Combining real time and post-acquisition quality control (QC) for metabolomics workflows

Aiko Barsch, Patrick Groos, Nikolas Kessler, Matthias Szesny, Sven W. Meyer, Ilmari Krebs, Heiko Neuweger, Matthew R. Lewis

Bruker Daltonics GmbH & Co. KG, Bremen, Germany

Introduction

Global profiling of metabolites in complex biological samples benefits from mass spectrometry-based techniques coupled to chromatographic separation and/or ion mobility spectrometry for enhanced selectivity. Together, these hyphenated systems produce complex multidimensional data for hundreds to thousands of metabolites across similar numbers of samples, with both random and systematic error possible in each dimension of measurement. The resulting data is complex, and consequently, evaluation of data quality is challenging (Figure 2). To help tackle this challenge, we outline a suite of quality control tools for the assessment of metabolomics data (Figure 1), inclusive of real-time quality of monitoring QC analytes, post-acquisition review, and data correction of complete feature-sets.

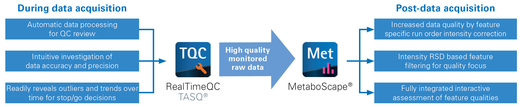

Figure 1. Workflow for combining real-time and post-acquisition QC in metabolomics studies.

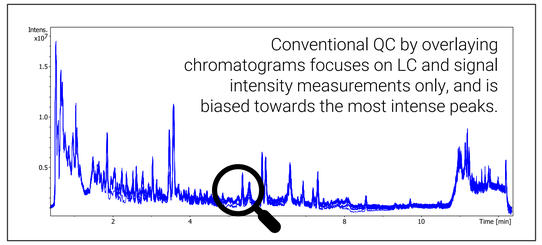

Figure 2. Overlay of Total Ion Chromatograms of every tenth sample analysis (designated "QC sample") from 300 consecutive analysis of human urine NIST SRM 3672, representing conventional quality review.

Methods

Three-hundred injections of human urine (NIST SRM 3672) were made using an Agilent 1290 Infinity II UHPLC and Bruker Impact II VIP mass spectrometer, and every 10th injection was labelled as a "QC sample" to mimic the classical QC design of metabolomics studies. Data quality was monitored in parallel with acquisition by inspection of endogenous metabolites previously chosen to represent the broader metabolome ("QC analytes") using the RealTimeQC module of TASQ® 2023b software. Following data acquisition, MetaboScape® was used to perform post-acquisition retention time alignment, mass recalibration, and compensation for run-order effects in peak intensity measurements using a soft LOESS correction curve based on feature-specific intensities in designated QC samples.

Results

- Visualization of data acquisition in real-time allows the inspection of random and systematic variation across all key measurements (LC peak area, peak intensity, retention time, *m/z*, and isotope pattern fidelity/mSigma) during data acquisition (Figures 3 and 4).
- Automated flagging of measurement outliers facilitates rapid and informed decision making on whether to intervene in an ongoing analysis (Figure 4).
- On the basis of RealTimeQC feedback, the sample analysis was allowed to finish uninterrupted, with excellent measurement performance results tabulated across all 300 replicate analyses (Figure 5).

Figure 3. TASQ RealTimeQC visualization of *m/z* accuracy and isotope pattern fidelity (mSigma) for the QC analyte isoleucine across the full analysis of study (green) and QC (pink) samples. The results demonstrate the absence of systematic deviation (e.g. run-order effects) while highlighting infrequent outliers owing to random variation in a manner that is otherwise impractical with manual data quality review.

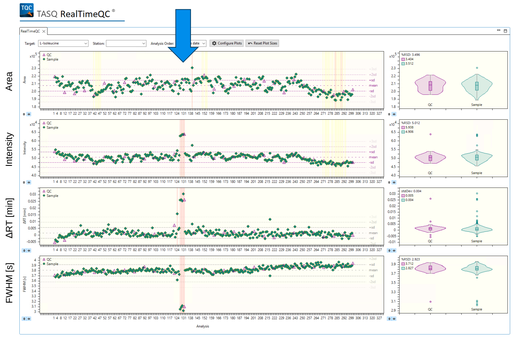

Figure 4. Fast decision making to stop or continue analysis is enabled by comprehensive measurement visualization of the QC analyte isoleucine. Despite a QC sample being automatically flagged as an outlier in peak intensity, retention time and peak width (FWHM), it is also shown to maintain a consistent peak area. On this basis, the analysis is allowed to continue, and subsequently all measurements return to being within expected limits of variation. The study concludes with 3.4% RSD in peak area across all QC samples.

A

	Peak Area [% RSD]	StdDev Retention time error [s]	StdDev mass error [ppm]	mSigma (median score of max 1000)
Butyryl-L-carnitine	7.63	0.063	0.483	5.12
Glutamine	3.38	0.016	0.625	3.04
Isoleucine	3.50	0.240	0.470	1.17
Isovalerylcarnitine	9.86	0.067	0.486	2.75
Leucine	2.96	0.077	0.496	1.23
O-Desmethyl-cis-tramadol	7.69	0.089	0.490	3.05
Phenylalanine	4.39	0.069	0.442	2.90
Uric acid	3.67	0.051	0.440	8.33

B

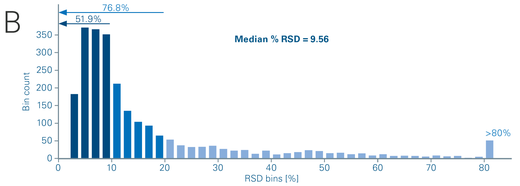

Figure 5. Summary statistics compiled from RealTimeQC data illustrate the excellent overall data quality of all tracked QC analytes across all 300 injections (A). These results are broadly representative of the precision observed across the entire detected metabolome as illustrated by post-acquisition analysis (histogram) of the peak area precision for all MetaboScape-extracted features (B).

- Where necessary, MetaboScape software allows for the comprehensive review of global profiling data quality and feature-specific correction of run-order-based effects in quantitation (Figure 6).
- Finally, more than 350 metabolites were automatically annotated using a Target List of metabolites described to be present in human urine (https://hmdb.ca/). Each metabolite in the dataset met the annotation quality scoring criteria of <1ppm mass accuracy and isotope pattern fidelity (mSigma) <20.

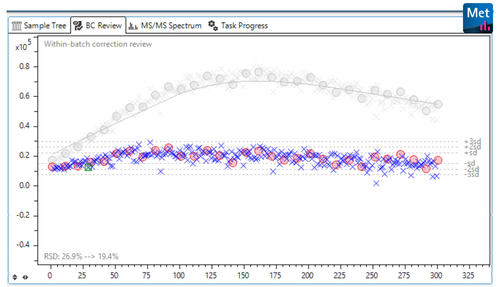

Figure 6. Within-batch correction in MetaboScape for accurate quantitation of a feature showing run-order based intensity drift.

Conclusions

- Augmentation of a traditional post-acquisition QC workflow for metabolomics with real-time monitoring of QC analytes improves control of data quality by enabling comprehensive measurement monitoring and informed decision making during the analysis.
- Potential applications include the early detection of instrument malfunctions, allowing for quick corrective action to be taken to avoid compromised results and lost sample.
- Post-acquisition data visualisation and correction of run-order-based effects is still a necessary component of QC in LC-MS-based global profiling studies.

Metabolomics