# Speeding up Mass Spectrometry

## Determining the Effect of Mass Spectrometry Scan Speeds on Data Reproducibility for Metabolomics Analyses

**Recently, there has been a trend towards faster chromatography for qualitative mass spectrometry applications, resulting in narrower chromatographic peak widths. As a result, slow-scanning mass spectrometers have been challenged with trying to acquire sufficient data to define these narrow peaks adequately. In this work, we sought to evaluate the mass spectrometer scan speed required to provide accurate, reproducible data when coupled to high speed LC systems in a metabolomics workflow.**


In non-targeted metabolomics experiments, scientists try to find real biological variation between groups of samples. To do this, researchers find and identify compounds that are either uniquely present or are just specifically up-/down-regulated in a particular group of samples, so-called biomarkers. The biomarkers are often quite minor and can easily be masked by other differences between the samples that are unrelated to biology, and could be artefacts related to sample handling, data acquisition or data processing.

To find biomarkers, researchers often perform numerous complex statistical analyses of the data to differentiate one sample "group" from another. The underlying basis of these statistical approaches is that the variation within a group will be much less than the variation between different groups. However, underlying this approach is the assumption that the data within each sample group is highly reproducible so as not to introduce artificial biases.

This work will focus on the requirements for data collection and analysis by evaluating the compatibility of various mass spectrometry scan speeds with fast chromatography methods. We will evaluate the ability of Principle Component Analysis (PCA) to group data accurately, and the corresponding reproducibility and confidence in the results, as mass spectrometry scan speed is changed.

### Materials and Methods

Urine samples were collected in the morning from three different male and three different female subjects and frozen until analysis. A 1:10 dilution in mobile phase was done prior to analysis. Samples were divided into nine aliquots, resulting in nine replicate injections from each of six different subjects.

Chromatographic separation was performed using a PE 100 series autosampler and micro pumps. A gradient from 95 % to 5 % aqueous at 250 µl/min through a BetaBasic C18 50 x 2.1 mm, 3 µm column was done over a three minute analysis. A QSTAR Elite System (Applied Biosystems/MDS Sciex) was used for collection of MS data over a 1000 amu mass range and all detection was performed in positive ion mode using TurbolonSpray Source. Dynamic AutoCalibration was used to maintain mass accuracy throughout the data collection. Data processing was performed using MarkerView Software.

MS data were acquired at several different scan speeds, ranging from 1–20 spectra/s. Data analysis was carried out using MarkerView Software to determine the effect of MS scan speed and chromatographic peak width, and to determine the required number of points across a peak to ensure accurate, reproducible and interpretable data.

### High Throughput Chromatography

Precise, accurate and reproducible gradients are critical for LC/MS/MS analyses, especially when comparing results from sample-to-sample, run-to-run, and lab-to-lab. This is especially true for metabolomics analysis where it is critical to remove any possible variation due to chromatographic irregularities or imprecision prior to statistical analysis. For this analysis, a three minute runtime with a 60 second gradient (95 % to 5 % aqueous) was employed. This chromatography resulted in peak widths in the order of 4 seconds at the base, and although this can be considered high throughput, there is room to improve; one to two second wide peaks are possible and will be explored in our future work.

### High Speed Accurate Mass Spectrometry

As the chromatographic peak widths narrow, obtaining more spectral scans per peak should increase the confidence in statistical data used to evaluate large populations of samples. The system can easily acquire data at a rate of 20 spectra/s. This makes it possible to obtain sufficient information across a chromatographic peak that may only be four seconds wide at the base. With further reductions in peak width and intensity, scan speed becomes even more critical to ensure accurate results. As mentioned previously, the biomarkers are likely to be very minor compounds and have correspondingly small peaks.

### Results

MS scan speed and fast chromatography – how does this affect a metabolomics analysis?

Clearly, acquiring data at a faster rate gives more data points across a chromatographic peak. But is it necessary – does this affect the results? To investigate further the effect of scan speed on the ability to separate the samples

into their respective groupings and determine the accuracy and interpretability of the data, MarkerView Software was used to perform PCA.

The software is designed to allow the data from several samples to be compared so that differences can be identified. The program uses multivariate analysis (MVA) techniques to compare the samples and provides both supervised and unsupervised methods.

For this study, an unsupervised method, PCA, was used to review the collected data and analyse the effect of MS scan speed on the groupings found. Figure 1 shows the scores plot for data collected at 1 and 20 spectra/s, respectively. For the 1 spectrum/s data, it appears that there are only three different groups within the data. For the 20 spectra/s data, however, six distinct groups are clearly seen, easily separating the individuals from one another. Once the groups have been assigned different coloured points, as shown in figure 2, the challenge in interpreting the 1 spectrum/s data becomes clearer. EF and BKM seem to be in one group while BKF, AW and GI seem to form a second group. Only the JW group is distinct. Interestingly, AW and GI are samples from males while BKF is a sample from a female. So at 1 spectrum/s, even the most basic differentiation between males and females is not possible. To try to separate the 1 spectrum/s data further into correct groups, we looked at PC1 vs PC3 and PC2 vs PC3. Neither of these alternate ways of viewing the data provided separation of the samples into appropriate groups.

In addition to the ability to acquire MS and MS/MS spectra rapidly, another advantage of the QSTAR Elite for metabolomics analyses is its high resolution and accurate mass capabilities. Unlike FTMS-like instruments, the resolution and mass accuracy of the system are not affected by scan speed. Whether data is acquired at 1 spectrum/s or 20 spectra/s, the device can generate data at better than 12,000 resolution and sub-2 ppm mass accuracy (data not shown due to space limitations). These characteristics, coupled with fast data acquisition are critical in allowing potential biomarkers to found and identified.

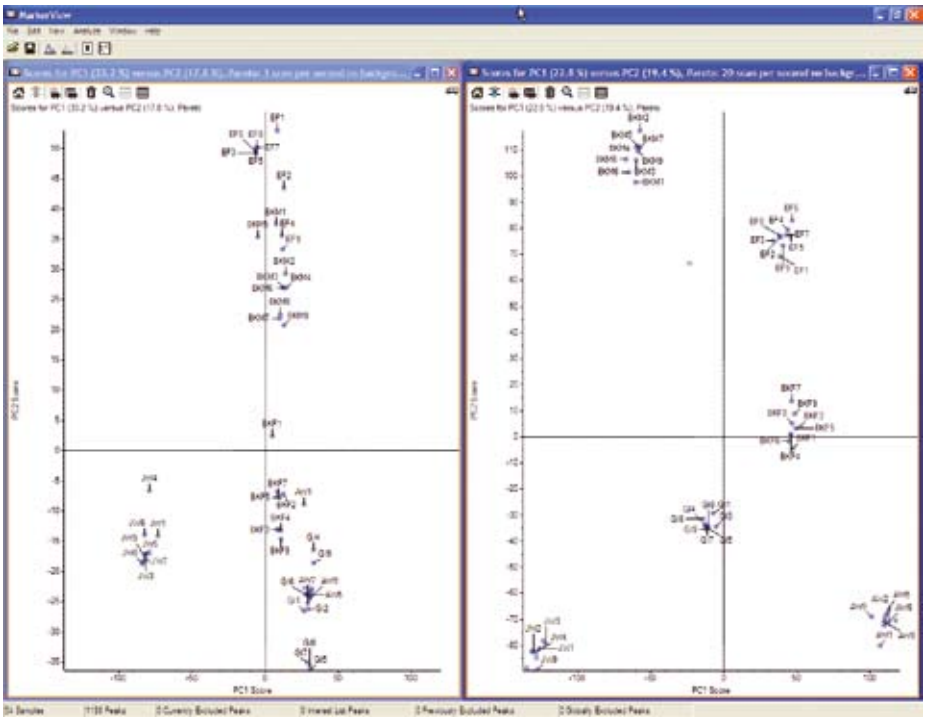

**Fig. 1: A) Scores plot for the 1 spectrum/sec data acquisition.**
**B) Scores plot for the 20 spectra/sec data acquisition. For the 1 spectrum/sec, it appears that there are only 3 different groups within the data. For the 20 spectra/sec data, however, 6 distinct groups are clearly seen, easily separating the individuals from one another.**

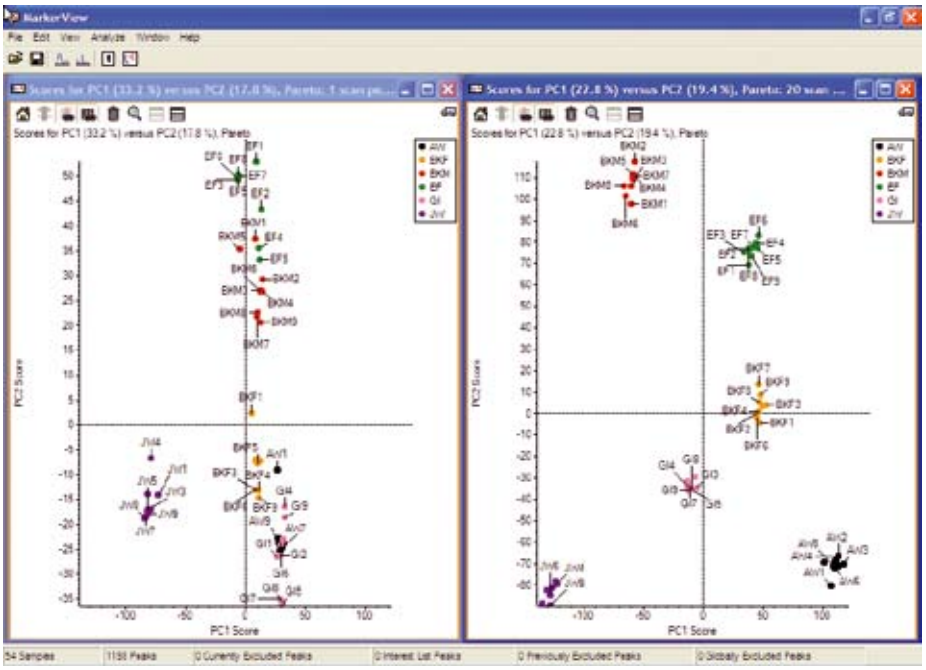

**Fig. 2: A) Scores plot for the 1 spectrum/sec data colour-coded to highlight groupings better.**
**B) Scores plot for the 20 spectra/sec data colour-coded for comparison. The challenge in interpreting the 1 spectrum/sec becomes clearer compared to figure 1. However, the amount of data collected over a chromatographic peak at 20 spectra/sec has provided better quality data, allowing for easy differentiation of the groups of samples.**

## Summary

- 1–3 spectra/s data acquisition rates that are typical of FTMS-type instruments are not sufficient to allow groups to be distinguished by PCA when coupled to high speed chromatography
- fast scanning of the system allows sufficient data collection across very narrow chromatographic peaks for a metabolomics analysis
- resolution and mass accuracy are not affected by scan speed – data can be collected at a rate of 20 spectra/s at better than 12,000 resolution and sub-2 ppm mass accuracy.

## Future Work

- Data acquisition using other LC systems to achieve narrower peak widths, in the order of 1–2 seconds at the peak base
- acquisition of MS and MS/MS data simultaneously using Information Dependent Acquisition and Dynamic Background Subtraction to evaluate whether data can be correctly grouped and interpreted when some scan time is dedicated to MS/MS acquisition.

*Gary Impey*
*Applied Biosystems/MDS Sciex*
*Concord, ON Canada*
*www.appliedbiosystems.com*