MS acquisition in one ionization mode) allow good peak capacity with moderate throughput.[7] Faster separations are possible with shorter columns, narrower diameters, alternative types of column material (such as monolithic and fused-core), higher temperatures and higher flow rates. However, the latter may lead to poorer ionization while higher temperature may result in metabolite stability problems. The use of ultrafast chromatography and ballistic gradients with very short cycle time has been reported.[26] In the case of targeted analysis, the incorporation of fast polarity switching for data acquisition in two ionization modes simultaneously in combination with ultrafast chromatography is possible using advanced triple quadrupole mass spectrometers. However, its application for global profiling will still require separate acquisition in positive and negative ion modes (for high resolution data) and will not preclude some of the LC related problems. Thus, the use of DIMS for large-scale epidemiological studies could be an optimal means of accelerating MS analysis whilst keeping associated costs affordable.

DIMS can be employed by using a flow injection analysis mass spectrometry (FIA-MS) system or a fully automated chip-based nanoelectrospray (nESI) approach. The chip consists of an array of 400 nanoelectrospray emitter nozzles providing a stable spray for up to 30 min for a few microliters of sample. The NanoMate system has several potential advantages over LC-MS comprising high-throughput, lack of sample-to-sample carry-over, low sample amounts required for the analysis (nanoliters in comparison with microliters required for LC-MS or FIA-MS), wide dynamic range and a sensitivity comparable to that

of hyphenated techniques. Another important advantage of very small electrospray emitter diameters (< 10 µm) is the reduction of ion suppression effects.[27-29] The main disadvantage of nESI-HRMS (and all other DIMS methods) is the inability of the technique to distinguish between isobaric and isomeric species in single stage MS mode, which requires chromatographic separation or additional MS/MS experiments. The main characteristics of the nESI-DIMS method together with a comparison to chromatographic approaches for large-scale applications are presented in Table 1. **Table 1** Summary of main advantages of nESI-HRMS method in comparison with LC-MS approaches

Characteristic of nESI-DIMS method	**LC-MS comparison (alternative)**
Fast analysis in both ionization modes. Very high-throughput.	Use of fast chromatography with fast polarity switching for targeted analysis in MRM mode.
No cross contamination. Lack of carryover.	Possibility of carryover between the samples. Need of injection of blanks and needle washing cycles.
Low cost of analysis (*ca.* 400 *per* chip)	Increased solvent requirements. Need to change a chromatographic column due to deterioration of its performance in long-term use.
Low sample amount (nanoliters)	Microliters of sample
Low ion suppression with low flow rates (nL/min) and sample dilution.	Decreased ion suppression due to the chromatographic separation.
Coverage of a wide range of metabolites (information about all ionized components in a sample): polar metabolites, fatty acids, lipids, acylcarnitines, *etc*.	Need of using different stationary phases for metabolites of different polarities (HILIC and RP) and lipids.

The nESI method has been successfully applied for quantification of serotonin and related compounds in urine samples,[30] free fatty acids in serum using the exact mass and isotopic distribution,[31] as well as for the global profiling of plasma.[28] Application of nESI methods and computational workflows for global metabolic profiling of various tissue extracts using the FT-ICR mass spectrometer has been significantly developed and optimized over the last few years.[32-34] However, literature relating to urine metabolic profiling studies employing nESI-MS, or in general, DIMS methods, is limited.[24] Recently, different urine sample preparation procedures (including solid phase and liquid-liquid extraction) along with different data normalization approaches were evaluated using the DI-ESI-MS.[35] The optimized approach based on simple sample dilution and statistical data normalization has been shown to have the highest potential.

In the current study we present the development of a nESI-HRMS method for targeted multi-analyte quantitative analysis and parallel global metabolic profiling of urine samples for the prospective application in large-scale epidemiological studies. The general workflow of the proposed method is depicted in the Supplementary Figure S1. The method was developed on the TriVersa Advion NanoMate system in infusion mode coupled to HRMS QTOF Synapt G2-*Si* (Waters Corporation, Manchester, UK). The method uses the stable isotope labelled analogues of all the targeted metabolites as internal standards. These can also be used as internal references for the m/z scale to correct for any shifts and to verify mass accuracy. The method has been validated according to the relevant FDA guide lines.[36]

As a proof-of-concept the method was applied for the analysis of a set of 101 urine samples from female patients with different pregnancy outcomes (Pregnant, Non-Pregnant and Early Pregnancy Loss (EPL)). The quantification was achieved by standard addition of authentic standards into a composite (pooled) urine sample accounting for any matrix effects. The identification of the peaks of interest was based on the accurate mass measurements, spiking experiments (for both, non-labelled and labelled standards) and off-line MS/MS measurements for the targeted masses. Additionally the full-scan MS data were subjected to multivariate analysis for metabolic fingerprinting.

EXPERIMENTAL

Materials

All chemicals and solvents were of the highest purity grade. Labelled and non-labelled standards used for quantification (See Table 2 for details) were purchased from Sigma-Aldrich (Isotec, Gillingham, UK) or CK Isotope (Cambridge Isotope Laboratories, Ibstock, UK). HPLC-grade methanol and water were purchased from Sigma-Aldrich (Gillingham, UK).

Preparation of standard solutions

The labelled and non-labelled standard stock solutions were prepared at concentration of 1 mg/mL in methanol. The multi-analyte mixture of labelled internals standards and the multi-analyte mixture of non-labelled calibrators were prepared by mixing 1 mL of each stock solution in the total volume of 10 mL. Working solutions used for the calibration curve were prepared from stock solutions at concentrations of 100, 10, 5, 1, 0.5, 0.05, and 0.05 μg/mL.

The labelled internal standard multi-analyte mixture was prepared at 10 μg/mL for all the standards by serial dilution of the stock solutions with methanol.

Samples

A total of 101 first early morning human urine samples from 19 female patients aged 27-32 with previous difficulties conceiving were obtained on the days 18, 22, and 26 during three consecutive cycles and stored at -80° C. The patients had different outcomes: Pregnant (P), Non-Pregnant (NP) and Early Pregnancy Loss (EPL). See Supplementary Table S1 for details.

Ethical approval for the study was obtained from the Utrecht Medical Centre research ethics committee and written informed consent was obtained from all participants.

Sample Preparation

Thawed urine samples were diluted by a factor of 20. An aliquot of 10 μL of each sample were mixed in order to obtain the pooled urine sample used for the preparation of calibration series and quality control (QC) samples. The pooled urine sample was also diluted by a factor of 20.

An aliquot of 100 µL of each diluted sample was pipetted in a randomized order into the well-plates. To each well containing 100 µL of a test sample, 50 µL of the multi-analyte mixture of labelled internal standards at fixed concentration were added for quantification. Then, 150 µL of ultrapure methanol were added in order to maintain the water-methanol proportion of 1:2 which provided the most stable signal in the optimization experiments.

The calibration series and the QC samples were prepared in a similar way pipetting 100 µL of the pooled urine samples, adding 50 µL of the multi-analyte mixture of labelled internal standards at fixed concentration and 50 µL of the corresponding calibrator, and making up the total volume to 300 µL by adding methanol.

In this way, the urine samples were diluted by a factor of 60 in a well, while the calibrators used for the calibration series and QC samples were diluted 6 times. All these factors have been taken into account for the calculations of the limits of detection (LOD) and quantification (LOQ) as well as the quantification results.

The sample plates were sealed and subjected to 1 min ultrasonication followed by 10 min centrifugation at 1500g and 4 °C before MS analysis.

Chip based nanoelectrospray system

Chip-based electrospray infusion analysis was performed using the TriVersa NanoMate system (Advion BioSciences, Ithaca, New York). This device uses 5 µm nominal diameter nozzles on disposable chips and a new tip from a 384 tip rack for each sample from the 96-well plate. The injection volume of 5 µL of a sample was aspirated and delivered to the inlet side of the ESI chip. The nanoelectrospray was created and maintained by applying 1.4 kV high voltage and 0.8 psi nitrogen flow. The sample plate temperature was maintained at 4 °C.

Total data acquisition time was of 50 seconds for each ionization mode (first negative, and then positive ion mode) but the overall turnaround time for each sample was 2.25 minutes to let the instrument automatically switch the polarity and the voltage settle before acquiring the data in the second ion mode. The data for the negative and positive mode were acquired in two separate files in the MassLynx™ software. The total time for the analysis of a 96 well-plate was 5 hours.

Mass Spectrometry

High-resolution TOF MS data were acquired on a Waters Synapt® G2-Si (Waters MS Technologies, UK) in negative and positive modes with automatic polarity switch. The data were collected in high resolution (R= 35000) continuum mode with the scan time of 1 s over the mass range of 40 – 600 m/z. The sampling cone voltage was set at 40 V, and the source offset at 80 V. The temperature of the source was set to 100 °C.

Sodium formate solution was used to calibrate the mass spectrometer on a daily basis. The lock-mass function was turned off but data in both modes were recalibrated after the acquisition by in-house software using reference signals of endogenous metabolites present in all urine samples (such as creatinine in positive ion mode and hippuric acid in negative ion mode) as well as of the labelled standards added to each sample.

For the identification of compounds, MS/MS was performed on targeted peaks in Resolution mode (with Enhanced Duty Cycle) using a pooled urine sample. Since the urine samples were diluted by a factor of 20 and then by a factor of 3 in a well (total dilution factor of 60, see Sample Preparation section), the acquisition time for the tandem mass spectrometric measurements was extended up to 10 minutes for some low concentration metabolites. The optimal CID energy was selected for each peak between 10-60 eV. The MS/MS experiments were repeated for the pooled urine sample spiked with the standards. The two sets of spectra were compared to the metabolite fragmentation patterns available in online databases (HMDB[11] and Metlin[37]) and spectra acquired for the standards in neat solvent.

Quantification

Each well-plate contained the test samples, the calibration series (consisting of six points plus zero point), six QC samples prepared at three different concentration levels. The blanks consisting of water and methanol were injected after the calibration series and each QC sample in order to avoid any carry-over of metabolites or standards.

A standard addition into a pooled urine sample (which should reflect the averaged nature of all the samples from a study) was used by spiking the multi-analyte series of non-labelled standards and using isotope-labelled analogues as internal standards spiking them at a known and fixed concentration into each calibrator, QC and test sample. The MS quantification was performed using the ratio of intensities of the metabolite to the intensity

of its internal standard. The response for each test sample was compared to the generated calibration curve to give urinary concentration values.

In the case of relative quantification, normalized metabolite signal intensities were compared between different groups, using either an internal standard or another metabolite as a basis for normalization. The internal standard for each metabolite was selected analysing the calibration series and QC samples, in which their concentration was constant. The ratio of intensity of each of them to different available internal standards was assessed in these samples. The standard giving the lowest variance was selected as an internal standard for normalization of each metabolite.

The accuracy and intra- and inter-day precision during the method validation were calculated as follows:

Accuracy = Mean (measured concentration)/ Theoretical concentration ^{x} 100

Precision = Mean SD/ mean (measured concentration) ^{x} 100.

Method LOQ was determined as 10 times the standard deviation of the y-intercept divided by the slope of the calibration curves obtained in water.

Data Analysis

Basic data visualization and quality control was achieved using the MassLynx 4.1 software (Waters Corporation, UK). For analysis, the raw data were converted to the mzML format with ms convert followed by processing using in-house scripts written in python 2.7.4[33, 34, 37]. These included the sum of the scans acquired, correction of the mass shifts using the observed masses for reference compounds, library of internal reference masses, interpolation of the data to a final resolution of 0.00056 m/z, local baseline estimation and peak picking using the Savitzky-Golay algorithm.[39] For the untargeted analysis, a matrix containing all the spectra in one ion mode interpolated on a common m/z scale was used.

Construction of the calibration curves and quantification result calculation was performed in Microsoft Excel. The obtained concentrations and relative quantification results were normalized for creatinine in each individual sample, log transformed and subjected to one-way ANOVA, Kruskal-Wallis or t-test using R 3.0.3[9]. For the global metabolic profiling, the full scan data were probabilistic quotient normalized[40] to reduce any variance between the samples due to any possible difference in their extent of dilution. Principal Component

Analysis (PCA) and Partial Least Square - Discriminant Analysis (PLS-DA) were performed in MATLAB R2014a (Mathworks^{TM}), and SIMCA 13.0.2 (Umetrics, Sweden) to establish inherent groupings in urine samples relating to biological class[41], in this case pregnancy outcome (pregnant, non-pregnant or early pregnancy loss).

RESULTS AND DISCUSSION

nESI-HRMS analysis of the human urine samples resulted in spectra featuring several hundreds of signals from different metabolic constituents in both negative and positive ion modes. Parameters of the instruments (TriVersa NanoMate and Synapt-G2 Si) were optimized to achieve the best mass resolution, spray stability and sensitivity. An example of full-scan data is presented in the Supplementary Figure S2.

The list of stable isotope standards and non-labelled calibrators along with the information about the MS detection mode and the m/z values are presented in Table 2. Since the method allows for the acquisition of full-scan data, and information about all metabolites undergoing ES ionization in a sample is potentially available, relative quantification was possible for an extended panel using the internal standards already present in the sample. Thus, several other metabolites have been included in the list presented in the Table 2 for relative quantification. The selected compounds shown in the Table 2 cover different classes of metabolic constituents including amino acids, TCA cycle metabolites, fatty acids, acylcarnitines, gut microbial metabolites, and metabolites of the NAD+ pathway. Many of these metabolites have been previously reported in several metabolic profiling studies related to pregnancy disorders.[42-44]

Method Validation

As a first step in the method development, different dilution factors (non-diluted (1), 1:5, 1:20, 1:50, and 1:100) of the urine samples were tested along with the sample filtration using the 96-well filter plates. The filtration of the diluted urine samples before the analysis showed neither significant difference in the quality of the data acquired nor in the number of detected signals or instrument performance. Thus, only sample dilution was used as a sample preparation step. The optimal dilution factor was characterized in depth assessing effect of dilution on the intensities of endogenous metabolites along with the intensities of added internal standards. The effect of dilution on the obtained spectra of the pooled urine

sample is shown in Supplementary Figure S3 (a). It was observed that sample dilution resulted in an increase in the number of detected components. This can be explained by the improved ionization efficiency and diminished matrix suppression effect caused by the presence of inorganic salts. It should be noted that the analysis of undiluted urine samples did not allow the detection of any of the targeted metabolites. It was found that the optimal dilution factor for the urine samples was between 1:20 and 1:50. The intensities of the added internal standards increased with the dilution of the urine samples and were at their highest point at 1:100 dilution. The endogenous metabolite intensities decreased from 1:20 to 1:100 (see examples on the Supplemental Figure S3 (b)). In order to minimize the loss of information (especially regarding the trace level metabolites) a dilution factor of 1:20 was selected for the analysis of these samples. However, it must be pointed out that the loss of intensity of endogenous metabolites at 1:50 dilution is smaller than the increase in the intensities of the internal standards, due to increased ionization efficiency. Higher dilutions did not produce significantly fewer ions, while lower dilutions did not cause an increase in the intensities of endogenous compounds as expected. A simple calculation shows that this method requires a very small amount of sample. For example, if the volume of infusion is of 5 μL, and the sample is diluted 1:20 times followed by 3 fold dilution with methanol in a well-plate (see Experimental section, Sample preparation), the amount of original sample that is actually injected into the mass spectrometer is of several nL. In comparison, in the LC-MS measurements, typically human urine samples are either not diluted or diluted 1:1 with water, and the injection volumes are in the range of of 1-5 μL. Such a small sample amount injected during nES ionization results in the absence of any carry-over and contamination. An essential part of any method development is its validation to reassure its robustness and reliability. Additional experiments analysing multi-analyte QC samples according to US FDA guidelines[36] were carried out to validate the nESI-HRMS method. The analytical selectivity of the method was confirmed by the ability to detect targeted signals and quantify the metabolites of interest in the presence of many other compounds in the urinary specimens.

The most challenging step in the development of any multiplexed assays of urine is to determine individually the concentration levels for each of the targeted metabolites as the composition of the matrix and the concentration range of the analytes can be highly variable. For each metabolite of interest, the linear range (LLOQ and ULOQ as the lower and upper extremes of the range), accuracy, precision were determined from the replicates

of calibration curves and QC analysis (inter- and intra-day) prepared using the pooled urine sample. Calibration curves were also recorded for purely aqueous matrix in order to examine the matrix effect and to obtain the values of the LOQ using the neat solvent as the blank matrix. The results are presented in the Table 3 and in Supplementary Tables S2 and S3.

Comparison of the calibration curves obtained for pooled urine and water indicated that the linear range in water is relatively narrow. The quantification performance was found to be better in the low concentration range since the curves quickly saturate at higher concentrations (for the majority of metabolites starting from 0.167 μg/mL, see Supplementary Table S2). When the standard addition method is used in pooled urine samples, the linear range extends to higher concentrations, but it is limited at lower values by the baseline concentration of metabolite already present in the urine being the main drawback of standard addition method (see Table 3). The slopes of the curves obtained for urine and water are practically identical. This observation implies that it is possible to extrapolate the results for the test samples to values lower than the baseline concentration in the pooled urine sample. Using the slopes obtained by the standard addition approach to the pooled urine sample represents the only strategy which accounts for matrix effects for the calibration series and the sample and allows extending linear range to higher concentrations.

The LOQ values obtained for water as a matrix were all within the reported normal ranges reported in literature and presented in the Supplementary Table S2.

The determination of accuracy and precision values are reported in Supplementary Table S3. This analysis helped to refine and confirm the linear ranges for each metabolite. A potential source of inaccuracy for the quantification of some metabolites (CV higher than 20%) in this study was associated with the initial use of the same level of calibrator concentration for all the metabolites selected for quantification. An improvement in efficiency was achieved through an iterative process of finding the optimal linear ranges for each targeted metabolite, confirming the selectivity of the method and finding the right concentration of the labelled internal standard (5S/N) for a reliable quantification and identification. It has been previously discussed and reported[45] that the validation assay should be developed and adapted to an intended purpose of a study; it must be "fit-for-purpose". Different levels of assays have been proposed, each of them obeying a certain level of rigor of validation. The multiplexed studies, targeting tens or hundreds of

compounds in a single analysis, present a particular challenge, especially in large scale population studies with a high degree of within- and between-individual variability of urinary metabolites. For biomarker studies, FDA guidelines are slightly more "relaxed"[46] (20 and 30% rule for accuracy and precision of the QC samples at high and LLOQ levels, respectively) which is more appropriate for metabonomics.

Table 2 Metabolites included in the list for quantification by standard addition method and relative quantification.

Metabolite	IS standard	nESI ion mode	m/z [M-H]-	m/z [M+H]+
Adipic acid*	Adipic-d8-acid	Negative	145.0501	
Leucine*	Leucine-d3	Negative	130.0868	
Phenylalanine*	Phenylalanine-β-^{13}C	Negative	164.0712	
Glutaric acid*	Glutaric acid (2,2,4,4-d4)	Negative	131.0344	
Succinic acid*	Succinic-d4-acid	Negative	117.0188	
Palmitic acid*	Palmitic acid-1-^{13}C	Negative	255.2324	
Acetylcarnitine*	Acetyl-d3-L-carnitine	Positive		204.1236
Nicotinamide*	Nicotinamide-2,4,5,6-d4	Positive		123.0558
4-cresol	Succinic-d4-acid	Negative	107.0502	
Aconitic acid	Succinic-d4-acid	Negative	173.0086	
Asparagine	Phenylalanine-β-^{13}C	Negative	131.0457	
Benzoic acid	Phenylalanine-β-^{13}C	Negative	121.0295	
Citric acid	Succinic-d4-acid	Negative	191.0197	
Glutamic acid	Succinic-d4-acid	Negative	146.0459	
Glutamine	Phenylalanine-β-^{13}C	Negative	145.0613	
Hexoses	Phenylalanine-β-^{13}C	Negative	179.0561	
Hippuric acid	Phenylalanine-β-^{13}C	Negative	178.051	
Indoxyl sulfate	Phenylalanine-β-^{13}C	Negative	212.0023	
N-acetylneuraminic acid	Phenylalanine-β-^{13}C	Negative	308.0982	
Nicotinic acid	Phenylalanine-β-^{13}C	Negative	122.0248	
Phenylacetylglutamine	Phenylalanine-β-^{13}C	Negative	263.1037	

Taurine	Phenylalanine-β-^{13}C	Negative	124.0068	
Tyrosine	Phenylalanine-β-^{13}C	Negative	180.0661	
L-Carnitine	Acetyl-d3-L-carnitine	Positive		162.1125
Creatinine	Acetyl-d3-L-carnitine	Positive		114.0662
Creatine	Acetyl-d3-L-carnitine	Positive		132.0773
Dopamine	Acetyl-d3-L-carnitine	Positive		154.0868
Hypoxanthine	Nicotinamide-2,4,5,6-d4	Positive		137.0463
Phenylacetic acid	Acetyl-d3-L-carnitine	Positive		137.0603
Phenylethylamine	Acetyl-d3-L-carnitine	Positive		122.0964
Propionylcarnitine	Acetyl-d3-L-carnitine	Positive		218.1392
Tyramine	Acetyl-d3-L-carnitine	Positive		138.0892

* Quantification by standard addition (absolute)

Table 3 Calibration curves **ax+b=y** (average of 7 replicates), linear range and the concentration of the metabolites obtained in the pooled urine matrix and compared to the literature values.

Metabolite	**a**	**b**	**R^2**	**Linear range (µg/mL) expressed as the concentration of a calibrator in a well.**
Succinic acid	0.0639	0.5187	0.9399	0.833-16.67
D,L-Leucine	0.1306	0.1911	0.9995	0.833-16.67
Glutaric acid	0.0560	0.3067	0.9735	0.833-16.67
Adipic acid	1.2416	0.4317	0.9994	0.167-16.67
D,L-Phenylalanine	0.0729	0.2650	0.9899	0.167-1.667
Palmitic acid	0.0761	0.1728	0.9712	0.167-0.833
Nicotinamide	0.0653	0.0413	0.9968	0.083-16.67
Acetylcarnitine	0.0744	0.0348	0.9986	0.083-0.833

Metabolite quantification in the urine samples by a nESI-HRMS method

The selection of metabolites for quantification was based on the previously reported data linking them to various pregnancy disorders.[42-44]

The metabolites shown in Table 2 (with an asterisk) were quantified by the standard addition method described above using the calibration curve obtained on each well-plate in the pooled urine sample. The peaks of the metabolites of interest were identified using the accurate masses of their ions, tandem mass spectrometry and confirmed by accurate mass and MS/MS patterns of authentic standards (an example spectrum I provided in Figure S4). For certain metabolites, some of the samples produced higher than the reported normal range concentration values, in a patient dependent manner (note that there were several samples obtained from the same patient in different cycle days, to a maximum of 9 samples per patient). As mentioned above, the main drawback of the developed method is its inability to selectively differentiate structural isomers. For instance, one of the structural isomers of succinic acid is methylmalonic acid, thus the quantitation of the former can suffer from interference from the latter. Their selective analysis would require complicated derivatization and chromatographic separation procedures.[47] However, succinic acid is the more abundant metabolite, and elevated concentration of methylmalonic acid is expected only in some rare cases related to genetically inherited disease - methylmalonic acidaemia for example- in which the body is unable to process correctly certain amino acids and fats. Also, it can be related to severe vitamin B12 deficiency, but the incidence of this condition is also rare. Therefore, the values obtained based on the signal at 117.0188 are more likely to reflect the concentrations of succinic acid.

In the context of hypothesis generating metabolic profiling studies, the limitation related to resolution of isomeric species, in some cases, maybe only apparent because many isomers are often part of the same metabolic pathway, and in this case, it is possible to estimate the sum effect. This could be true in the case of sugars (glucose, fructose, mannose, galactose) or branched chain amino acids (leucine and isoleucine). However, in hypothesis testing studies, for each target metabolite the information of its chemical structure and possibility to have structural isomers must be first assessed. For example, dimethylglycine and gamma-aminobutyric acid are structural isomers but they belong to different biochemical pathways and have very different functions in the body. Thus, an attempt of their analysis by DIMS may result uninformative.

The method presented here also enabled the relative quantification of non-targeted species. All the metabolites listed in Table 2 were quantified in all urine samples. Supplementary Table S4 shows the values in μg/mL for the metabolites quantified using the standard addition method. These values were normalized to the creatinine signal quantified relatively in each sample using acetyl-d3-L-carnitine as an internal standard (see Table 2) present in each sample well at fixed concentration. The choice of internal standards for relative quantification was based on the coefficient of variance of the ratio of intensity of each metabolite to different available internal standards in the QC samples and calibration series. Normalization to creatinine was also performed for other metabolites listed in the Table 2, and these data were subsequently used for statistical analysis. The log transformed values were subjected to one-way ANOVA analysis, $p<0.05$ (in the case if the Shapiro-Wilk test for normality failed, the non-parametric Kruskal-Wallis test was applied instead of ANOVA). Table 4 presents the p-values for different sets of data tested. The differences can be also visualized in the boxplots in Figures 1 and 2.

In order to demonstrate the validity of presented method, several metabolites (succinic acid, phenylalanine, acetylcarnitine and nicotinamide) were quantified in a small number of samples by both a nESI-HRMS and a nESI-MS/MS method developed on the same mass spectrometer in a similar way to that reported earlier in the literature.[23] The method comparison experiments included a calibration series obtained in a pooled urine sample and the analysis of 15 samples (5 from each of three groups of patients, see Supplementary Table 2). The calibration and quantification results for the two methods were compared using Pearson correlation. The correlation coefficients for the calculated calibration series varied in the range of 0.92-0.98 for four metabolites. In case of quantification, the correlation coefficients were as follows: succinic acid 0.81; phenylalanine 0.91; acetylcarnitine 0.79, and nicotinamide 0.94. The nESI-HRMS method tended to give slightly lower concentration values compared to the MS/MS method. This is in agreement with the results reported for the comparison of the HRMS method developed on an Orbitrap mass spectrometer and a traditional ESI-MS/MS method.[23] However, the slight shift of the quantification values obtained by nESI-HRMS method compared to MS/MS quantification did not compromise the diagnostic ranges and discrimination of the study groups.

Table 4 One-way ANOVA (or Kruskal-Wallis) test comparing three pregnancy outcomes with t-test values for pairwise comparison of the outcomes for the quantified metabolites, *** p<0.05**

Metabolite	**3 outcomes (one-way ANOVA)**	**NP vs P (t-test)**	**NP vs EPL (t-test)**	**P vs EPL (t-test)**
Glutaric acid	0.145	0.225	0.479	0.0674
Succinic acid	*** 0.0122**	*** 0.0412**	0.539	*** 0.00207**
Palmitic acid	0.288	0.231	0.735	0.163
Adipic Acid	0.093	0.15	0.515	0.0516
Leucine	0.139	0.263	0.392	0.0614
Phenylalanyne	0.096	*** 0.039**	0.534	0.132
Acetylcarnitine	0.88	0.711	0.65	0.894
Nicotinamide	*** 0.0433**	*** 0.0261**	0.65	0.0703
4-Cresol	0.513	0.47	0.196	0.724
Aconitic acid	*** 0.021**	*** 0.023**	0.928	*** 0.012**
Asparagine	0.415	0.227	0.69	0.69
Benzoic acid	*** 0.005**	0.086	0.112	*** 0.001**
Citric acid	*** 0.025**	0.625	0.077	*** 0.004**
Glutamic acid	*** 0.008**	*** 0.011**	0.877	*** 0.006**
Glutamine	*** 0.005**	0.071	0.192	*** 0.001**
Hexoses	0.129	0.966	0.059	0.08
Hippuric acid	0.461	0.8	0.203	0.33
Indoxyl sulfate	*** 0.002**	*** 0.015**	0.485	*** 0.0005**
N-acetylneuraminic acid	*** 0.043**	0.188	0.274	**0.01**
Nicotinic acid	0.261	0.327	0.598	0.1
Phenylacetylglutamine	*** 0.006**	*** 0.04**	0.402	*** 0.0015**
Taurine	0.931	0.676	0.876	0.841
Tyrosine	0.845	0.871	0.503	0.705
Carnitine	0.119	0.816	0.108	0.0695
Creatine	0.668	0.332	0.84	0.628
Hypoxanthine	0.6264	0.4672	0.84	0.368

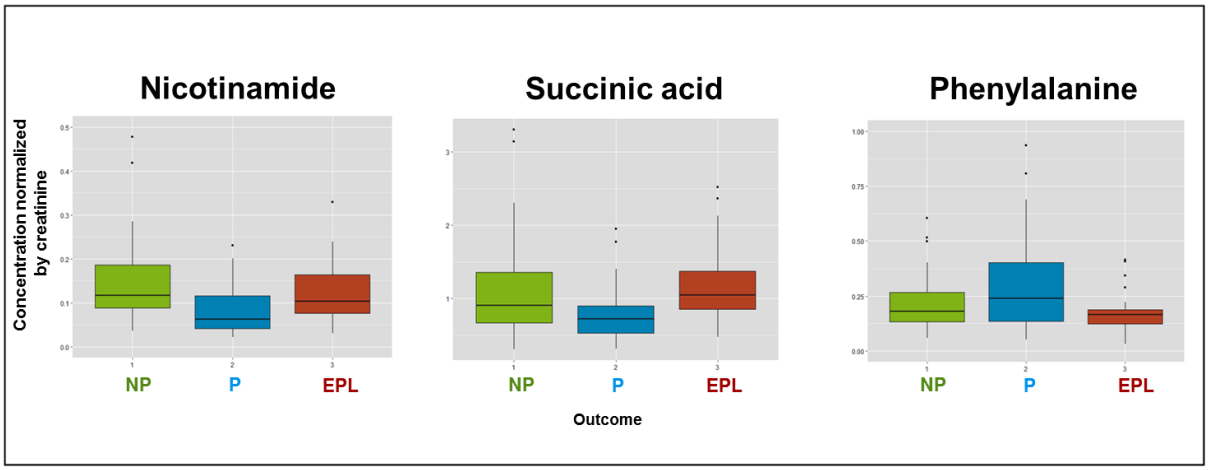

Fig.1 Box plots for three metabolites quantified by standard addition method showing significant differences between P and NP or EPL groups; one-way ANOVA p<0.05 (NP-Non-Pregnant, P-Pregnant, and EPL – early pregnancy loss).

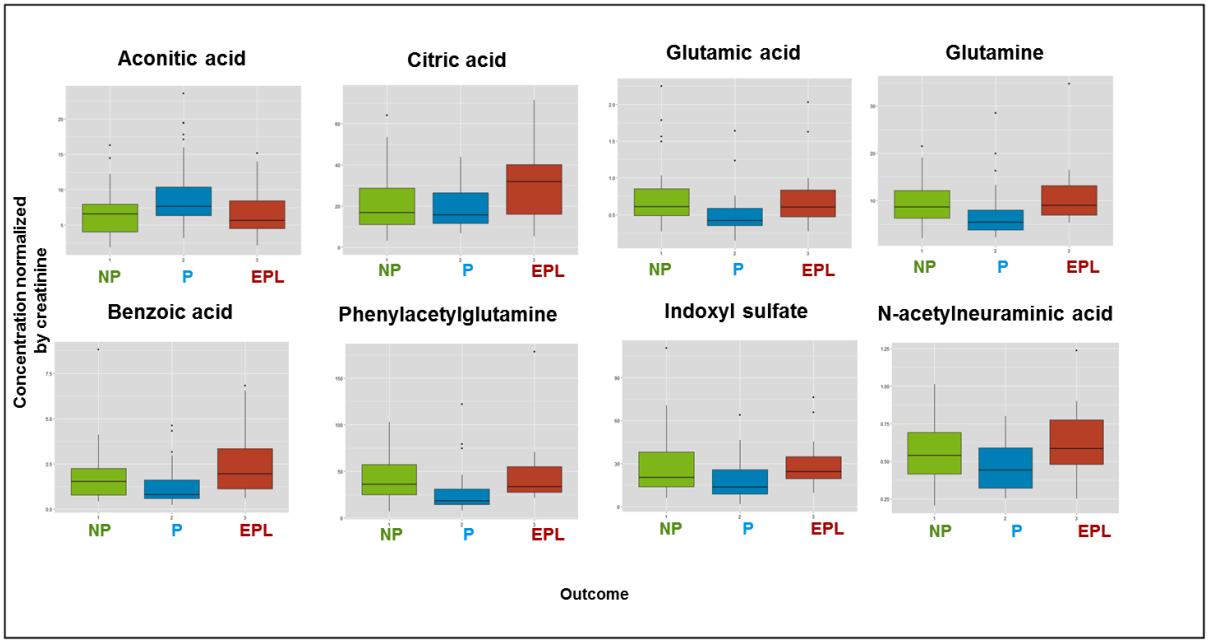

Fig.2 Box plots for eight metabolites, relative quantification showing significant differences between P and NP or EPL groups; one-way ANOVA p<0.05 (NP-Non-Pregnant, P-Pregnant, and EPL – early pregnancy loss).

Maternal urine profiling has been already used for prenatal diagnostics and early prediction of several poor pregnancy outcomes.[43, 44] The urine samples assayed in the present study have

been collected at a very early stage of pregnancy (in the first few weeks when pregnancy was confirmed by HCG test). All significant differences in metabolite concentrations presented by the p-value in the Table 4 and in the boxplots were observed for comparisons of Pregnant *versus* Non-Pregnant and *versus* EPL groups. No significant differences were found for Non-Pregnant *versus* EPL group in this targeted analysis. By characterizing metabolic differences between pregnancy, non-pregnancy and early pregnancy loss, it was hoped that we might uncover new information relating to the metabolic conditions indicating likelihood of a female patient for conception and implantation of an embryo.

Higher excretion levels of succinic and citric acids along with lower levels of aconitic acid in EPL compared to P group could be associated with shifts in energy demand and production that might be one of the reasons of the poor pregnancy outcome. Alterations in succinate levels, relating to perturbations in Krebs cycle have been associated with adverse pregnancy outcomes. For example, increased succinate levels in maternal urine have been reported in cases of central nervous system fetal malformations[43], whereas lower concentrations in the amniotic fluid have been associated with fetal malformation cases[44]. Citrate is converted to isocitrate in Krebs cycle through intermediate *cis*-aconitate. The lower excretion of aconitate and higher excretion of citrate observed for EPL group could indicate the truncation of this step in Krebs cycle due to the inhibition of the enzyme aconitase which could be caused by oxidative stress.[48]

The urinary concentration of nicotinamide was higher in the NP and EPL groups than the P group. This compound, also called vitamin B3, comes from the diet sources but can be also produced in the cells from NAD+. It is responsible for the redox metabolism which is essential for the energy release from nutrients during beta oxidation, glycolysis, and Krebs cycle.

Elevated levels of glutamine, observed here for EPL and NP groups compared to P group, can also reflect perturbations in energy production (through glutaminolysis) as well as elevated immune response.[49] The increase in glutamine serum concentration has been previously observed in metabolic profiling of women with idiopathic recurrent spontaneous miscarriages and it has been associated with exaggerated inflammatory response and vascular dysfunction leading to poor endometrial receptivity.[42] In our study, another indication of some inflammatory process could be the increased excretion of N-acetylneuraminic acid observed for EPL group. The adverse inflammatory immune response has been previously associated with various reproductive failures during ovulation, fertilization, implantation and pregnancy.[50] A recent ^{1}H NMR study of urinary metabolic profiles in early pregnancy (end of

the first trimester) in a prospective mother-child cohort Rhea study intending to discover metabolites associated with future birth outcomes, such as preterm birth and fetal growth restriction, has shown that women with induced preterm birth had a significant increase in a N-acetyl glycoprotein resonance. This resonance is frequently associated with inflammation-induced acute phase proteins that could be a result of release of cytokines by adipose tissue and it was positively correlated with BMI and significantly associated with preterm birth in obese and overweight women.[38]

The other metabolites, for which significantly higher excretion levels were observed for EPL and NP groups (see Table 4 and Figure 2), could be related to the activity of gut microbiome (benzoic acid, phenylacetylglutamine, indoxyl sulfate) and potentially oxidative stress in the case of indoxyl sulfate.

Although the biochemical differences between pregnancy outcomes observed here are interesting and warrants further investigation, the results need to be interpreted with caution based on the relatively small number of participants in each group.

To obtain additional information, the HRMS global profiles obtained for these samples by nESI-HRMS method were subjected to multivariate analysis.

Untargeted analysis of full-scan high resolution data

The nESI-HRMS method offers the ability to perform untargeted statistical analysis using the full-scan data and the combination of targeted and untargeted approaches offers an attractive solution to obtaining information on specific pathways within a systems environment.

The representative Principal Component Analysis (PCA) scores plots of pre-processed (binned, normalized, and scaled to unit variance) full-scan data in positive and negative ion modes presented in the Figure 3a, b do not show significant clustering between three different outcomes, but clearly show outlying samples, most of which from P group. It was previously shown that DIMS had comparable classification and prediction capabilities to LC-MS in the biomarker discovery study of serum samples from kidney cancer patients.[51] Interestingly, in our study preliminary results of PCA of the data obtained for the same set of samples by global profiling UPLC-MS (work in progress), have shown a very similar picture finding the same samples to be outliers as by nESI-HRMS. It has been assumed from both techniques that these samples were more diluted as they were distributed in the direction of the dilution QC samples. The application of supervised techniques such as PLS-DA did not significantly enhance the class separation. Both values, R^2X and R^2Y, goodness of fit for X (spectral variables) and Y (class variables) matrices respectively, (42.6% and 37.6% in positive ion

mode, and 71.4% and 29.6% in negative ion mode, respectively) and Q^2Y, predictive ability, (13.3% in positive and 11.5% in negative ion mode) are not high underscoring the low robustness of the model, most likely due to low statistical power. However, it is important to note that the EPL samples obtained from the patients with several recurrent pregnancy losses (57, 16, and 61) differentiate more from the other classes. Further in depth analyses comparing the outcomes pairwise applying supervised techniques such as OPLS-DA may uncover discriminant features, the identification of which can be then performed by tandem MS.

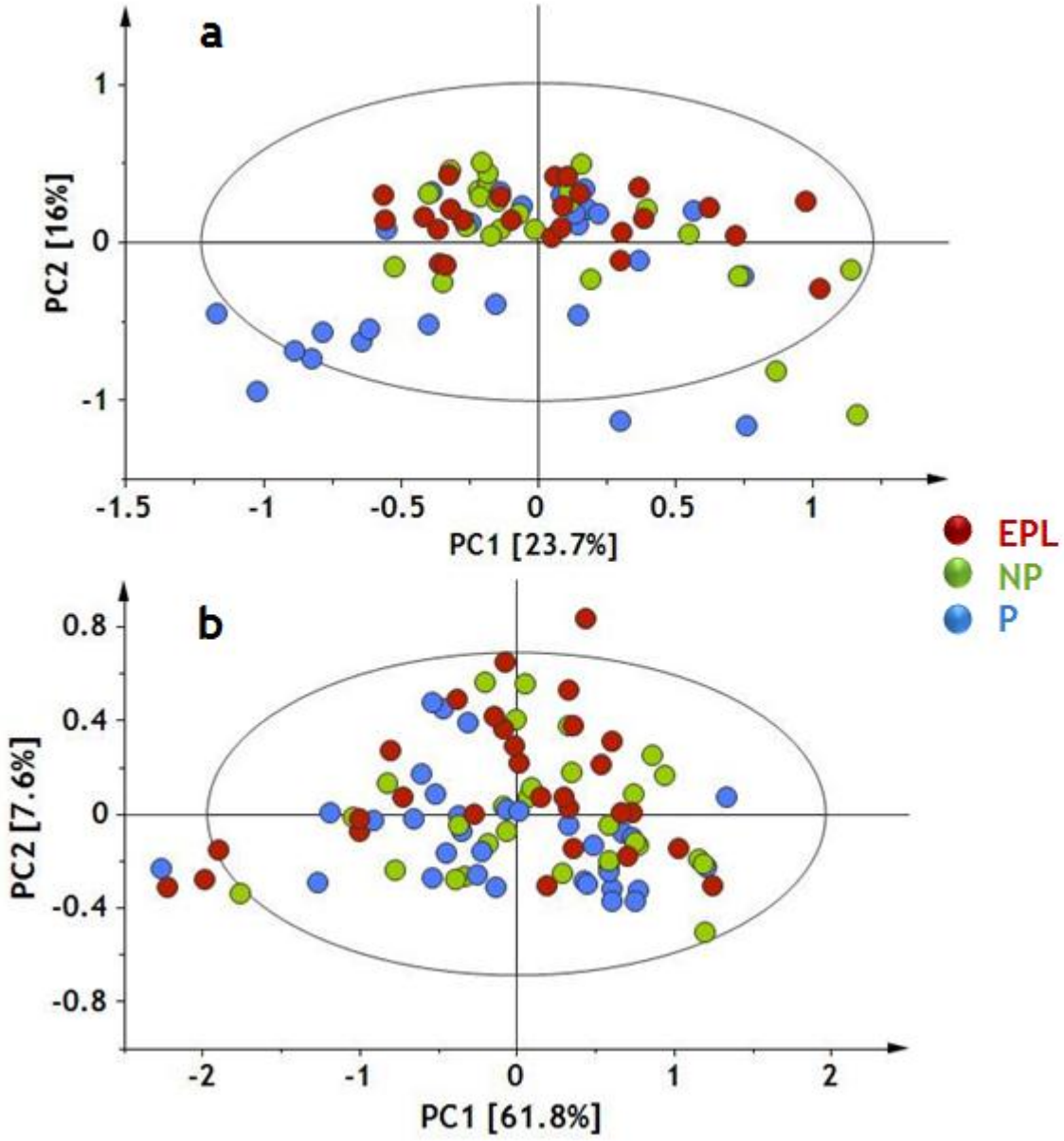

Fig.3 Scores scatter plot of PCA of the data obtained in positive ionization mode, R^2X=39.8%, Q^2X=32.7% (a), and in negative ionization mode, R^2X=73%, Q^2X=69.9% (b).

CONCLUSIONS

We present an alternative to a conventional UPLC-MS (MS/MS) procedure, which has the potential to measure hundreds of compounds in parallel with excellent reproducibility and little effort. Automated nESI-HRMS technology presents advantages such as high sensitivity, high-throughput, wide dynamic range due to low ion suppression, and lack of carry-over. We show that the sample preparation for the analysis is simple, consisting of sample dilution and maintenance of optimal water-organic solvent proportion. The method requires minimal amounts of sample and the injectable volume is of the order of nL. Apart from low specimen consumption, it also ensures a stable instrument performance and lack of any contamination. Such an easy protocol can be easily adopted and optimized for other biofluids such as plasma, serum, CSF, and saliva. We have already undertaken some preliminary experiments with serum, plasma, CSF, saliva, cord blood using this method. The main optimization step in sample preparation consists in deproteination (by using cold methanol) and filtration through 1μm glass fibre filter plate (unpublished data).

Use of high resolution mass spectrometry, instead of the low resolution instruments preferred for high throughput metabolic fingerprinting based on DIMS, increases the selectivity of qualitative and quantitative analyses. The tentative identification of compounds can be made directly from the exact mass values. The following off-line or even in-line fragmentation experiments together with isotopic pattern analysis (available from the high resolution data) can help in confirming the identities. The main drawback of the method lies in the inability to differentiate between the structural isomers, but at the same time it allows high resolution full-scan data to be acquired that can be processed any time post analysis for quantification, semi-quantification of accurate mass peaks or untargeted fingerprinting.

In order to develop and validate the method, a cohort of human urine samples (n=101) was used to measure a selected set of metabolites. We show the application of the method for the targeted quantification of several metabolites by the method of standard additions as well as the ability to perform the relative quantification of other metabolites and untargeted fingerprinting of high resolution full scan data.

This method is not intended to replace the hyphenated LC-MS techniques but it may become a valuable complementary tool for a large scale metabolic phenotyping laboratories. The above results indicate that the HRMS method offers new interesting possibilities for multiplexed quantitative analysis of large population studies in the fields of epidemiology and

stratified medicine. The quantification possibilities provided by current method, although not of the level of decision-making or clinical action, allow characterization of population phenotypes, estimate the "healthy" and "disease" levels of metabolites, differentiate between case and control and select representative samples for further LC/MS and MS/MS analysis. Moreover, nESI chip based mass spectrometry provides additional information in the form high resolution accurate mass global profile of the samples.

The current method is still in a phase of further development to optimize experimental and computational workflows which will allow for improvements in method selectivity, data analysis, interpretation, and metabolite identification but merits validation across multiple disease cohorts and biological matrices.

ACKNOWLEDGEMENTS

The research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

1. E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. S. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. C. Zhao, J. K. Nicholson and P. Elliott, *Nature*, 2008, 453, 396-U350.
2. J. K. Nicholson, E. Holmes, J. M. Kinross, A. W. Darzi, Z. Takats and J. C. Lindon, *Nature*, 2012, 491, 384-392.
3. O. Beckonert, H. C. Keun, T. M. D. Ebbels, J. G. Bundy, E. Holmes, J. C. Lindon and J. K. Nicholson, *Nat Protoc*, 2007, 2, 2692-2703.
4. O. Beckonert, M. Coen, H. C. Keun, Y. L. Wang, T. M. D. Ebbels, E. Holmes, J. C. Lindon and J. K. Nicholson, *Nat Protoc*, 2010, 5, 1019-1032.
5. A. Roux, Y. Xu, J. F. Heilier, M. F. Olivier, E. Ezan, J. C. Tabet and C. Junot, *Anal Chem*, 2012, 84, 6429-6437.
6. D. Ryan, K. Robards, P. D. Prenzler and M. Kendall, *Anal Chim Acta*, 2011, 684, 17-29.
7. E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes and J. K. Nicholson, *Nat Protoc*, 2010, 5, 1005-1018.

8. E. C. Y. Chan, K. K. Pasikanti and J. K. Nicholson, *Nat Protoc*, 2011, 6, 1483-1499.
9. E. C. Horning and M. G. Horning, *Clin Chem*, 1971, 17, 802-&.
10. E. C. Horning and M. G. Horning, *J Chromatogr Sci*, 1971, 9, 129-&.
11. S. Bouatra, F. Aziat, R. Mandal, A. C. Guo, M. R. Wilson, C. Knox, T. C. Bjorndahl, R. Krishnamurthy, F. Saleem, P. Liu, Z. T. Dame, J. Poelzer, J. Huynh, F. S. Yallou, N. Psychogios, E. Dong, R. Bogumil, C. Roehring and D. S. Wishart, *Plos One*, 2013, 8.
12. K. Dettmer, P. A. Aronov and B. D. Hammock, *Mass spectrometry reviews*, 2007, 26, 51-78.
13. C. Junot, F. Fenaille, B. Colsch and F. Becher, *Mass spectrometry reviews*, 2013, DOI: 10.1002/mas.21401.
14. J. F. Xiao, B. Zhou and H. W. Ressom, *Trends in analytical chemistry : TRAC*, 2012, 32, 1-14.
15. Z. T. Lei, D. V. Huhman and L. W. Sumner, *J Biol Chem*, 2011, 286, 25435-25442.
16. R. S. Plumb, G. Fujimoto, J. Mather, W. B. Potts, P. D. Rainville, N. J. Ellor, C. Evans, J. R. Kehler and M. E. Szapacs, *Bioanalysis*, 2012, 4, 605-615.
17. Y. Q. Xia, J. Lau, T. Olah and M. Jemal, *Rapid Commun Mass Sp*, 2011, 25, 2863-2878.
18. N. R. Zhang, S. Yu, P. Tiller, S. Yeh, E. Mahan and W. B. Emary, *Rapid Commun Mass Sp*, 2009, 23, 1085-1094.
19. E. I. Hamelin, W. Bragg, R. L. Shaner, L. L. Swaim and R. C. Johnson, *Rapid Commun Mass Sp*, 2013, 27, 1697-1704.
20. J. G. Stroh, C. J. Petucci, S. J. Brecker, N. Huang and J. M. Lau, *J Am Soc Mass Spectr*, 2007, 18, 1612-1616.
21. M. Krauss, H. Singer and J. Hollender, *Anal Bioanal Chem*, 2010, 397, 943-951.
22. S. Lacorte and A. R. Fernandez-Alba, *Mass spectrometry reviews*, 2006, 25, 866-880.
23. J. Denes, E. Szabo, S. L. Robinette, I. Szatmari, L. Szonyi, J. G. Kreuder, E. W. Rauterberg and Z. Takats, *Anal Chem*, 2012, 84, 10113-10120.
24. J. Draper, A. J. Lloyd, R. Goodacre and M. Beckmann, *Metabolomics*, 2013, 9, S4-S29.
25. D. P. Enot, W. Lin, M. Beckmann, D. Parker, D. P. Overy and J. Draper, *Nat Protoc*, 2008, 3, 446-470.
26. K. W. Dunn-Meynell, S. Wainhaus and W. A. Korfmacher, *Rapid Commun Mass Sp*, 2005, 19, 2905-2910.

27. A. Schmidt, M. Karas and T. Dulcks, *J Am Soc Mass Spectr*, 2003, 14, 492-500.
28. K. O. Boernsen, S. Gatzek and G. Imbert, *Anal Chem*, 2005, 77, 7255-7264.
29. E. R. Wickremsinhe, B. L. Ackermann and A. K. Chaudhary, *Rapid Commun Mass Sp*, 2005, 19, 47-56.
30. M. Moriarty, M. Lehane, B. O'Connell, H. Keeley and A. Furey, *Talanta*, 2012, 90, 1-11.
31. Y. P. Zhang, L. Qiu, Y. M. Wang, X. Z. Qin and Z. L. Li, *Analyst*, 2014, 139, 1697-1706.
32. W. R. J. M. Kirwan J. A., Broadhurst D. I., Viant M. R., *Scientific Data*, 2014, DOI: doi:10.1038/sdata.2014.12.
33. T. G. Payne, A. D. Southam, T. N. Arvanitis and M. R. Viant, *J Am Soc Mass Spectr*, 2009, 20, 1087-1095.
34. J. A. Kirwan, D. I. Broadhurst, R. L. Davidson and M. R. Viant, *Anal Bioanal Chem*, 2013, 405, 5147-5157.
35. R. Gonzalez-Dominguez, R. Castilla-Quintero, T. Garcia-Barrera and J. L. Gomez-Ariza, *Analytical biochemistry*, 2014, 465C, 20-27.
36. US Department of Health and Human Services FDA, *Guidance for Industry, Bioanalytical Method Validation* 2001.
37. C. A. Smith, G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan and G. Siuzdak, *Ther Drug Monit*, 2005, 27, 747-751.
38. L. Maitre, E. Fthenou, T. Athersuch, M. Coen, M. B. Toledano, E. Holmes, M. Kogevinas, L. Chatzi and H. C. Keun, *Bmc Med*, 2014, 12.
39. A. G. Savitzky, M. J. E. , *Anal Chem*, 1964, 36, 1627-1639.
40. F. Dieterle, A. Ross, G. Schlotterbeck and H. Senn, *Anal Chem*, 2006, 78, 4281-4290.
41. J. Trygg, E. Holmes and T. Lundstedt, *Journal of proteome research*, 2007, 6, 469-479.
42. P. Banerjee, M. Dutta, S. Srivastava, M. Joshi, B. Chakravarty and K. Chaudhury, *Journal of proteome research*, 2014, 13, 3100-3106.
43. S. O. Diaz, A. S. Barros, B. J. Goodfellow, I. F. Duarte, E. Galhano, C. Pita, C. Almeida Mdo, I. M. Carreira and A. M. Gil, *Journal of proteome research*, 2013, 12, 2946-2957.
44. S. O. Diaz, J. Pinto, G. Graca, I. F. Duarte, A. S. Barros, E. Galhano, C. Pita, C. Almeida Mdo, B. J. Goodfellow, I. M. Carreira and A. M. Gil, *Journal of proteome research*, 2011, 10, 3732-3742.

45. J. W. Lee, V. Devanarayan, Y. C. Barrett, R. Weiner, J. Allinson, S. Fountain, S. Keller, I. Weinryb, M. Green, L. Duan, J. A. Rogers, R. Millham, P. J. O'Brien, J. Sailstad, M. Khan, C. Ray and J. A. Wagner, *Pharm Res*, 2006, 23, 312-328.
46. W. B. Dunn, I. D. Wilson, A. W. Nicholls and D. Broadhurst, *Bioanalysis*, 2012, 4, 2249-2264.
47. M. M. Kushnir, G. Komaromy-Hiller, B. Shushan, F. M. Urry and W. L. Roberts, *Clin Chem*, 2001, 47, 1993-2002.
48. L. J. Yan, R. L. Levine and R. S. Sohal, *Proceedings of the National Academy of Sciences of the United States of America*, 1997, 94, 11168-11172.
49. P. Newsholme, *J Nutr*, 2001, 131, 2515s-2522s.
50. J. Kwak-Kim, S. H. Bao, S. K. Lee, J. W. Kim and A. Gilman-Sachs, *Am J Reprod Immunol*, 2014, 72, 129-140.
51. L. Lin, Q. A. Yu, X. M. Yan, W. Hang, J. X. Zheng, J. C. Xing and B. L. Huang, *Analyst*, 2010, 135, 2970-2978.