Quality Control in LC-MS Metabonomics Analysis. Approaches, issues, guidelines

H. Gika[1,2], G. Theodoridis[1,2], J. Wingate[3], I. Wilson[1]

[1]AstraZeneca DMPK, Alderley Park, UK [2]Aristotle University of Thessaloniki, Greece [3]Applied Biosystems|MDS Sciex, Canada

Introduction

Self evidently research in areas supporting "systems biology" such as genomics, proteomics and metabonomics are critically dependant on the generation of sound analytical data. Metabolic phenotyping using LC-MS-based methods is currently at a relatively early stage of development and approaches to ensuring data quality are still developing. As part of studies on the application of LC-MS in metabonomics the **within day reproducibility** of LC-MS, with both +ve and -ve electrospray ionisation (ESI), has been investigated using a standard "quality control" (QC) sample.

Based on these findings an analytical protocol for the metabonomic analysis of human urine has been developed with proposed acceptance criteria based on a step-by-step assessment of the data.

Overall, by careful monitoring of the QC data it is possible to demonstrate that the "within-day" reproducibility of LC-MS is sufficient to assure data quality in global metabolic profiling applications.

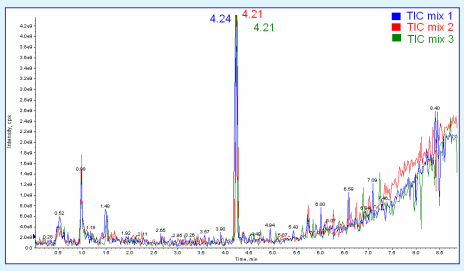
Experimental

Samples: Mid-stream urine samples: 30 male, 30 female healthy volunteers.
Centrifuge 13000 rpm, dilution 1:4 with 0.1 % aqueous formic acid. Standard mix sol. for initial "system suitability". A "quality control" (QC) sample: "pooled" urine to provide a representative "mean" sample. Run duration: 17 hours.

LC/ESI-MS analysis : Symmetry C18, 3.5um (2.1x100 mm) column at 40 (±0.2)°C, gradient acetonitrile (B) over 0.1% formic acid in water (A), flow rate (0-0.5 min) 100 % A, (0.5-4 min) 80% A, (4-8 min) 5% A, (8-9) 5% A. 400 uL/min. Injection 10 uL
Hybrid Triple Quadrupole Linear Ion Trap, 4000 QTRAP®, (Applied Biosystems|MDS Sciex). (+),(-) ESI TurboIonSpray® at 350 °C, ±4500 V separately. 12 min EMS 100 to 850 m/z (1000 amu/s), Ion Trap: Dynamic Fill Time mode.
"QC" sample injected 4 times at the beginning of the run and then every ten samples.

Data analysis : MarkerView™ software 1.1.0.1, Simca P 11 Umetrics

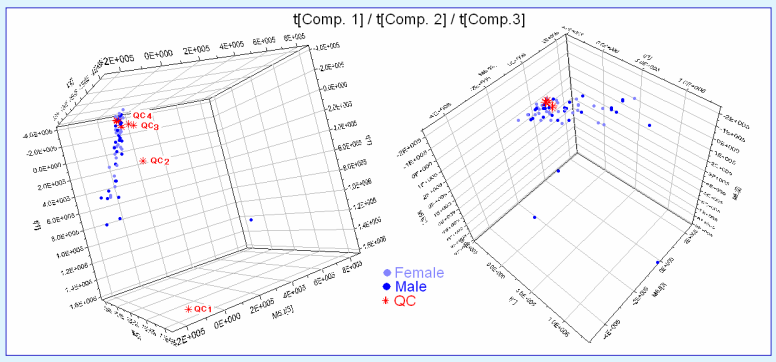
Results



The **test mix** of pure standards provides an initial screen : **Rapid visual examination of the system performance** by overlay of the 3 runs at the **beginning**, **middle and end of the batch**. Any dramatic change, in either chromatographic or detector performance would have been evident in the form of changes in retention time, peak shape, peak height/area and mass accuracy etc.

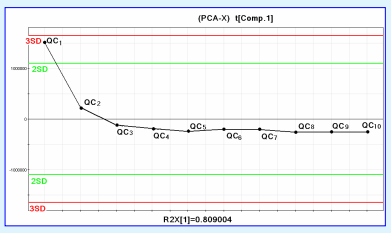
A **"biological QC" approach** is used to investigate the stability of the analytical methodology within a run:

The tighter the **clustering of the QC samples** in the scores plot the more reliable the data: differences between the samples from different individuals more likely to reflect real differences in metabolite profiles rather than analytical variation.



The first few injections on the system are not representative and are excluded

PCA scores plots



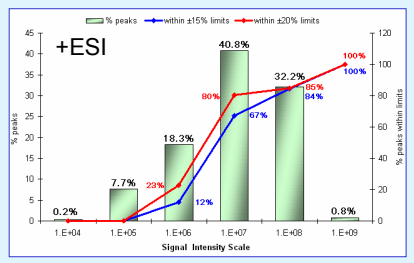
The time series properties of PCA 1st component shows how the QC samples behaved as the run (+ESI) progressed

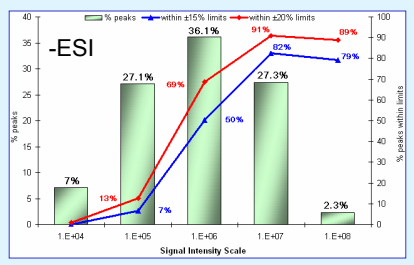
Investigation of **variability** of peaks using a range of **acceptance criteria**.

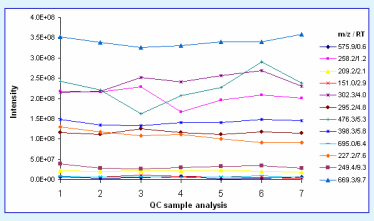
The **distribution of peaks** found by the software (noise threshold set at 10^4) together with the % acceptable peaks in QC data obtained determined for two **critical limits** (±15 and ±20 of the average response) were examined.

The degree of variability is not constant over the range of signal intensities detected.

The bulk of unacceptable results are concentrated in the minority of low intensity peaks.

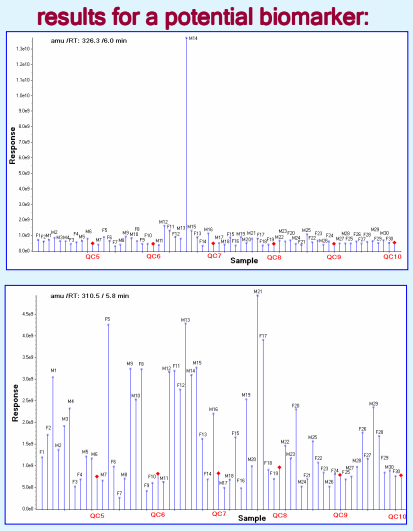






Responses of 15 randomly selected peaks in the QC samples: **no obvious systematic variation** or drift in the signal over the period of the analysis

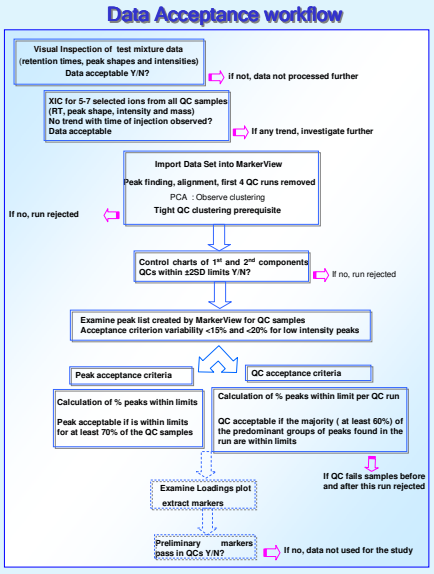
Use of QC data to "validate" the results for a potential biomarker:



Plot profiles for significant ions

QC data on specific ions confirm that the differences in the population are real.

Data Acceptance workflow



Suggested workflow for accepting LC-MS-generated metabonomics data as fit for in depth, statistical analysis as part of biomarker discovery: failure to pass any of these stages should trigger a re-analysis of the sample set.

Conclusions

The results showed that the system's stability needs to be controlled by QC runs, initial QCs are not representative, and should be excluded.

Variability in both mass and retention time were not observed to be high (no more than 0.2 amu and 0.08 min respectively for randomly selected peaks).

Reproducibility was critically dependant on signal intensity after chromatography was equilibrated. In particular the variability of lower intensity peaks was significantly higher than those of higher intensity.

References

1. Sangster, T.; Major, H.; Plumb, R.; A.J. Wilson, A.J.; Wilson, I.D. *Analyst,* **2006**, 131, 1075-1078
2. FDA Guidance for Industry, Bioanalytical method Validation, Food and Drug Administration, centre for Drug valuation and Research (CDER), May 2001.a guidance
3. J.C. Lindon et al , Nature Biotechnology, 2005, **23** 833-838
4. Castle, A.L.; Fiehn, O.; Kaddurah-Daouk, R.; Lindon, J.C.; *Briefings in Bioinformatics,* **2006,** *7*, 159-165.

Acknowledgments

EU TOK-IAP 2960 Grant for funding