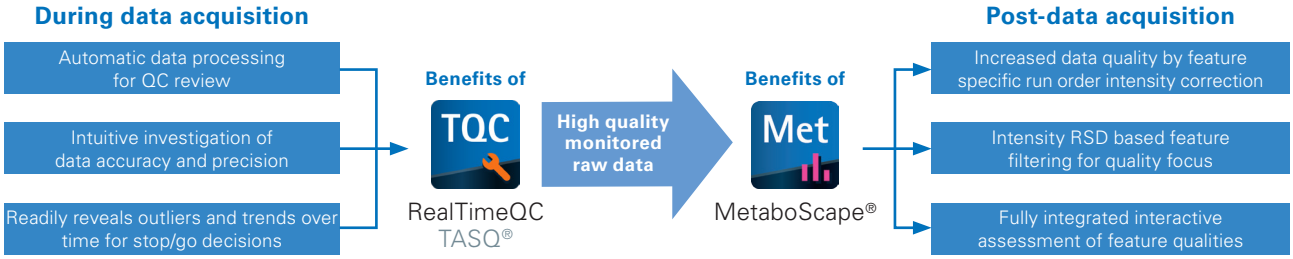
Real-time and post-acquisition quality control (QC) in metabolic phenotyping studies

Abstract

- Metabolic profiling of biofluids using hyphenated mass spectrometry-based techniques including LC-MS and LC-TIMS-MS is a workflow of prime importance for biomarker research.
- High precision measurements are a prerequisite for deriving meaningful conclusions from multidimensional data.
- Both random and systematic errors are possible in each dimension of measurement, creating a need for thorough quality control (QC) monitoring and post-acquisition data correction of complete feature-sets.
- The RealTimeQC module of TASQ® 2023b software runs in parallel with data acquisition, providing detailed insight to instrument performance and data quality based on user-defined QC analytes.
- Following data acquisition, MetaboScape® enables correction of feature-specific run order effects and facilitates in-depth analysis of individual feature quality.

Keywords:
Metabolomics, Quality control, Real-time QC, MetaboScape, TASQ



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

Modern methods for the global profiling of metabolites in complex biological samples utilize powerful bioanalytical systems that combine high resolution mass spectrometry (HRMS) with ultra-high performance liquid chromatography (UHPLC) and/or trapped ion mobility spectrometry (TIMS) for highly selective metabolite measurement. Together, these systems produce complex multidimensional data for hundreds to thousands of metabolites across similar numbers of samples, with both random and systematic errors possible in each dimension of measurement. The resulting data is complex, and consequently, evaluation of the data quality is challenging. To make matters worse, as outlined by Broadhurst *et al.* in 2018 [1], compared to targeted assays there are no QC acceptance criteria for all the detected metabolites in untargeted metabolomics. For these workflows, the repeated analysis of pooled QC samples can provide relative measures of precision and be used to correct for systematic measurement error including run order-based drift in signal intensity. However, such analysis is conventionally performed after data acquisition is completed, by which time any substantial measurement drifts or sudden deviations may be beyond remedy, potentially compromising the analysis. For these reasons, researchers may wish to more closely monitor the quality of their profiling analyses in real-time (i.e. in parallel with data acquisition). However, automated tools to support meaningful real-time investigation of complex multidimensional data are severely lacking in the field.

To help tackle these challenges, we introduce here the RealTimeQC module of TASQ 2023b software for in-depth analysis of data quality during acquisition. We further demonstrate Bruker's complete QC workflow which includes MetaboScape for post-acquisition review and correction of profiling data.

Methods

Six vials were prepared with a 1:5 dilution of standard reference material urine (NIST SRM 3672) in ultra-pure water and 50 injections were performed per vial. This resulted in a total of 300 injections of the diluted sample acquired with an Agilent 1290 Infinity II UHPLC and Bruker impact II VIP QTOF high-resolution mass spectrometer. HRMS acquisition was performed in positive ionization mode featuring a VIP-HESI source (details see Table 1). A 15-minute reversed-phase UHPLC gradient was adapted from the National Phenome Center open LC-MS platform [2, 3] (Details see Table 2).

Every tenth analysis was labeled (and is herein referred to) as a QC sample, mimicking the conventional QC format of global profiling studies. Data quality was monitored in parallel with acquisition by inspection of endogenous metabolites chosen to represent the broader metabolome ("QC analytes") using the RealTimeQC module of TASQ 2023b software. Following data acquisition, MetaboScape 2023b 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.

Table 1
MS acquisition parameters

MS	impact II VIP	
Source	**VIP-HESI source**	
	End Plate Offset	500 V
	Capillary	4500 V
	Nebulizer	2.0 Bar
	Dry Gas	8.0 L/min
	Dry Temp	230°C
	Probe Gas Temp	400°C
	Probe Gas Flow	4.0 L/min
Ionization	Positive ion mode	
Acquisition mode	MS – 5 Hz	
Transfer parameters	Scan range	20-1300 *m/z*
	Funnel 1 RF	200 Vpp
	Funnel 2 RF	200 Vpp
	Hexapole	50 Vpp
	Quadrupole Low Mass	60 *m/z*
	Quadrupole Ion Energy	5 eV
	Collision RF	450 V
	Transfer Time	80 µs
	Pre Pulse Storage	5 µs
Calibration	Automatic internal mass calibration using sodium formate	

Results and discussion

Conventional QC is tedious and ultimately not informative

Interspersed acquisitions of pooled QC samples have become a routine practice in laboratories conducting metabolic profiling experiments. Figure 1 illustrates how QC sample data are typically reviewed during acquisition by overlaying the Total Ion Chromatograms (TIC) of all QC samples in the 300 consecutive human urine analyses. Despite being a common approach for data quality assessment, this form of inspection can be laborious and time consuming. Furthermore, this approach is not deeply informative as investigation is typically biased towards gross chromatographic and signal intensity-based trends and fails to detect possible mass shifts, differences in isotopic fidelity or changes in peak areas of characteristic marker metabolites (QC analytes).

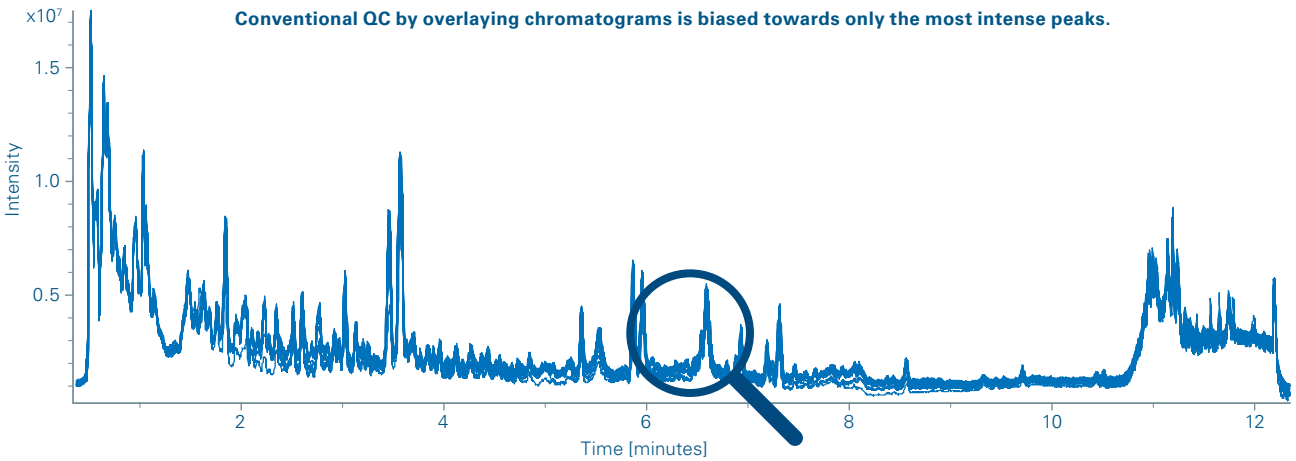


Figure 1

Overlay of Total Ion Chromatograms of every tenth datafile from 300 consecutive analysis of a human urine sample. Conventional QC is performed by manually overlaying chromatograms from the repeated analyses of QC samples. This is often laborious, time consuming, and ultimately lacks information as investigation is typically biased towards only the most intense peaks and lacking the depth and dimensionality produced by modern hyphenated techniques.

RealTimeQC enables real-time monitoring of multidimensional data

QC monitoring can be improved using RealTimeQC by first selecting QC analytes endogenously present in or synthetically added (i.e. internal standards) to samples that act as representative markers for the broader metabolic profile and then monitoring them in depth across the analysis. Here we investigated eight such QC analytes selected to cover a representative range of masses and retention times.

For each newly analysed sample RealTimeQC provides a customizable array of data quality metrics for each QC analyte (see Figures 2 and 3). Here we investigated the peak area, peak width, mass accuracy, retention time accuracy, and isotopic fidelity across all samples. Table 3 highlights the high raw data quality for the eight investigated QC analytes among the 300 total injections. For all QC analytes, the RSD of peak areas is below 10%. The standard deviations for retention time and mass error are below 0.25 seconds and 1 ppm, respectively.

	Peak Area [% RSD]	StdDev Retention time error [sec]	StdDev mass error [ppm]	mSigma (median score of max 1000)
Butyryl-L-carnitine	7.63	0.063	0.483	5.12
Glutamine	3.38	0.01602	0.625	3.04
Isoleucine	3.50	0.240	0.470	1.17
Isovalerylcarnitine	9.86	0.0672	0.486	2.75
Leucine	2.96	0.0768	0.496	1.23
O-Desmethyl-cis-tramadol	7.69	0.0888	0.490	3.05
Phenylalanine	4.39	0.069	0.442	2.90
Uric acid	3.67	0.0507	0.440	8.33

Table 3
Overview of QC results from 300 consecutive injections of a human urine sample.
Following automatic peak detection and integration in TASQ, the Statistics Table provides a high-level overview on data quality for selected QC analytes (e.g. internal standards or endogenous metabolites chosen to represent the broader metabolome).

Additionally, the median score of the mSigma value, which expresses the isotopic fidelity, was very low with a mSigma value <10 of a maximum possible value of 1000, with lower values indicating better isotopic fit.

Despite the overall high quality of the raw data, monitoring in real-time was essential to assessing the nature of potential deviations. RealTimeQC provides intuitive visualization for inspection of random and systematic variation revealing possible trends in data over time and highlighting measurement outliers.

TASQ RealTimeQC

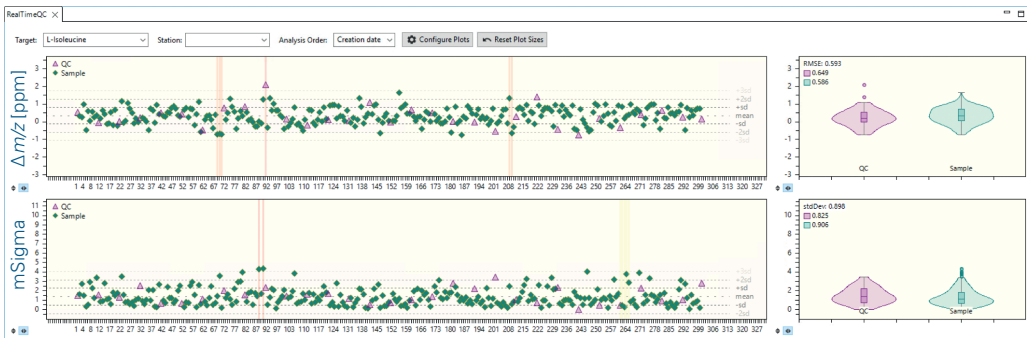


Figure 2
TASQ RealTimeQC visualization – isoleucine
TASQ RealTime QC visualization of mass accuracy and isotopic fidelity for isoleucine in 300 consecutive injections. Both plots reveal stable mass accuracy and isotopic fidelity acquisitions for Isoleucine over 300 injections.

For example, the left part of Figure 2 shows mass error (delta *m/z*) and isotopic fidelity (mSigma) plotted for the selected QC analyte isoleucine across the 300 datafiles acquired showing no cause for concern in the ongoing analysis. In Figure 3A the peak area, peak intensity, delta of retention time (RT) and LC peak Full Width at Half Maximum (FWHM) are plotted for the selected QC analyte, isoleucine. The right part of both figures illustrates the overall variation using violin plots and summary statistics.

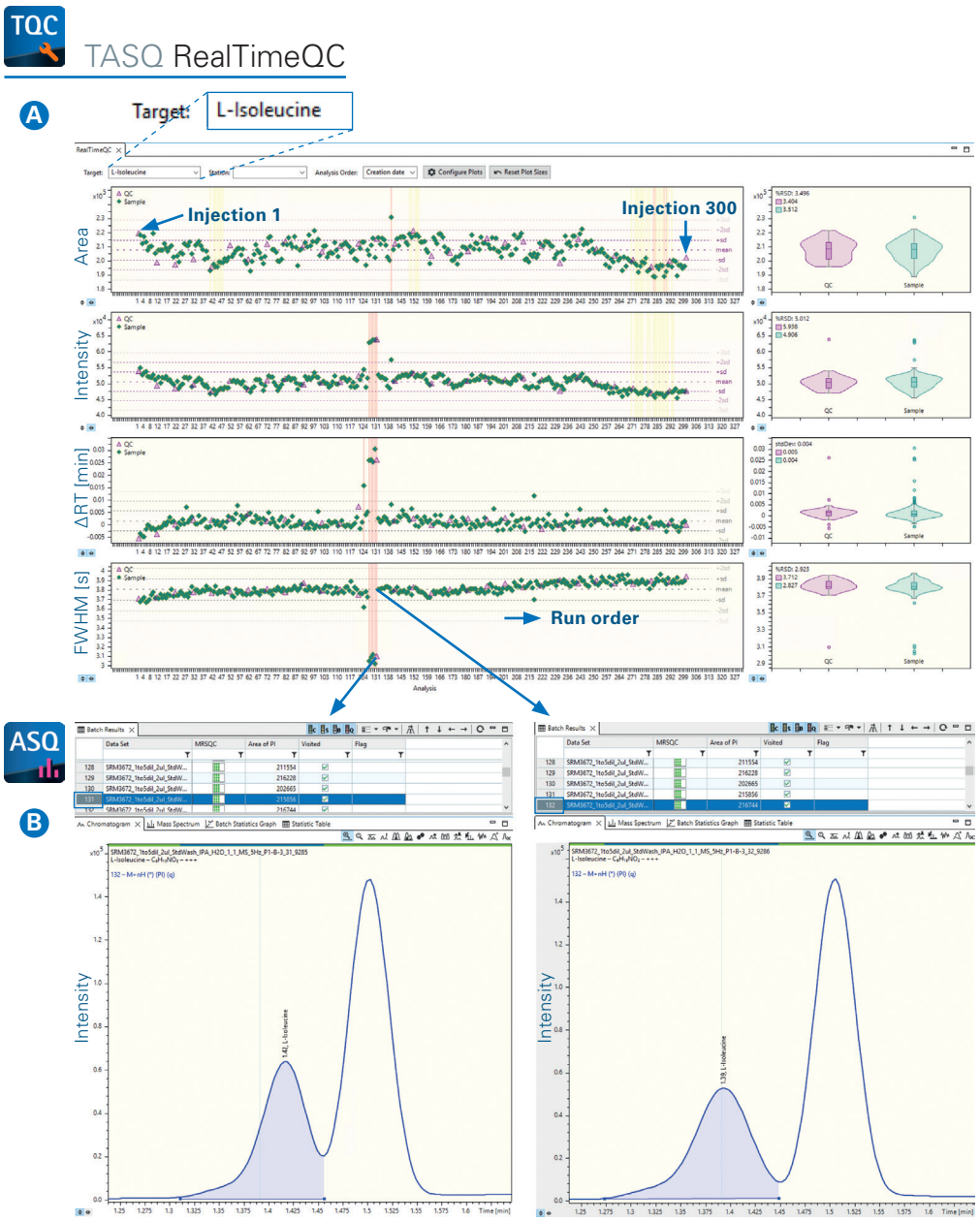


Figure 3
TASQ RealTimeQC allows rapid assessment of QC analyte accuracy and precision across all samples. Ⓐ TASQ RealTimeQC visualization of peak area, peak intensity, difference in retention time, and peak width (FWHM) vs. run order for the QC analyte isoleucine. The scatter and violin plots (top) clearly illustrate stable peak area across three days measurement time with 3.4% and 3.5% relative standard deviation (RSD) across QC and all other samples, respectively. However, the peak intensity (second from top) plots show a change in peak height for several sequential injections in the sequence. Similar differences were also observed in retention time and peak width (FWHM). Ⓑ Analysis of the chromatographic data in TASQ supported the data reported by RealTimeQC, clearly showing a changing peak shape that ultimately had no observable effect on the peak area.

However, at injection 130, RealTimeQC revealed an anomaly in the peak height (intensity), retention time and peak width of isoleucine during the acquisition sequence (Figure 3 A), triggering inspection of isoleucine in TASQ (Figure 3 B). Analysis revealed that while peak width and retention time had changed, the peak area remained constant. Based on this observation the analysis was allowed to proceed, resulting in a subsequent return to normal and excellent overall data quality.

Feature-specific signal correction by MetaboScape

Real-time monitoring of data quality during acquisition ultimately led to successful completion of the 300-sample batch and helped demonstrate the excellent precision of the Agilent 1290 Infinity II UHPLC and Bruker impact II VIP QTOF mass spectrometer combination. To investigate how representative of the global profiling data our selected QC analytes were, a complete

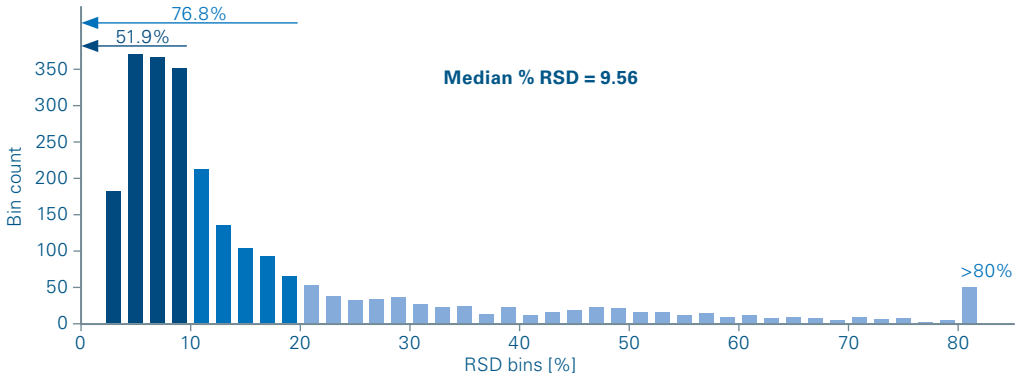


Figure 4

Relative Standard Deviation distribution for feature peak areas extracted by T-ReX® in MetaboScape.

The % RSD of feature peak areas determined by MetaboScape across all 300 injections. No within-batch correction was applied. The bin count is plotted vs. RSD [%]. The last bin contains all values >80%.

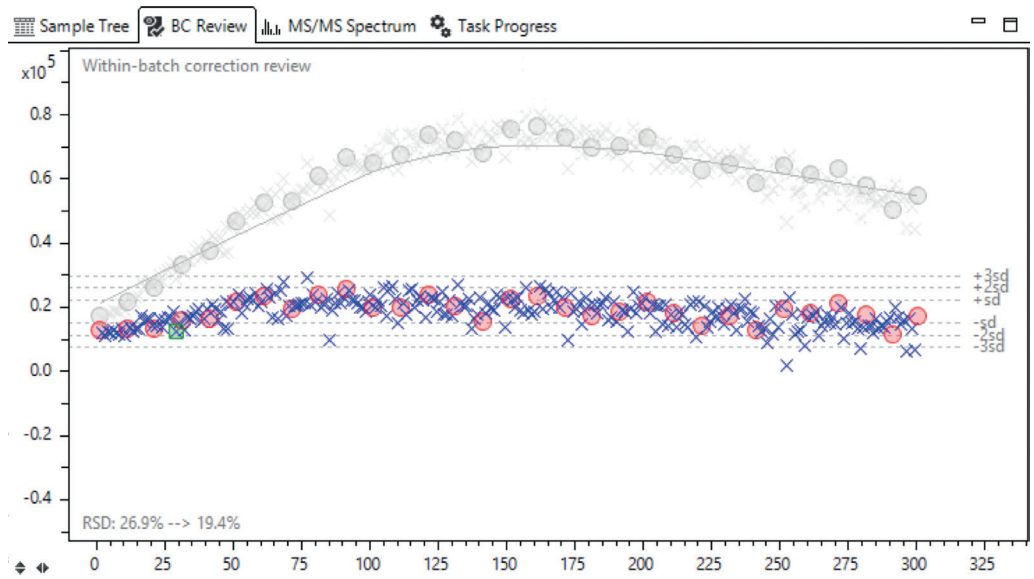


Figure 5

Within-batch correction in MetaboScape.

The Batch Correction (BC) Review plot highlights a feature showing systematic variation across the batch, and the successful application of feature-specific signal correction.

post-acquisition analysis of the data was performed using Bruker MetaboScape software, the results of which are shown in Figure 4. The median % RSD of features peak areas determined by MetaboScape across all 300 injections was 9.55%. In total 2448 features were extracted, with 51.9% of all features having an RSD below 10% and 76.8% of all features having an RSD below 20% highlighting the raw data quality without in-batch correction, and the suitability of RealTimeQC and the analyte QC approach to accurately represent a global profiling dataset.

Finally, even with high precision measurement, feature specific run order effects can and do occur. Figure 5 highlights such a feature that was automatically within-batch corrected in MetaboScape using a soft LOESS correction curve based on feature areas in QC samples. This selected example is a reminder that even with high quality monitored raw data, systematic variation correction can still positively impact overall data quality in global profiling applications.

Table 2
LC parameters

LC	Agilent 1290 Infinity II UHPLC		
Column	Avantor ACE Excel 2 AQ (150 x 2.1 mm, 2.0 µm)		
Column Oven Temp.	45°C		
Mobile phase	A: Water + 0.1% Formic acid B: Acetonitrile + 0.1% Formic acid		
Pump	Seal wash	Isopropanol / Water 10% / 90%	
Gradient	**Time [min]**	**Flow [mL/min]**	**%B**
	0.0	0.60	1
	0.1	0.60	1
	10.0	0.60	55
	10.15	0.61	65
	10.30	0.63	75
	10.45	0.67	85
	10.60	0.75	95
	10.70	0.8	100
	11.00	1.00	100
	11.55	1.00	100
	11.65	1.00	1
	11.70	0.90	1
	11.80	0.80	1
	11.90	0.70	1
	12.00	0.65	1
	12.10	0.61	1
	12.15	0.60	1
	12.65	0.60	1
Multisampler	Temperature	4°C	
	Injection volume	2 µl	
	Use vial/well bottom sensing	Yes	
	Draw speed	100 µL / min	
	Eject speed	400 µL / min	
	Wait time after draw	1.2 sec	
	Wash solvent	Isopropanol / Water 50% / 50%	
	Standard Wash	Flush port, 3 seconds	

Conclusion

- The Bruker impact II VIP QTOF hyphenation with the Agilent 1290 LC provides high quality raw data across hundreds of injections (median % RSD of feature peak areas across all 300 injections was 9.55%).
- The RealTimeQC module of TASQ software provides a visualization of the data in real-time for selected QC analytes. This allows for at-a-glance statistics and trends to be inspected conveniently across analytes, outliers to be automatically detected, and informed decision making (to stop or continue analysis, potentially sparing precious samples) to be made when it matters most.
- Correction of feature specific run order effects remains a critical component of post-acquisition data analysis, provided conveniently and automatically by MetaboScape software as a key part of the Bruker QC ecosystem.

References

[1] Broadhurst et al. (2018). *Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies.* Metabolomics **14**(72) https://doi.org/10.1007/s11306-018-1367-3

[2] Lewis et al. (2022). *An Open Platform for Large Scale LC-MS-Based Metabolomics*, ChemRxiv https://chemrxiv.org/engage/chemrxiv/article-details/61ebd6fa0716a8529e3823dc

[3] https://github.com/phenomecentre/npc-open-lcms/blob/main/Protocols/NPC.SOP.MS005_v2.1_PP_RP%20urine.pdf

For Research Use Only. Not for use in clinical diagnostic procedures.

Bruker Switzerland AG

Fällanden · Switzerland
Phone +41 44 825 91 11

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

ms.sales.bdal@bruker.com – www.bruker.com

Learn more at **www.bruker.com/impact**