Journal of Chromatography B, 871 (2008) 299–305

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Evaluation of the repeatability of ultra-performance liquid chromatography–TOF-MS for global metabolic profiling of human urine samples☆

Helen G. Gika[a,b,*], Euan Macpherson[c], Georgios A. Theodoridis[a,b], Ian D. Wilson[a]

[a] AstraZeneca, Department of Clinical Pharmacology, Drug Metabolism and Pharmacokinetics, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK
[b] Laboratory of Analytical Chemistry, Chemistry Department, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece
[c] AstraZeneca, UK Biostatistics, Parklands, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK

ARTICLE INFO

Article history:
Received 28 February 2008
Accepted 30 May 2008
Available online 6 June 2008

Keywords:
Metabolite profiling
Ultra-performance liquid chromatography
Time-of-flight mass spectrometry
Metabonomics
Metabolomics

ABSTRACT

The application of reversed-phase ultra-performance liquid chromatography, based on the use of sub 2 μm particles, combined with time-of-flight mass spectrometry has been investigated for the production of global metabolite profiles from human urine. The stability and repeatability of the methodology, which employed gradient elution, was determined by the repeat analysis of a pooled quality control (QC) sample. As seen in previous studies conducted with conventional LC–MS an element of system conditioning was required to obtain reproducible data, as the initial injections were unrepresentative. However, once the system had equilibrated excellent repeatability in terms of retention time, signal intensity and mass accuracy was seen providing confidence that for this matrix, the within-day repeatability of UPLC–TOF-MS was sufficient to assure data quality in global metabolic profiling applications.

1. Introduction

The detection of biomarkers by metabonomics ("the quantitative measurement of the dynamic multiparametric response of living systems to pathophysiological stimuli or genetic modification") [1,2] is currently performed using a range of analytical techniques such as NMR spectroscopy or, increasingly, separation techniques (CE, LC and GC) as reviewed elsewhere [3]. Recently LC, particularly when coupled to mass spectrometry has emerged as a very promising method for global metabolite profiling (e.g. [3–6] and references therein). Experience with the application of LC–MS-based methods [6–8] has however, shown that repeatability (within- and between-day) is harder to achieve compared to e.g., NMR spectroscopy, which in contrast has been shown to be both stable and reproducible (e.g. [9]). Previously we have examined the repeatability, suitability and potential of the LC–MS analytical platform for the metabonomic analysis of human urine, assessing this on the basis of a biological QC [8,10] approach. In this way we were able to demonstrate a level of stability of this analytical methodology within a typical analytical run of ca. 20 h that gave some confidence in the ability of LC–MS to provide meaningful data [8]. This work was based on the use of conventional reversed-phase gradient chromatography on a 3.5 μm C18 bonded stationary phase. More recently chromatography on sub 2 μm stationary phases, with solvent delivery at conventional flow rates but at elevated pressures has been introduced (ultra-performance liquid chromatography or UPLC) offering much greater chromatographic efficiency. UPLC offers improved resolving power and detection sensitivity compared to conventional LC. When coupled to a mass spectrometer that can provide rapid scanning and data acquisition, as found with time-of-flight (TOF-MS) instruments, the system offers a powerful tool for global metabolite profiling. There are already a number of examples of the use of UPLC in metabonomic applications [11–13]. Given that, as we have noted elsewhere [8], for LC–MS-based metabonomics data to become fully accepted in e.g., regulatory submissions or for systems biology studies data quality is paramount and must be demonstrated for each analytical platform. The investigations reported here were therefore designed to provide an indication of within-day analytical performance of UPLC–MS analysis for human urine samples.

2. Experimental

2.1. Reagents and materials

All solvents used were of HPLC grade and obtained from Fisher Scientific (Loughborough, Leicestershire, UK). Formic acid

☆ This paper is part of a special volume entitled "Hyphenated Techniques for Global Metabolite Profiling", guest edited by Georgios Theodoridis and Ian D. Wilson.

* Corresponding author. Tel.: +30 2310 999251/997718; fax: +30 2310 999221/997719.
E-mail address: elegik@hotmail.com (H.G. Gika).

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.05.048

of analytical grade was also purchased from Fisher Scientific. The metabonomics MS system test mix was supplied by Waters Corporation (Milford, USA). Water (18.2 MΩ) was obtained from a Purelab Ultra system from Elga (Bucks, UK).

2.2. Samples

Urine samples collected from 114 female subjects, and stored at −80 °C, were used for this study. Sample treatment prior to analysis was limited to thawing at room temperature, centrifugation at 17,900 × *g* (13,000 rpm) for 5 min to remove particulates and dilution with water (250 μL of the sample was mixed with 750 μL of water and vortexed). 300 μL aliquots of the diluted samples were then placed in autosampler vials and were then centrifuged again at 2700 × *g* (3000 rpm) for 7 min before the analysis. In parallel a quality control (QC) sample was prepared by mixing equal volumes (100 μL) from each of the 114 samples [7,10]. An aliquot of 200 μL of this pooled urine sample was diluted in the same way as the test samples with the addition of 750 μL of water and was analysed several times throughout the run in order to provide a measure of the system's stability and performance.

Further to this a synthetic test mixture (metabonomics MS system test mix; Waters) was analysed in the beginning, the middle and the end of the run as an additional means of controlling the MS system. The residue contained in the vial was reconstituted in 1 mL of water: acetonitrile (98:2, v/v). The resulting solution consisted of theophilline (*m*/*z* 181.1; 120 μg/mL), caffeine (*m*/*z* 195.2; 120 μg/mL), hippuric acid (*m*/*z* 180.2; 120 μg/mL), 4-nitrobenzoic acid (*m*/*z* 168.1; 60 μg/mL) and nortriptyline (*m*/*z* 264.4; 45 μg/mL).

2.3. UPLC–MS analysis

2.3.1. Ultra-performance liquid chromatography

Chromatography was performed with an ACQUITY UPLC™ System on a 2.1 mm × 100 mm Acquity BEH C18 column, 1.7 μm particle size (Waters Corporation, Milford, USA) at 50 °C. Gradient elution, at a flow rate of 400 μL/min, was performed using 0.1% (v/v) aqueous formic acid as solvent A and acetonitrile (also containing 0.1% formic acid (v/v)) as solvent B. Initial gradient conditions were 100% A for 1 min with a linear rise to 90-10 A:B, v/v over the next minute. These conditions were held constant for a further 1 min, followed by a further linear increase to 100% B over the next 6.5 min. The eluent was then held constant at 100% B for 0.5 min, after which it was returned to 100%A over 0.1 min followed by 2 min equilibration with this eluent prior to the next injection. For analysis 10 μL of sample was injected on to the column.

Injection valves and syringes were washed twice before each injection with 600 μL water–methanol 98:2, v/v) and 600 μL water–methanol, 20:80, v/v. During the analysis the samples were kept at 4 °C at the autosampler.

2.3.2. ESI mass spectrometry

Mass spectrometry was performed using a Waters Micromass® Q-TOF Micro (Milford, MA, USA) operating in positive ion electrospray mode. The capillary and cone voltages were set at 3 kV and 30 V. The desolvation temperature was set to 250 °C and the source temperature to 120 °C. The cone gas was set to a flow rate of 10 L/h and the desolvation gas flow was maintained at 700 L/h. For mass accuracy a LockSpray™ interface was used with leucine-enkephalin (556.2771 amu) solution (0.25 μg/L) infused at 30 μL/min as the lock mass. Full scan data were collected from 80 to 850 *m*/*z* over a period of 9 min with a scan time of 0.3 s and interscan delay of 0.1 s. MassLynx™ software (Waters) was used for system controlling and data acquisition.

2.3.3. Sample analysis

The samples were analysed in a single run, which in total lasted for ca. 29 h, to obtain metabolite profiles. A standard test mix (see above) was analysed at the beginning, in the middle and at the end of the run. The pooled QC samples were injected five times at the beginning of the run in order to condition or equilibrate the system and then every ten samples to further monitor the stability of the analysis. The acquired QC data were used to investigate the analytical variability in the whole run. This was necessary in order to evaluate whether the analytical system had changed (and to what extent) over the time course of the analysis. This is critical for evaluating the variation in the analytical results and therefore the reliability of the metabolite profiling data.

2.4. Data processing

The raw spectrometric data acquired were processed by MarkerLynx™ application manager (Waters). MarkerLynx uses ApexTrack peak integration to detect chromatographic peaks. The track peak parameters were set as follows: peak width at 5% height 15 s, peak-to-peak baseline noise 80, intensity threshold 100, mass window 0.05 amu, retention time window 0.2 min, noise elimination level 6 and mass tolerance 0.50 amu.

Peak list data obtained by MarkerLynx were further processed by Simca P version 11 from Umetrics (Windsor, UK) for multivariate data analysis. Basic applications such as Principal Components Analysis and other statistical tools were implemented. Further statistical analysis of the data was carried out using the R statistical software package [14].

3. Results and discussion

As we have shown elsewhere, UPLC provides superior results compared to conventional HPLC in terms of resolution and increased numbers of ions detected [11,12]. In previous studies using LC–MS we implemented an analysis time of ca. 10 min per sample as this provides a reasonable compromise between the number of ions detected and throughput (e.g. [7]). We therefore opted for a comparable analysis time and reversed-phase gradient conditions in the current work to aid comparison with this earlier study. In Fig. 1 a heat map from the analysis of a quality control sample from the present study is shown, giving an indication of the distribution of masses and retention times encountered for this typical sample. The chromatogram clearly illustrates the complexity of the analysed sample, with a large number of ions detected in a short analysis time.

3.1. PCA visualisation

Having established from an examination of the results of the test mixture (see Ref. [7]) that the run seemed repeatable for the stability of retention time, peak shape, mass accuracy and signal intensity we proceeded to investigate the data derived from the QC samples.

As in previous studies [7,8], the preliminary evaluation of the quality of the data derived from the analysis of urine samples was based on the behaviour of the pooled sample, which was used as a QC sample to assess the analytical variability across the run. This pooled sample provided a sample that was representative of the study samples, essentially containing all of the analytes that would be encountered during the course of the analysis. Simply stated, the hypothesis was that, if repeatability was absolute, all of the QC sample injections would give the same UPLC–MS profiles. Clearly in practice this cannot be fulfilled since there are many parameters which make run-to-run repeatability challenging. However, as

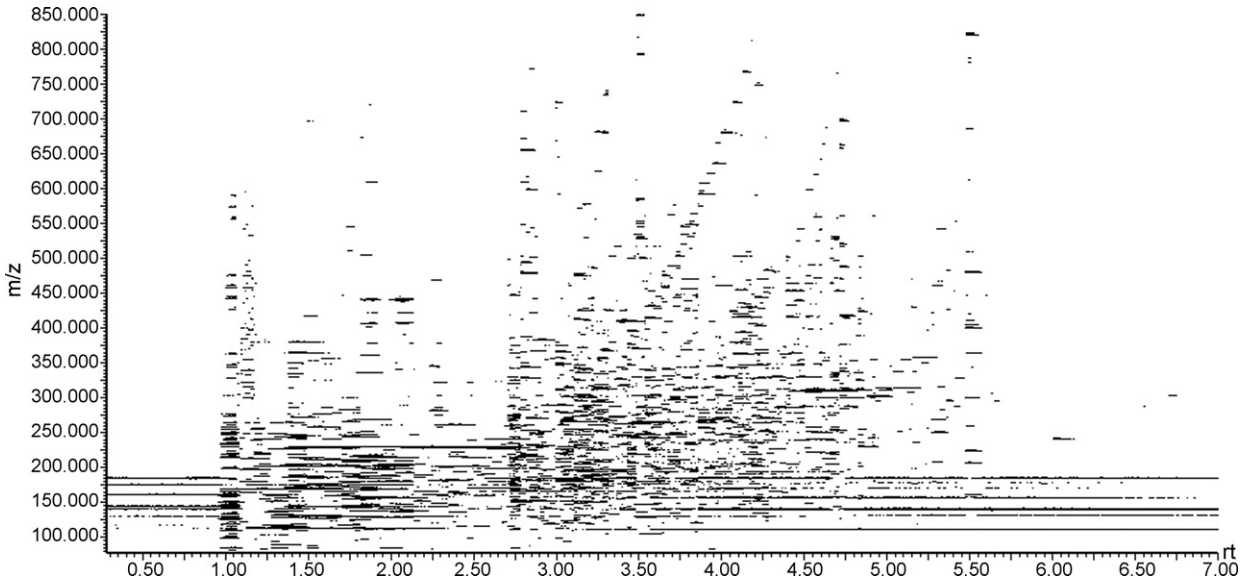

Fig. 1. A heat map from the UPLC–MS analysis of the pooled quality control (QC) sample showing the distribution of ion (m/z) with elution time.

these QC samples were analysed regularly through the run, examination of the data derived from them provided a representative indication of sample-to-sample variability during the course of the whole analysis.

An initial overview of the quality of the run was obtained by PCA of the whole data set including all the QC injections (i.e., including the initial conditioning injections). Based on the hypothesis described above the closer the QC samples appear on the scores plot the more reproducible the performance of the UPLC–MS system should be. One thing that should be taken into consideration when using this pragmatic approach to assessing run quality is that, since the scaling in a PCA scores plot does not offer absolute quantitative measurements, there is always the risk for these plots by themselves to be misleading. Thus, where test samples are very different from one another, the first two Principal Components may represent a large proportion of the variability in the data (perhaps more than 50%) and therefore differences present in the QC data (caused by e.g., instrument instability, column ageing or sample decay, or all these three factors simultaneously) may be suppressed by the overall variability of the sample set. This would result in the points in the scores plot that represent the QC samples being tightly clustered. On the other hand, in samples with smaller differences (e.g., in samples from normal healthy animals) the overall variation is likely to be much smaller and hence the variation observed in the analysis of the QCs is not negligible. In such a case the QCs could appear more scattered in the scores plot. In Fig. 2A the PCA plot of the whole dataset (applying pareto scaling to the raw data) is given. Fig. 2B gives a 2D PCA plot of the data set after the removal of the first five QC samples. As can be seen the QCs (coloured in red) are tightly clustered in the middle of the plot in both figures. This type of result thus provides some evidence that the UPLC–MS system was stable and therefore providing reliable data suitable for further statistical analysis. The first and second components, including the first conditioning injections, accounted for 10.5% and 8.04% of the total variability in the scaled data respectively.

By exploring the time dependency of the PCA scores we can obtain additional insight into trends and drifts over the course of the analysis of the batch. This is illustrated in Fig. 3 which displays the first component t[1] as a plot versus the samples in run order. In this way it shows the t[1] as it evolved with time (with the 2 and 3σ limits also shown). Clearly, in this case, after the initial condition-

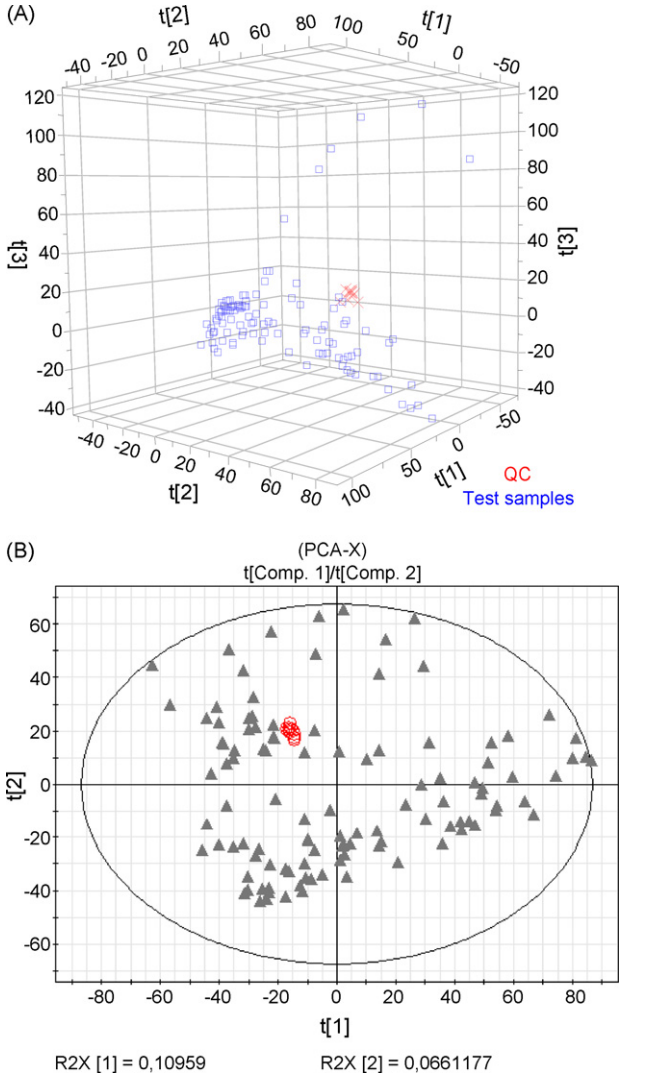

Fig. 2. (A) PCA 3D plot of the whole data set; (B) PCA 2D plot with the omission of the first five QCs. QCs are shown in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

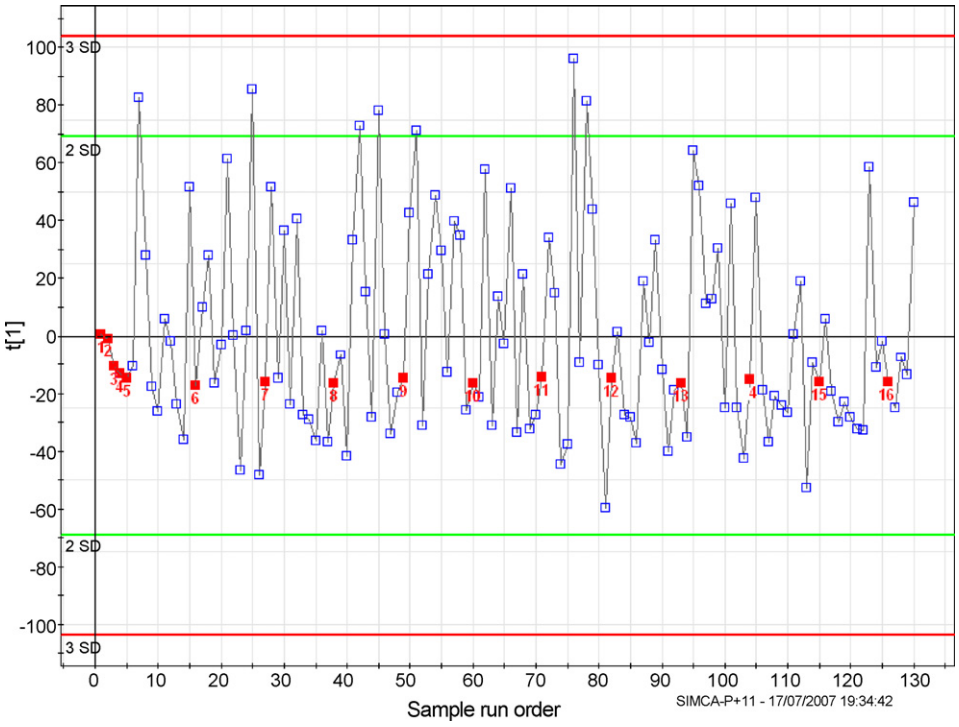

Fig. 3. Time series plot of the first PCA component (t[1] vs sample run order). QCs are shown as solid squares (coloured in red) and test samples are open squares (coloured in blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

ing phase, there was little variation in the QC samples throughout the whole of the 29 h analytical run. This type of result gives some confidence that the analysis was stable for the duration of the run, after the initial settling in period, and provides a further means of assessing the quality of the data. The figure also illustrates the wide variability between the individual urine samples which, compared to the variability of the QC population is significant, thus providing re-assurance (together with data such as that shown in Fig. 2) that the differences observed between the different samples are correctly attributed to biological variation and not to the variability of the analytical system.

3.2. Examination of mass chromatograms

Armed with the information from the test mixture analysis and the PCA of the QCs a more detailed assessment of the quality of the data was performed. Initially the stability of the chromatographic system was examined with respect to the retention time repeatability of selected ions present in the QCs. Obviously urine is a very complex sample (Fig. 1) containing a very wide range of components including neutral, acidic, basic and amphoteric compounds. In an unstable separation system, where subtle modifications of the stationary phase might be taking place (e.g., the exposure, or masking, of acidic silanol groups etc.) it can be anticipated that retention times may also alter, or even that elution order may change. To obtain specific information on the run-to-run repeatability of the retention times a number of peaks ($n = 11$) were selected from the mass chromatogram and these were monitored as extracted ion chromatograms along the whole data set of 16 QC injections. The ions that were examined covered the retention time range from 1 to 6 min. They were: m/z 142.0, 1.03 min; m/z 169.03, 1.83 min; m/z 152.06, 2.81 min; m/z 313.08, 3.26 min; m/z 265.1, 3.5 min; m/z 180.1, 3.51 min; m/z 205.12, 4.12 min; m/z 286.19, 4.36 min; m/z 302.23, 4.91 min; m/z 400.19, 5.51 min; m/z 242.28, 6.04 min. When these ions were examined from the injection of QC5 to QC16 (i.e., after the equilibration period) the retention time variation was found to be negligible with CV values less than 1% in all these cases. In addition peak area repeatability for the same ions expressed as CV values ranged between 7% and 9%. In Fig. 4 the repeatability of retention time and peak height of a representative ion is depicted in successive QCs samples, covering a run time of 14 h (67 injections). The extracted ion chromatograms for all the QCs were remarkably stable but for purely illustrative purposes results are shown for only seven of the QCs. The repeatability in retention time and signal intensity seen here was very satisfactory, giving further confidence that the UPLC–TOF-MS system was operating robustly providing reproducible QC sample fingerprints. We also observed good mass accuracy for these ions with variability lower than 0.005 amu.

3.3. Processed data analysis

Thereinafter the analytical variability in the processed data from the peak picking and alignment software was examined. The raw data set of 16 QCs was extracted with the MarkerLynx software (using the parameters given in the experimental section) to provide a peak list file containing 3414 variables (data set A). A separate dataset of 5522 peaks (data set B) was generated when the same settings in MarkerLynx were applied to the raw data of all 114 samples plus the 16 QCs. Each of these peaks corresponds to an intensity measurement at a point in ion mass-retention time space. These features are considered to be metabolite ions present in the QC sample and were examined in terms of consistency in intensity all through the run as a proof (besides the retention times) that the system was operating robustly.

The FDA recommends that a coefficient of variation (CV) of 15% of the nominal value represents an acceptable degree of repeatability (except for concentrations close to the LOQ where 20% is considered to be adequate) in bioanalytical methods for drugs [15]. Whilst no criteria have yet been suggested for metabolic profiling data it nevertheless seems reasonable to start the debate by aim-

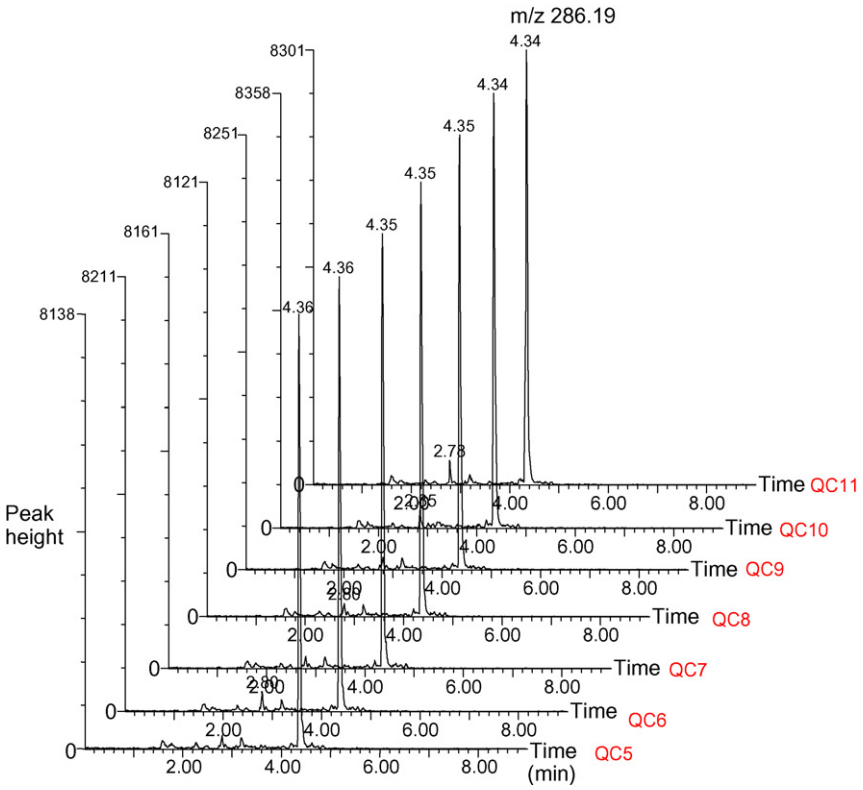

Fig. 4. Extracted ion chromatograms of an ion (*m/z* 286.19, retention time 4.3 min) in 7 sequential QCs samples run in between of 67 test samples demonstrating the repeatability in peak height and retention time.

ing for this level of repeatability if possible. We therefore examined the UPLC–MS results using similar acceptance criteria as a means of determining the quality of the data.

To determine how many of the detected peaks might be acceptable under the FDA criteria mentioned above, the coefficient of variation of intensities was calculated for all peaks across the QC samples (zero intensity data was treated as missing). Various filters of the selected peaks were considered to illustrate how possible filtering could improve the number of peaks being acceptable under the criteria and the quality of the data. Table 1 summarises the results of the QC data when data set A was considered. Thus, even with the five start-up QCs included, it can be seen that 666 (19.5%) of the 3414 peaks were acceptable using the ≤15% CV criterion, whilst 1054 (30.9%) were acceptable at the ≤20% level. The number of peaks meeting these criteria improved when focussing only on the peaks which were present, i.e., had a positive intensity, across all 16 QC samples: 285 (35.5%) out of 803 at CV ≤ 15% and 424 (52.8%) at the CV ≤ 20% level respectively.

Not unexpectedly, given the obvious differences between the bulk of the QCs and some of the initial start-up QCs, when these were left out of the data analysis a higher percentage of the peaks met the acceptance criteria. So, when only the last 11 QCs were considered, 974 (28.5%) of all 3414 peaks met the CV ≤ 15% and 1356 peaks (39.7%) for the CV ≤ 20% criteria. With the five start-up QCs removed, there were 1089 peaks for which a positive intensity was recorded in all of the 11 remaining QCs. For 495 (45.5%) of these peaks the CV ≤ 15% criteria was met and in 652 (59.9%) peaks the CV ≤ 20% was met.

A similar pattern was seen in the calculations of the CVs across the QCs in data set B (see Table 2). In addition, for this dataset we considered how many peaks met the FDA criteria out of those which always had a positive intensity across the 114 samples. It was noted that a very high percentage of these ever present peaks met the criteria when the start-up QCs were removed (85.1% at the ≤15% and 93.1% at the ≤20% criteria respectively). This procedure was used in the context of examining analytical repeatability only. Using peaks that had a positive intensity across all the samples is not recommended for biomarker detection for there may be examples of useful biomarkers which are not present in all samples (e.g., only present in a disease group and not present in controls or *vice versa*).

These results demonstrate again, if such a demonstration was still necessary, the clear need both to condition the LC–MS system with injections of the matrix under investigation prior to

Table 1
Percentage of detected peaks fulfilling the 15% and 20% CV criteria

QCs considered	Peaks considered	Number of peaks	Number (%) of peaks with % CV ≤ 15%	Number (%) of peaks with % CV ≤ 20%
All QCs	All peaks	3414[a]	666 (19.5%)	1054 (30.9%)
	Ever present peaks in the QCs	803	285 (35.5%)	424 (52.8%)
QCs excluding the 5 start-up QCs	All peaks	3414[a]	974 (28.5%)	1356 (39.7%)
	Ever present peaks in the QCs	1089	495 (45.5%)	652 (59.9%)

Dataset derived by data extraction of the QCs only (data set A).

[a] Peaks were only considered to have met the criteria had they shown a positive intensity in at least 2 of the QCs and met the 15% or 20% CV criteria. This was to avoid peaks with a positive intensity in only one of the QCs being declared acceptable, as they would have a % CV of 0.

Table 2
Percentage of detected peaks fulfilling the 15% and 20% CV criteria

QCs considered	Peaks considered	Number of peaks	Number (%) of peaks with % CV ≤ 15%	Number (%) of peaks with % CV ≤ 20%
All 16 QCs	All peaks	5522	580 (10.5%)[a]	965 (17.5%)
	Ever present peaks in the QCs	971	303 (31.2%)	509 (52.4%)
	Ever present peaks in the 114 samples	101	49 (48.5%)	67 (66.3%)
QCs excluding the 5 start-up QCs	All peaks	5522	905 (16.4%)[a]	1296 (23.5%)
	Ever present peaks in the QCs	1278	582 (45.5%)	802 (62.8%)
	Ever present peaks in the 114 samples	101	86 (85.1%)	94 (93.1%)

Dataset derived from QCs and test samples.
[a] Peaks were only considered to have met the criteria had they shown a positive intensity in at least 2 of the QCs and met the 15% or 20% CV criteria. This was to avoid peaks with a positive intensity in only one of the QCs being declared acceptable, as they would have a % CV of 0.

the analysis of study samples, and the need to remove the data from these conditioning QC samples from the post run data analysis. Otherwise the data generated from the very start of the run, which does not appear typical or reliable, will adversely distort the subsequent data analysis. In addition the level of repeatability observed in the present study was similar to that of the previous study (LC–MS see Ref. [7]) where different instrumentation and software (analytical console and data treatment software) was used.

The above consideration of the CVs in the peak intensities across the QC samples with respect to the FDA recommendations was performed on the raw data from the peaks as these were collected from MarkerLynx. It may prove necessary to transform the raw intensity data prior to carrying out formal statistical analysis to take account of any distributional assumptions that the statistical methods might depend on.

3.4. Repeatability in relation to signal intensity

We have previously found it useful to examine the variability in the LC–MS data derived from this type of sample in relation to the signal intensities of the peaks detected by the software. A similar exercise was performed for the UPLC–MS data obtained here, applying the criteria of the CV to be less than 15% or 20% and the peak to be non-zero in at least 2 of the QCs. Fig. 5 shows the peak intensity distribution in the QCs. Using the peak list generated from dataset A it was found that 46% of the peaks had arbitrary areas ranging from 0 to 1, another 49% had peak areas ranging from 1 to 10 and only 5% of the total peaks had areas over 10. As we observed in the previous LC–MS study [7], and as would be expected, the peaks of lowest intensity showed the greatest variability. The percentage of the acceptable peaks is also depicted in Fig. 5. For instance

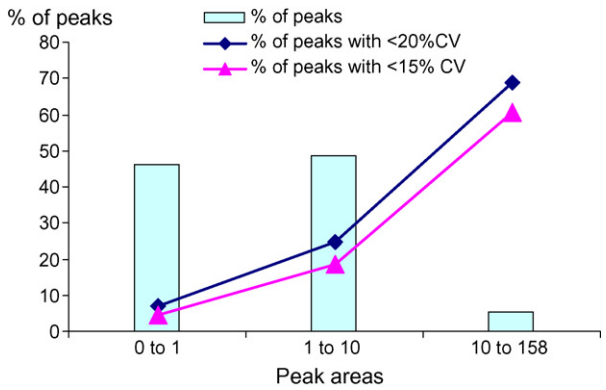

Fig. 5. Peak distribution based on peak areas (average peak areas of QCs, first 5 QCs are not included) and percentage of acceptable peaks according to the proposed criteria.

from the low intensity peaks which ranged in area from 0 to 1, only 7% (peaks with CV < 20%) were acceptable. This means that the main source of variability in the data is caused by the inconsistent detection of low intensity signals by the peak picking software. This is obviously a significant issue in this type of work since the statistical results very much depend on the quality of the extracted dataset (peak tables) generated by the peak picking/alignment software.

3.5. Impact of missing values

A problem encountered in MS data treatment is the occurrence of zero values in the data matrix. This can be either due to the absence of the peak (ion) in certain samples or due to the inability of the mass spectrometer to detect it (perhaps as a result of low intensity or ion suppression), or finally due to the inability of the algorithm to identify the peak. Processing in multivariate statistics a data matrix with several zero points may be problematic, although PCA can tolerate up to 50% zero points [16]. In that respect we have investigated the distribution of the zero signals (missing values) in the UPLC–MS data set A (including the QCs only). The majority of the missing values were observed (as it was indeed expected) for peaks of low intensity with mean areas between 0 and 1. When data analysis was performed considering only peaks with non-zero values, the repeatability was enhanced more than two fold for low intensity peaks (area 0–1). Next we investigated the distribution of the non-zero peaks in the 114 test samples to see if these were associated with particular areas of the chromatogram. If the distribution were non-random it might indicate regions that were for instance more subject to ion suppression, or just generally more noisy and therefore liable to be less reproducible. Fig. 6 depicts the distribution of the non-zero peaks for all of the test sample in m/z-retention time space over the first 6.6 min of the run where most of the peaks are found. The software classifies the 114 samples in ten groups which are illustrated as 10 different blocks. The numbers on top of each block in the figure. indicate the range of numbers of samples where the non-zero peaks are found. The plot does not seem to suggest that peaks which are well defined (i.e. have a non-zero intensity for all samples) differ greatly in their distribution in mass-retention time space compared with peaks which are less well defined (i.e., have a zero intensity for most samples). This indicates that there is no obvious dominating trend in the data set in connection to the missing values.

Whilst differences in samples, instrumentation and software make a direct comparison of the present UPLC–MS study with our previous LC–MS investigations [7] difficult, a number of general points are clear. For both UPLC and HPLC the use of a pooled QC, or similar matrix sample, is essential to condition the system. For both HPLC and UPLC, given retention time and detector stability, the major source of variability resulted from signal intensity, with

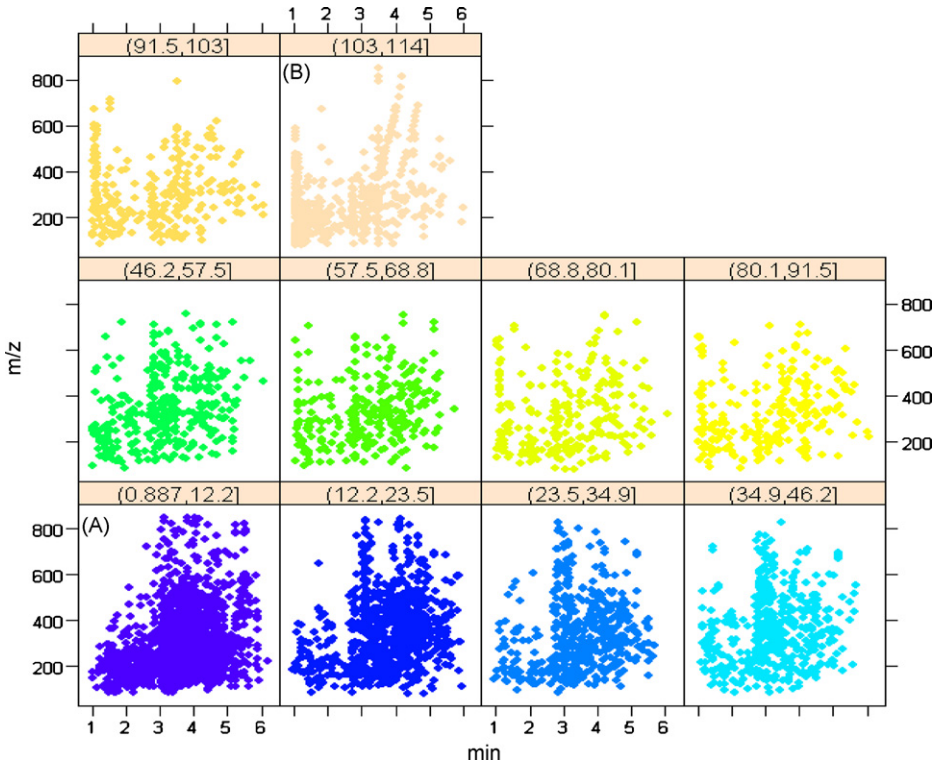

Fig. 6. Distribution of non-zero peaks in m/z versus retention time plot. The numbers on top of each block indicate the range of number of samples where the non-zero peaks are found. For example in box labelled (A) peaks with non-zero intensity for only a few of the samples (0–12 samples) are depicted whereas in the box labelled (B) peaks with non-zero intensity for nearly all samples (103–114 samples) are depicted.

low intensity ions showing more variability than more abundant components.

4. Conclusions

Based upon the assessment of QC samples, following the conditioning of the system, UPLC–MS provided an effective and repeatable method for the global metabolite profiling of human urine. UPLC provided excellent retention time stability, with variability lower than 1% over the 29 h required for the analysis of these samples. The use of appropriate QC samples, prepared from the study samples, enabled the performance of the system to be effectively monitored for drift in important analytical parameters such as retention time, detector response and mass accuracy. The main source of variability was the low intensity signals whilst the more abundant components were generally seen to give good repeatability according to the applied criteria for the acceptance of analytical data (CV less than 15%).

Acknowledgements

This work was carried out in the context of an EU committee Transfer of Knowledge Industry-Academia partnership grant (TOK-IAP 29640) between the Aristotle University Thessaloniki Greece and AstraZeneca UK.

References

[1] J.K. Nicholson, J.C. Lindon, E. Holmes, Xenobiotica 29 (1999) 1181.
[2] J.K. Nicholson, J. Connelly, J.C. Lindon, E. Holmes, Nat. Rev. Drug Discov. 1 (2002) 253.
[3] E.M. Lenz, I.D. Wilson, J. Proteome Res. 6 (2007) 443.
[4] R.S. Plumb, C.L. Stumpf, M.V. Gorenstein, J.M. Castro-Perez, G.J. Dear, M. Anthony, B.C. Sweatman, S.C. Connor, J.N. Haselden, Rapid Commun. Mass Spectrom. 16 (2002) 1991.
[5] I.D. Wilson, R. Plumb, J. Granger, H. Major, R. Williams, E.M. Lenz, J. Chromatogr. B 817 (2005) 67.
[6] G. Theodoridis, H. Gika, I.D. Wilson, TRAC Trends Anal. Chem. 27 (2008) 251.
[7] H.G. Gika, G.A. Theodoridis, J.E. Wingate, I.D. Wilson, J. Proteome Res. 6 (2007) 3291.
[8] T. Sangster, J. Wingate, L. Burton, F. Teichert, I.D. Wilson, Rapid. Commun. Mass Spectrom. 21 (2007) 2965.
[9] M.-E. Dumas, E.C. Maibaum, C. Teague, H. Ueshima, B. Zhou, J.C. Lindon, J.K. Nicholson, J. Stamler, P. Elliot, Q. Chan, E. Holmes, Anal. Chem. 78 (2006) 2199.
[10] T. Sangster, H. Major, R. Plumb, A.J. Wilson, I.D. Wilson, Analyst 131 (2006) 1075.
[11] I.D. Wilson, J.K. Nicholson, J. Castro-Perez, J.H. Granger, K. Johnson, B.W. Smith, R. Plumb, J. Proteome Res. 4 (2005) 591.
[12] R.S. Plumb, J.H. Granger, C.L. Stumpf, K. Johnson, B.W. Smith, S. Gaulitz, I.D. Wilson, J. Castro-Perez, Analyst 130 (2005) 844.
[13] J.H. Granger, R. Williams, E.M. Lenz, R.S. Plumb, C.L. Stumpf, I.D. Wilson, Rapid Commun. Mass Spectrom. 21 (2007) 2030.
[14] R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL: http://www.R-project.org, 2007.
[15] FDA Guidance for Industry, Bioanalytical Method Validation, Food and Drug Administration: A Guidance, Centre for Drug Valuation and Research (CDER), 2001.
[16] J. Trygg, E. Holmes, T. Lundsted, J. Proteome Res. 6 (2007) 469.