Thermo Fisher
S C I E N T I F I C

Metabolomics

Quality assurance and quality control in metabolomics: achieving high-quality data for high-quality results

Authors
Bashar Amer, Rahul R. Deshpande, Amanda Souza, and Susan S. Bird

Thermo Fisher Scientific, San Jose, CA

Summary
The need for quality assurance and quality control (QA/QC) in metabolomics has been recognized by the metabolomics community, in no small part due to the challenges of the increasing size and scope of large-scale metabolomics studies. This technical note aims to summarize important ongoing efforts and share the best practices for incorporating robust QA/QC strategies, thereby improving the integrity and robustness of untargeted metabolomics datasets and saving time and valuable samples by obviating the need for repeated analysis.

The main topics we will discuss are as follows: (1) QA and QC as defined in the metabolomics community; (2) QA tools for establishing system suitability prior to beginning sample analysis, and (3) QC tools that are routinely used to ensure reproducibility and data integrity during sample analysis.

1. Introduction
Metabolomics is a powerful analytical tool for the comprehensive detection of small molecules describing the biochemical phenotype in biological systems. As such, it finds application in various areas of research and industry. By profiling polar metabolites and lipid species in biofluids and tissue samples, researchers gain knowledge of the physiology and pathophysiology of phenotypical endogenous metabolites[1] and exogenous substances such as drugs or toxicants[2]. However, for metabolomics to provide valuable biological insights, robust reproducible measurements, and confident metabolite identifications are required. Here, we will explore critical steps before, during, and after LC-MS acquisition to ensure high-quality metabolomics data.

Implementation of metabolomics for large-scale projects highlighted the need for standardized protocols.[3-5] Interlaboratory studies have emphasized the requirement for harmonized protocols.[6] The need for quality assurance and quality control has been recognized by the broader community, and to address this need, several initiatives were launched including a metabolomics quality survey,[7] the mQACC consortium,[8,9] and the lipid standards initiative[10]. The goal of these initiatives is to raise awareness, define

thermo scientific

and disseminate standardized protocols, harmonize existing protocols, and establish reporting guidelines for the publication of metabolomic and lipidomic data.

The need for consistent disease diagnosis requires assay validation and mandates quality assurance and quality control[11] for the highest degree of reliable data. Newborn screening is an early implementation of tandem mass spectrometry-based metabolic profiling,[12] which incorporates targeted analysis of amino acids, biogenic amines, and acylcarnitines, to specifically detect inborn errors of metabolism. Newborn metabolic profiling provides a blueprint for QA/QC protocols in LC-MS-based methods. Many of these guidelines have been adopted by metabolomics researchers and practitioners while performing targeted analyses. These are yet to be widely adopted in untargeted metabolomics. Nevertheless, there is a clear overlap of QA/QC procedures in both targeted and untargeted analyses, employing system suitability testing, internal standards, and long-term reference material to ensure highly confident data. Thus, it is important that researchers adopt appropriate quality processes and measures to demonstrate high-quality, reliable data from untargeted metabolomics studies.

Experimental design for metabolomics studies requires analytical rigor to minimize technical variance obscuring the detection of biological variance and discrimination of metabolites. This article describes analytical controls to identify technical variations and appropriate mitigation steps.

2. Defining quality assurance and quality control in metabolomics

QA/QC processes (Figure 1) are required for ensuring analytical performance and reliability for both targeted and untargeted metabolomics studies.[1,5] QA comprises procedures done prior to data acquisition to test that the analytical system is suitable to obtain the required data quality. QC includes procedures done during and immediately after the sample analyses to test that the analytical data are reliable and reproducible.

Quality measures for metabolomics LC-MS data are designed specifically to:

- Ensure robust LC-MS system performance during every analytical batch
- Consistently obtain reliable and reproducible metabolomics data
- Increase untargeted metabolomics data quality and integrity

To ensure biological relevance for a large-scale metabolomics study, quality assurance processes are also needed to design and construct the experimental study so that the total analytical and sampling variation is much smaller than the expected biological variation in a human population.[1]

3. Quality assurance establishes system suitability and function

Quality assurance establishes instrument system suitability and function prior to starting the sample analysis (Figure 2). QA ensures that all components of the LC-MS system, individually and in combination, are performing optimally. This can typically be assessed with a simple LC-MS injection. Additional LC-MS injections may be used to evaluate the performance of a pre-defined assay. System suitability comprises testing the chromatographic system for retention time variability, chromatographic resolution, and maintaining acceptable peak shape. Mass spectrometer performance is assessed for mass measurement accuracy, mass spectral resolution, and isotopic fine structure. Additional testing should be performed to ensure good MS/MS performance required for molecular ion isolation and fragmentation, which is needed for reliable compound annotation and identification.

Quality Assurance (QA)
Prior data acquisition

- Thoughtful experimental design
- Staff training
- Construction of standard operating procedures (SOPs)
  o Sample collection and handling
  o Instrument operation
  o Instrument maintenance and calibration

Quality Control (QC)
During and after data acquisition

Samples collection   Data acquisition   Data analysis

- QC samples and standards
  o Pooled QC sample(s) for intra-study assessment
  o System suitability standards
  o System suitability blanks
  o Internal standards for correction, normalization, and assessment
  o Standard reference materials (SRMs) and long-term reference (LTR) QC for inter-study and inter-laboratory data assessment
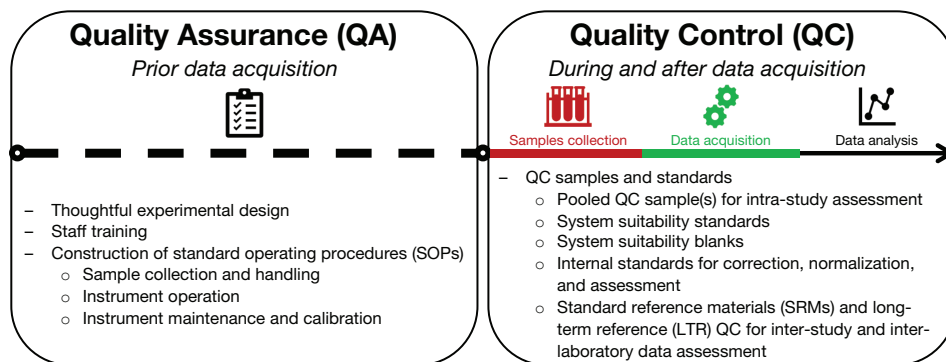
Figure 1. Definition of quality assurance and quality control in metabolomics

To assess system suitability, a blank sample is injected first to confirm that the LC-MS system is free of solvent impurities, contamination, or injection carryover. The blank sample generally consists of an injection medium or a similarly related solvent solution. Then, a neat mixture of reference standards that covers a broad range of molecular weight and metabolite classes is analyzed.[3] These mixtures may be custom created or sourced commercially. Table 1 shows an example of an in-house collection of endogenous metabolites for testing system suitability, while Table 2 provides a list of some commercially available mixtures. The latter includes unlabeled and stable isotope-labeled (SIL) amino acids, organic acids, acylcarnitines, and lipids.
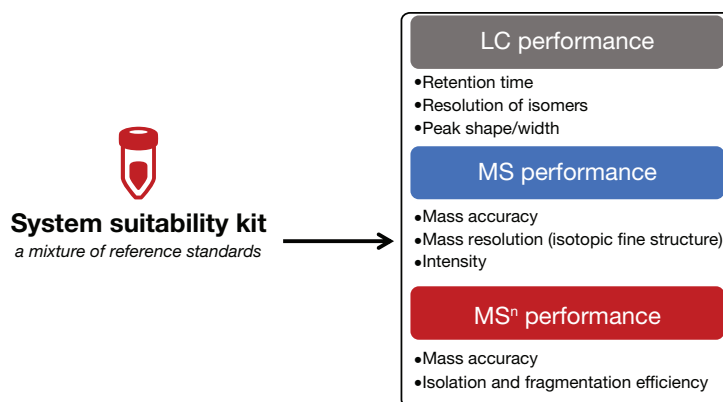

System suitability kit
a mixture of reference standards

LC performance
• Retention time
• Resolution of isomers
• Peak shape/width

MS performance
• Mass accuracy
• Mass resolution (isotopic fine structure)
• Intensity

MS^n performance
• Mass accuracy
• Isolation and fragmentation efficiency

Figure 2. System suitability for ensuring optimal instrument performance

Table 1. System suitability testing mixture for metabolomics (in-house)

Metabolite	Class	Formula	MW
Glycine	Amino acid	$C_2H_5NO_2$	75.0320
Isoleucine	Amino acid	C_6H13NO_2	131.0946
Leucine	Amino acid	$C_6H_{13}NO_2$	131.0946
Ornithine	Amino acid	$C_5H_{12}N_2O_2$	132.0899
Aspartic acid	Amino acid	$C_4H_7NO_4$	133.0375
Methionine	Amino acid	$C_5H_{11}NO_2S$	149.0511
Phenylalanine	Amino acid	$C_9H_{11}NO_2$	165.0781
Creatinine	Amino acid derivative	$C_4H_7N_3O$	113.0589
Pyruvic acid	Organic acid	$C_3H_4O_3$	88.0160
Succinic acid	Organic acid	$C_4H_6O_4$	118.0266
α-Ketoglutaric acid	Organic acid	$C_5H_6O_5$	146.0215
Kynurenic acid	Organic acid	$C_{10}H_7NO_3$	189.0426
Uracil	Nucleobase	$C_4H_4N_2O_2$	112.0273
Adenosine	Nucleoside	$C_{10}H_{13}N_5O_4$	267.0968
Glucose	Sugar	$C_6H_{12}O_6$	180.0634
Nicotinamide	Vitamin	$C_6H_6N_2O$	122.0480
Biotin	Vitamin	$C_{10}H_{16}N_2O_3S$	244.0882
Riboflavin	Vitamin	$C_{17}H_{20}N_4O_6$	376.1383
Folic acid	Vitamin	$C_{19}H_{19}N_7O_6$	441.1397
18:1(9Z) Lyso-PC	Lipid	$C_{26}H_{52}NO_7P$	521.3481
Glycodeoxycholic acid	Bile acid	$C_{26}H_{43}NO_5$	449.3141
Taurocholic acid	Bile acid	$C_{26}H_{45}NO_7S$	515.2917
17α-Hydroxyprogesterone	Steroid	$C_{21}H_{30}O_3$	330.2195
Thyroxine	Thyroid hormone	$C_{15}H_{11}I_4NO_4$	776.6867
Triiodothyronine	Thyroid hormone	$C_{15}H_{12}I_3NO_4$	650.7900

Table 2. Commercially available neat standard mixtures for use as QA and QC in metabolomics

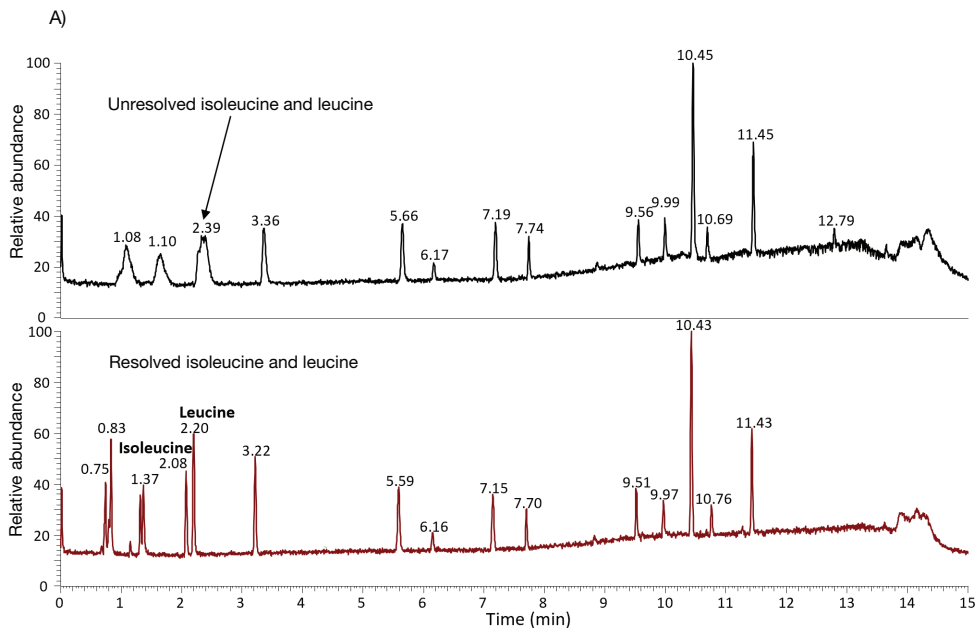| Product	Compounds
Thermo Scientific™ Pierce™ Amino Acid Standard H	Alanine; Arginine; Cystine; Glutamic acid; Histidine; Isoleucine; Leucine; Lysine; Methionine; Phenylalanine; Proline; Serine; Threonine; Tyrosine; Valine; Glycine; Aspartic acid
Thermo Scientific™ Pierce™ Small Molecule System Suitability Standard	Glycine, Atenolol, Flumetsulam, Atrazine, Tefenadine, Warfarin, Ultramark 1621, Methylmalonic acid, Rafoxinide
SPLASH™ Lipidomix™ Mass Spec Standard (Avanti Polar Lipids)	15:0-18:1(d7)-DG, -PA, -PC, -PE, -PG, -PI, and -PS; 18:1(d7)-LPE, -LPC, -MG, and cholesterol; 18:1(d9) SM; 15:0-18:1(d7)-15:0 TG; cholesterol (d7)
Metabolomics QC Kit (Cambridge Isotope Laboratories)	L-Alanine ($^{13}C_3$), L-Leucine ($^{13}C_6$), L-Phenylalanine ($^{13}C_6$), L-Tryptophan ($^{13}C_{11}$), L-Tyrosine (^{13}C), Caffeine ($^{13}C_3$), D-Glucose ($^{13}C_6$), Sodium benzoate ($^{13}C_6$), Sodium citrate ($^{13}C_3$), Sodium octanoate ($^{13}C_8$), Sodium propionate ($^{13}C_3$), Stearic acid, sodium salt ($^{13}C_{18}$), Succinic acid, disodium salt ($^{13}C_4$), D-Sucrose ($^{13}C_6$)
Metabolomics QReSS™ kit (Cambridge Isotope Laboratories)	Alanine (^{13}C, ^{15}N), 1,4-Butanediamine (putrescine) (^{13}C), Creatinine (N-methyl-D_3), Ethanolamine (D_4), Guanosine (^{15}N), Hypoxanthine (^{13}C), Leucine (^{13}C), Phenylalanine (^{13}C), Thymine (^{15}N), Tryptophan (^{13}C), Tyrosine (^{13}C), Nicotinamide (^{13}C), Citric acid (^{13}C), Fumaric acid (^{13}C), Indole-3-acetic acid (^{13}C), α-ketoglutaric acid (^{13}C), Palmitic acid (^{13}C), Pyruvic acid (^{13}C)
Carnitine standards set B (Cambridge Isotope Laboratories)	Carnitine, C1, C3, C4, C5 (isovaleryl), C8, C14, C16 acyl carnitines (labeled or unlabeled)
Organic acid mix (Cambridge Isotope Laboratories)	33 Organic acids (labeled or unlabeled): monoacids, diacids, aromatic acids, hydroxy acids, keto acids, and other

The reproducibility of mass accuracy, retention time, chromatographic peak shape, and peak area are compared to pre-established acceptance criteria to provide a quick pass-or-fail test of LC-MS system suitability. It is worth noting that, prior to performing sample analysis, especially when employing new HPLC columns, QC samples spiked with internal standards are run repeatedly (generally 10–20 times) to condition the UHPLC column and the entire LC-MS system to the sample matrix, and thus reduce unwanted variance.

3.1. Chromatography

A rapid system suitability check for chromatographic performance is illustrated in Figure 3 with the positive ion total ion chromatogram of a customized standard mixture. The early-eluting peaks region of the top chromatogram in Figure 3A consists of broad peaks width and poor peaks shape as evidenced by a lack of separation between the isoleucine and leucine isomers under these conditions, thus, failing the suitability test. After column maintenance, the system check was repeated and the early-eluting peak width and shape were determined to be acceptable as isoleucine and leucine were baseline separated, thus, passing the suitability criteria. In addition to peak width and isomer separation, retention time variability is another measure of chromatographic robustness.
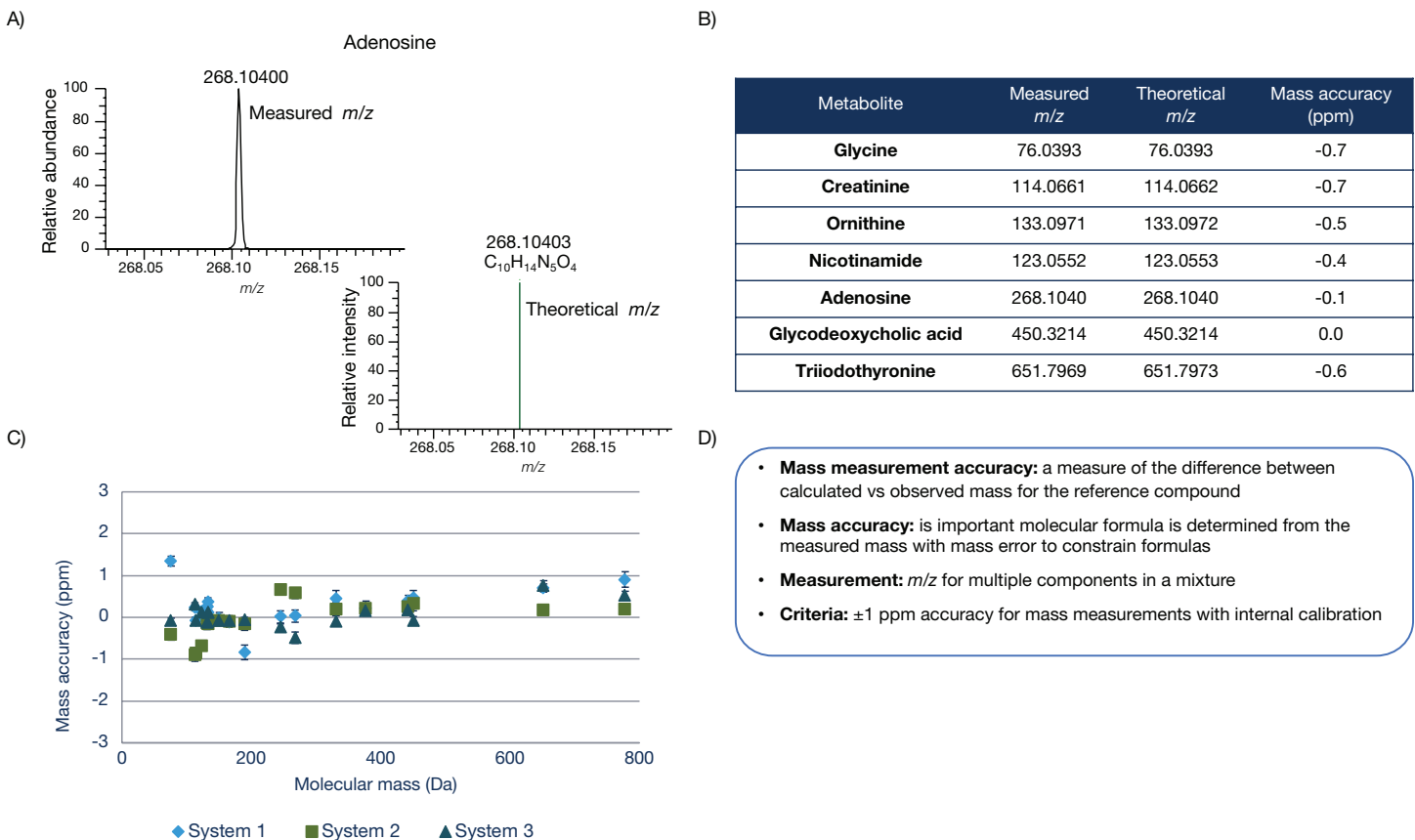
3.2. Mass accuracy

One of the fundamental properties of ultra-high-resolution accurate mass data is the measured mass-to-charge ratio (*m/z*) and the mass accuracy in parts-per-million (ppm) calculated for a known elemental composition. The mass accuracy measurement can be determined from a single injection of a reference standard for an individual metabolite (Figure 4A) or a collection of metabolites (Figure 4B). Further, these values can be recorded and stored as a reference for future instrument monitoring (Figure 4C). This test is critical to ensure accurate ion assignment to generate the reliable detection of expected target metabolites and the confident annotation of unknown species by prediction of elemental composition. For the latter, the elemental composition is assigned using mass tolerance to limit the possible molecular formulae during small molecule data analysis.

Figure 3. Reversed phase C18 UHPLC separation of a standard mixture. (A) Broad, early eluting peaks with no separation of isoleucine and leucine (top) compared to good peak width and shape with baseline separation of the two isomers (bottom). (B) Chromatographic metric and considerations for QA in metabolomics.

B)
- **Chromatographic reliability:** is required to annotate and quantitate hundreds of metabolites
- **Measurement:** retention time, peak width, peak shape, resolution
- **Criteria:** Rt CV ≤ 0.5%, isoleucine and leucine chromatographically resolved, good peak shape

B)

Metabolite	Measured m/z	Theoretical m/z	Mass accuracy (ppm)
Glycine	76.0393	76.0393	-0.7
Creatinine	114.0661	114.0662	-0.7
Ornithine	133.0971	133.0972	-0.5
Nicotinamide	123.0552	123.0553	-0.4
Adenosine	268.1040	268.1040	-0.1
Glycodeoxycholic acid	450.3214	450.3214	0.0
Triiodothyronine	651.7969	651.7973	-0.6

D)
- **Mass measurement accuracy:** a measure of the difference between calculated vs observed mass for the reference compound
- **Mass accuracy:** is important molecular formula is determined from the measured mass with mass error to constrain formulas
- **Measurement:** m/z for multiple components in a mixture
- **Criteria:** ±1 ppm accuracy for mass measurements with internal calibration
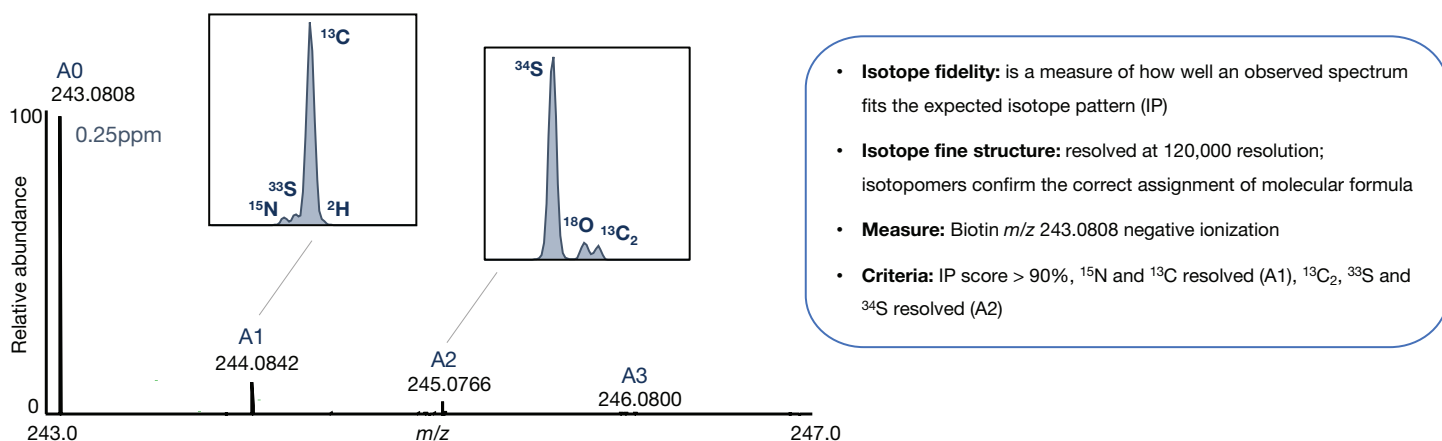
Figure 4. Instrument assessment for mass accuracy in a standard mixture. (A) Measured mass spectral accuracy (top) in the positive ion mode for adenosine compared to the theoretical exact mass (bottom) calculated by the elemental formula, $C_{10}H_{13}N_5O_4$. (B) Calculated mass accuracy for selected metabolites to assess instrument integrity. (C) Mass accuracy of a 25-component standard mixture in positive ion mode from three Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometers over six days of operation with built-in internal calibration. (D) Mass accuracy metric and considerations for QA in metabolomics.

3.3. Isotope fidelity and fine structure

A test for isotope fidelity and fine structure was conducted (Figure 5) to validate instrumental resolution for sufficient response to confirm elemental composition by matching the expected isotope pattern and resolving the isotopic peaks consistent with a specific elemental composition. In the negative ion mass spectrum of biotin, all the isotopic peaks were mass measured within ±1 ppm. Furthermore, the ^{13}C and ^{15}N isotopes were observed in the A1 peak (Figure 5, inset) and the A2 peak

was resolved with the expected exact masses and ratios for ^{34}S, $^{13}C/^{33}S$, and $^{13}C_2$ isotopologues. The detailed isotopic information provides a second important constraint, in addition to mass accuracy, for determining accurate elemental composition. This is critical since database searches are more reliable when searching molecular formulas in lieu of searching accurate mass alone. Thus, isotopic fine structure adds confidence to the elemental formula assignment and the subsequent annotation of unknown metabolites utilizing metabolite databases.

- **Isotope fidelity:** is a measure of how well an observed spectrum fits the expected isotope pattern (IP)
- **Isotope fine structure:** resolved at 120,000 resolution; isotopomers confirm the correct assignment of molecular formula
- **Measure:** Biotin m/z 243.0808 negative ionization
- **Criteria:** IP score > 90%, ^{15}N and ^{13}C resolved (A1), $^{13}C_2$, ^{33}S and ^{34}S resolved (A2)

Figure 5. A mass spectrum of biotin. Molecular ion was obtained at 120,000 resolving power. The monoisotopic peak (A0) measured at m/z 243.0808 and associated A1, A2, and A3 isotope clusters were also obtained. The inset displays isotopic fine structure at A1 and A2 clusters showing associated isotopologues ^{13}C, ^{15}N, ^{33}S, $^{13}C_2$, and ^{34}S in the ratios expected for $C_{10}H_{15}N_2O_3S$.
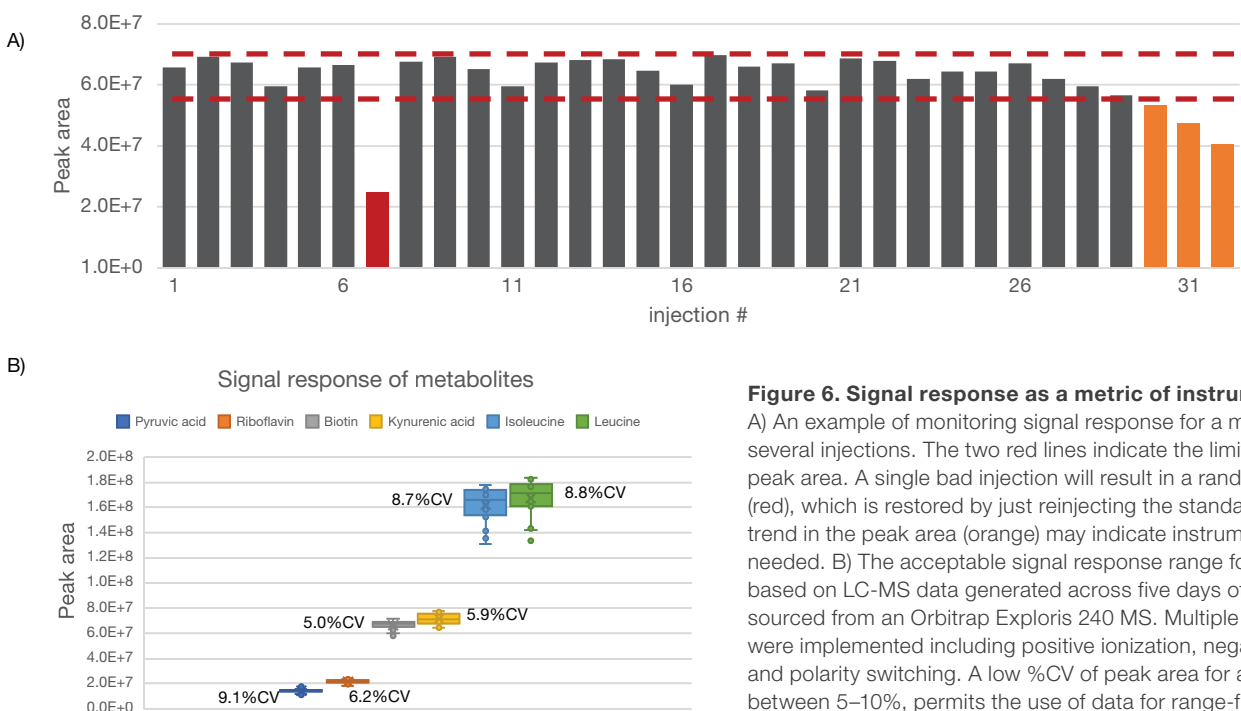
Figure 6. Signal response as a metric of instrument performance. A) An example of monitoring signal response for a metabolite across several injections. The two red lines indicate the limits of the acceptable peak area. A single bad injection will result in a random signal dropout (red), which is restored by just reinjecting the standards. A downward trend in the peak area (orange) may indicate instrument maintenance is needed. B) The acceptable signal response range for defined metabolites based on LC-MS data generated across five days of acquisition time sourced from an Orbitrap Exploris 240 MS. Multiple ionization modes were implemented including positive ionization, negative ionization, and polarity switching. A low %CV of peak area for all four metabolites, between 5–10%, permits the use of data for range-finding determination.

3.4. Signal response

The instrument response was tested to verify acceptable ion signals for selected metabolites. Using a known solution such as a neat standard mixture or an established matrix sample, measurements were compared against a pre-defined response range. The determination of an acceptable response range was derived from previous data (Figure 6). If signal response lies outside this range, lower or higher, appropriate steps must be taken to rectify the instrument status, for instance by cleaning the mass spectrometer front-end ion optics, cleaning, or position adjustment of the ion source metal needle, and/or replacing the LC column.

3.5. Activation and ion dissociation performance – HCD, CID, and MS[n]

Metabolomics assays may employ acquisition approaches that incorporate fragmentation spectra for the purpose of confirming the identification of known metabolites, quantitating target metabolites to generate absolute concentrations, or annotating unknown analytes. Consequently, fragmentation spectra should be evaluated during system suitability to ensure optimal instrument performance. Fragmentation by higher-energy collisional induced dissociation (HCD), for example, is compound dependent, with structural information varying by compound class and the selected adduct ion. To test the performance of HCD MS^2, the fragmentation spectrum of a standard is compared with its library spectrum and evaluated for the quality of structural information, mass accuracy, and library match score (Figure 7).
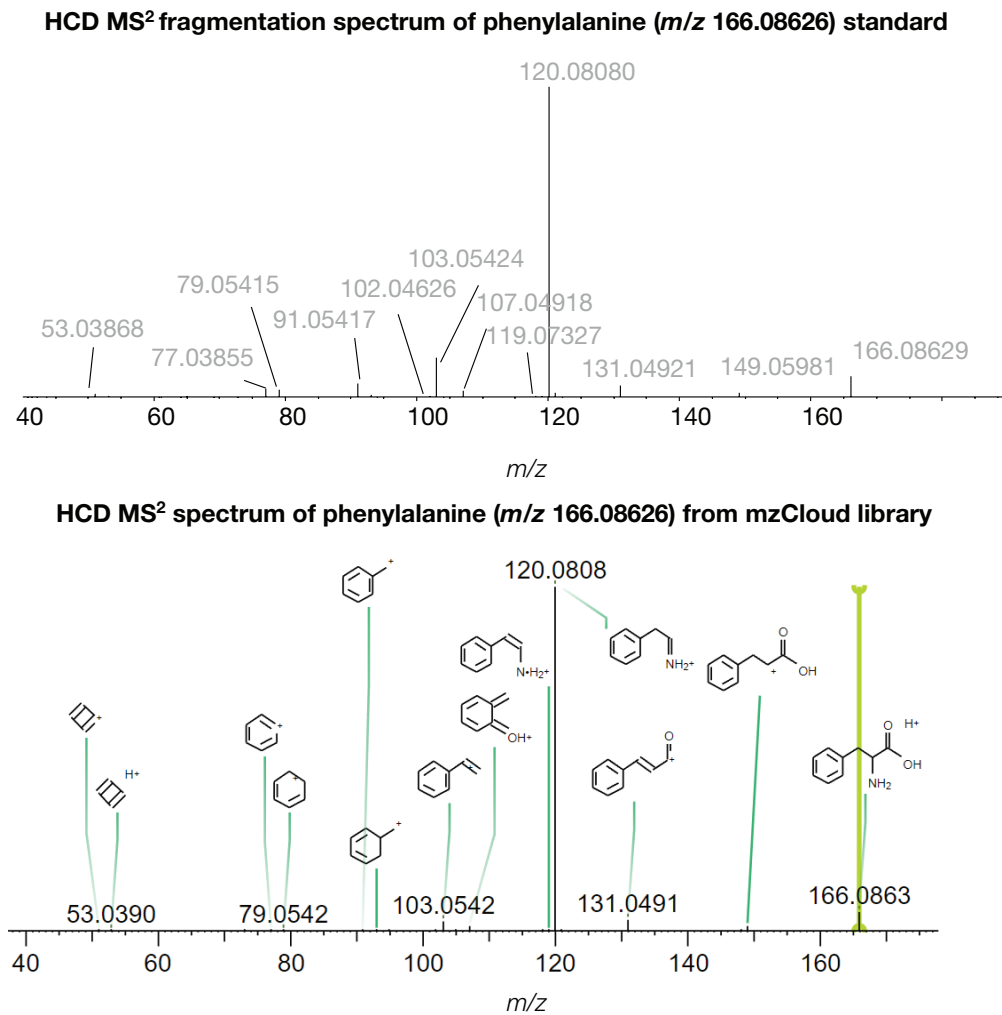
HCD MS^2 fragmentation spectrum of phenylalanine (*m/z* 166.08626) standard

HCD MS^2 spectrum of phenylalanine (*m/z* 166.08626) from mzCloud library

Figure 7. HCD MS^2 fragmentation of phenylalanine detected in human plasma and HCD MS^2 of phenylalanine pure standard from mzCloud™ library

As an alternative to HCD, fragmentation from resonant collision-induced dissociation (CID) can be achieved with instruments equipped with a linear ion trap mass analyzer such as the Thermo Scientific™ Orbitrap™ IQ-X Tribrid™ mass spectrometer. While HCD fragmentation of metabolites generally provides rich fragmentation spectra, additional complementary structural information may be needed. Further, depending on the application, the class of compounds investigated and the desired outcome, stepwise, multi-stage activation, MS^n may

be in order. Together, testing for CID and MS^n capability on associated systems ensures optimal performance. Expected product ions in the resulting CID or MS^n spectra were evaluated for mass accuracy and intensity. Improved lipid annotation can be achieved by utilizing HCD and CID fragmentation experiments. Figure 8 presents an example of using HCD MS^2 and CID MS^3 sequentially to annotate a triglyceride lipid (i.e., TG (16:0_18:3_18:1)).
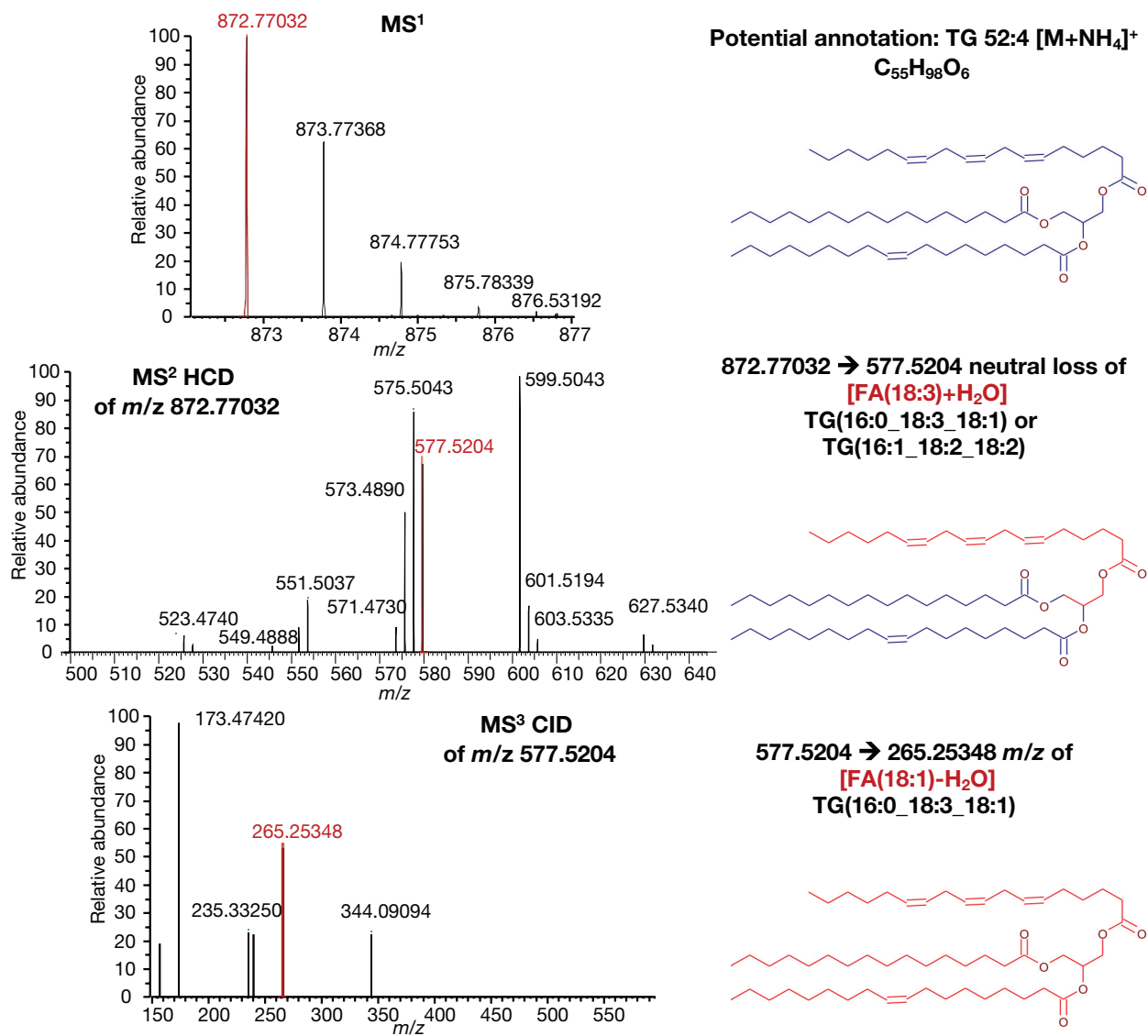

Figure 8. Confident lipid annotation using HCD MS^2 followed by CID MS^3 fragmentation

4. Quality control tools for monitoring system performance and data quality

Quality control tools monitor system performance during analysis and data quality during and after analysis. Common QC sample types include internal standards, pooled QC samples, and long-term reference materials (Figure 9).

The role of the internal standard is multipurpose, including measuring the reproducibility of representative analytes, monitoring signal response throughout the study acquisition, and detecting potential outliers resulting from injection error or sample preparation issues. Another common QC type in metabolomics studies is the use of a pooled QC sample, which is created by pooling an equal amount of all experimental samples or selecting a random subset of these samples in cases of larger studies. Intermittently injected throughout the injection sequence, the pooled QC samples demonstrate the reproducibility of multiple analytes as a function of variation. Reference materials, in addition, reflect matrix samples commonly analyzed, such as plasma or urine. These materials can be sourced commercially, e.g., NIST SRM 1950 human plasma, yet viable alternatives may include non-certified stock materials. With large volume stock, these materials are useful in tracking historical trends. Table 3 provides a good starting point for some of the commonly used reference materials that are commercially available today.
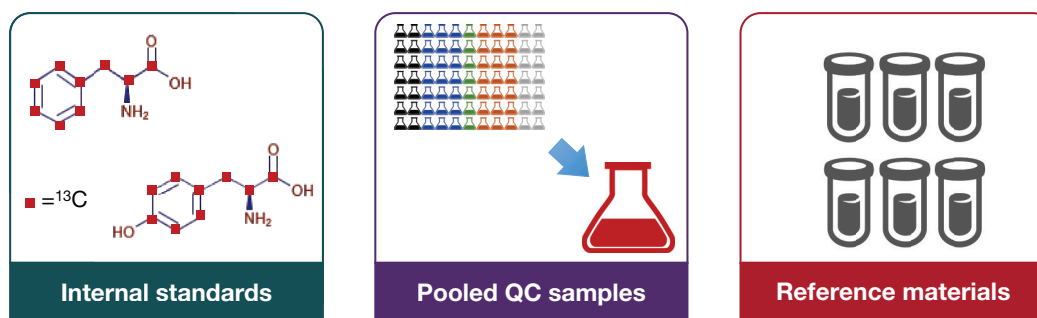
Figure 9. Accepted materials used for quality control measures in metabolomics experiments include internal standards, pooled QC samples, and reference materials representing matrix samples.

Table 3. QC reference mixtures for metabolomics

| Source	Product	Catalog #	Compounds
NIST	NIST SRM 1950	NIST1950	Metabolites in human plasma
	NIST SRM 3673	SRM 3673	Organic contaminants in non-smokers' urine
	NIST SRM 3672	SRM 3672	Organic contaminants in smokers' urine
	NIST SRM 3255	SRM 3255	Green tea (Camellia sinensis) extract
	NIST SRM 2378	SRM 2378	Fatty acids in human serum
CIL	Credentialed E. coli cell extract	MSK-CRED-KIT	E. coli extract unlabeled, U-^{13}C labeled metabolites
	Metabolite yeast extract	ISO1-UNL	Yeast extract with unlabeled metabolites
	Metabolite yeast extract U-^{13}C	ISO-1	Yeast extract with U-^{13}C labeled metabolites
Avanti Polar Lipids	E. coli Extract Total	100500	Total lipid extract
	Brain Extract Total	131101	Total lipid extract
	Heart Extract Total	171201	Total lipid extract
	Liver Extract Total	181104	Total lipid extract
	Yeast Extract Total	190000	Total lipid extract
	Soy Extract Polar	541602	Polar lipid extract

4.1. Internal standards

Internal standards are a valuable aid in detecting issues with different stages of sample preparation and those related to sample injection. Further, internal standards provide daily instrument monitoring over the course of data acquisition, which may be several days to over a week or longer. Analytes selected for internal standards should be, ideally, isotopically labeled and consist of materials with high chemical purity. An alternative source of internal standards when labeled compounds are not feasible or available is to use exogenous metabolites that are not present in the sample matrix of interest. Often, two to three labeled metabolites are selected as internal standards to be spiked into the extraction solution or directly added (spiked) to the experimental sample. Metrics to assess include mass accuracy, chromatographic retention time, and signal response. For lipids, employing multiple internal standards from several different compound classes is preferable. With the commercialization of labeled synthetic standard mixtures such as
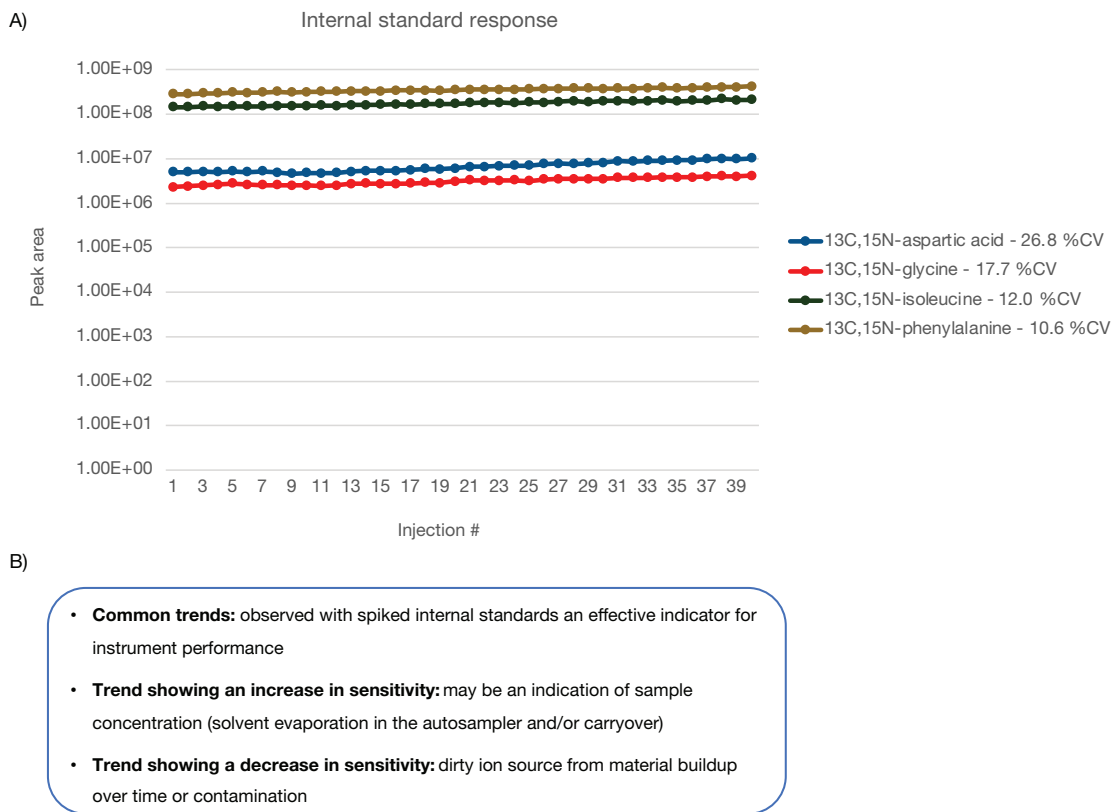
Avanti™ SPLASH™ Lipidomix™, it is becoming more economical to use stable isotope-labeled mixtures as internal standards or reference mixtures.

Figure 10 illustrates the value of spiking internal standards at different stages of sample preparation. In this analysis, d4-succinic acid reveals the presence of poor extraction recovery (outlier) while the d5-hippuric acid internal standard demonstrates a good *coefficient of variation* (CV) for sample reconstitution and LC autosampler injection. This type of QC information facilitates decision making such as re-injecting or re-extracting the sample in question or excluding this data point from subsequent data processing as an outlier.

Likewise, internal standards spiked into matrix samples are extremely useful for observing trends over time. Specific trends may indicate instrument issues requiring a course of action. Figure 11 illustrates the replicate analysis from the same autosampler vial of SRM 1950 human plasma spiked with $^{13}C,^{15}N$-labeled amino acids.
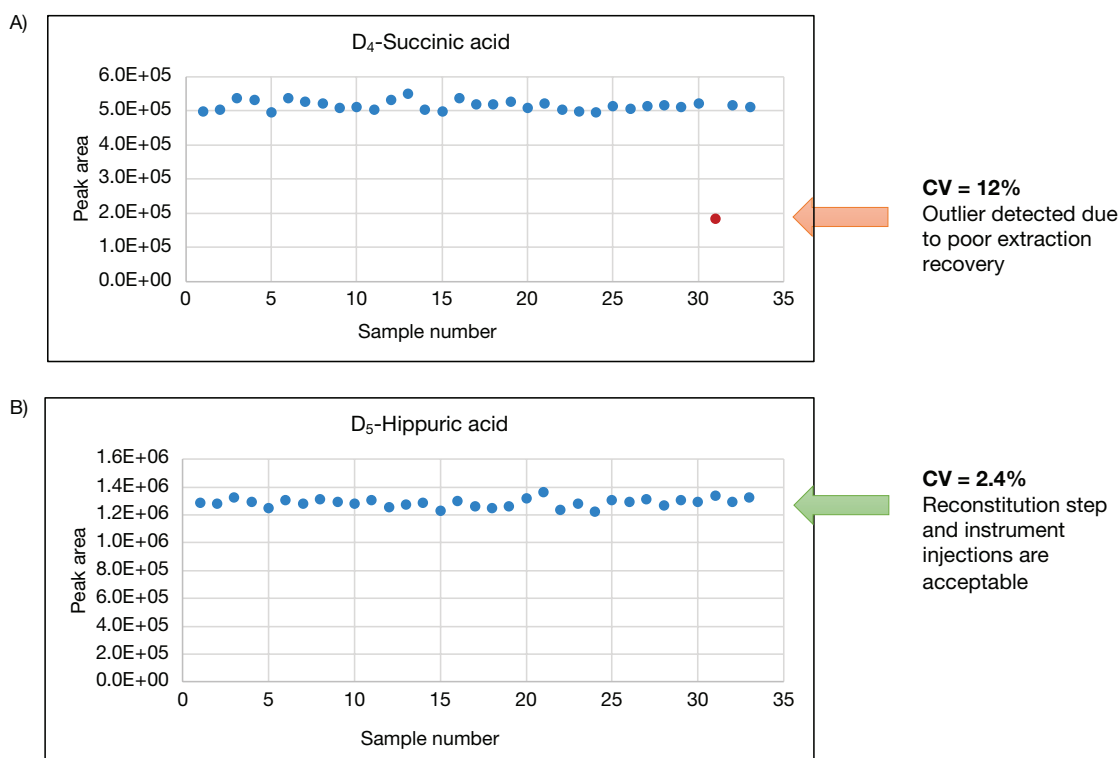
Figure 10. Internal standards spiked into SRM 1950 plasma during sample preparation. (A) D_4-Succinic acid, added during extraction, reveals outliers due to poor extraction recovery evaluated by peak response. (B) D_5-Hippuric acid, added during reconstitution, demonstrates a good CV for sample reconstitution and instrument injection evaluated by peak response.

A)

Internal standard response

Peak area vs Injection #

- 13C,15N-aspartic acid - 26.8 %CV
- 13C,15N-glycine - 17.7 %CV
- 13C,15N-isoleucine - 12.0 %CV
- 13C,15N-phenylalanine - 10.6 %CV

B)

- **Common trends:** observed with spiked internal standards an effective indicator for instrument performance
- **Trend showing an increase in sensitivity:** may be an indication of sample concentration (solvent evaporation in the autosampler and/or carryover)
- **Trend showing a decrease in sensitivity:** dirty ion source from material buildup over time or contamination
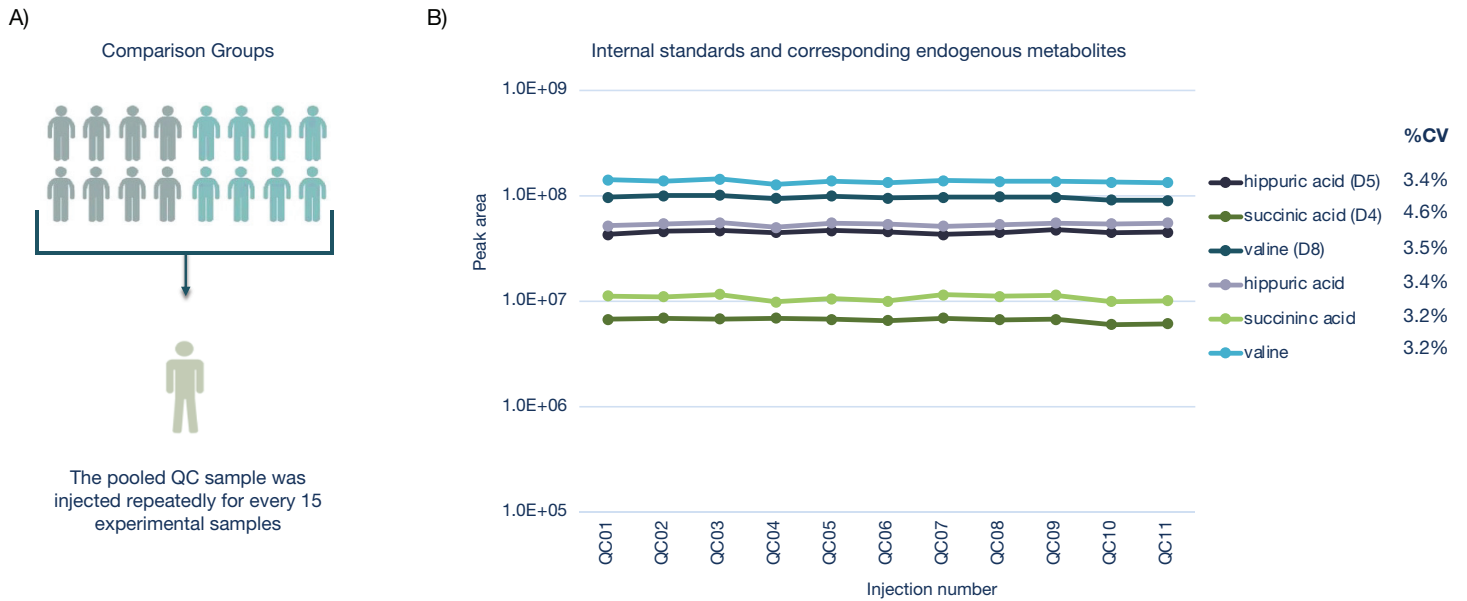
Figure 11. Amino acid internal standards spiked into SRM 1950 human plasma extract. (A) Replicate injections from the same autosampler vial of SRM 1950 human plasma extract spiked with $^{13}C,^{15}N$ amino acids showed a trend with increasing peak area for all four internal standards over 40 sample injections. Such a trend suggests the sample volume in the vial decreased, most likely due to evaporation resulting in the increased absolute peak area. (B) Internal standards are an effective indicator of instrument performance and can lead to corrective action such as ion source maintenance.

4.2. Pooled QC samples

A pooled QC sample stands as a representative sample for all experimental samples within a study. The pooled QC sample can be multi-purposed for QA and QC in addition to metabolite annotation and identification needs. Provided enough material, the pooled QC sample can be used in system suitability testing (mass accuracy, retention time, peak shape, etc.), conditioning a new LC column, or full system conditioning. When integrated into the sequence queue as part of the experiment analysis, this sample serves as a reproducibility assessment within the experiment since the same sample is injected repeatedly. In this case, the pooled QC samples can be analyzed for metabolite stability at the end of the analysis and compared against an expected range as another indicator for instrument performance and subsequent data quality. Figure 12 illustrates the use of pooled QC samples and internal standards to assess reproducibility for a human plasma study validating data integrity.

During untargeted MS-based metabolomics, full scan data is normally preprocessed to ensure unbiased peak detection. Here, pooled QC samples can be used to assess data quality and integrity using multivariate analysis and visualization tools like principal component analysis (PCA) plots. Figure 13 shows an example of pooled QC samples from an untargeted analysis arranged in a PCA plot with experimental samples indicating highly reproducible samples.

Figure 12. Amino acid internal standards and the corresponding endogenous metabolites detected in pooled human plasma QC samples. (A) The pooled QC sample was generated by combining the same volume of plasma from each experimental sample, creating a large, pooled stock plasma sample. This QC sample was then prepared and extracted the same way as the experimental samples and at the same time. Pooled QC samples were injected at intervals of every 15 experimental samples. (B) Signal response of the internal standards and their corresponding endogenous metabolites are plotted for all 11 QC samples. These six metabolites showed low variability with <5% CV, confirming the reproducibility of this same sample over the duration of the experiment. Data was acquired with an Orbitrap Exploris 240 LC-MS system.
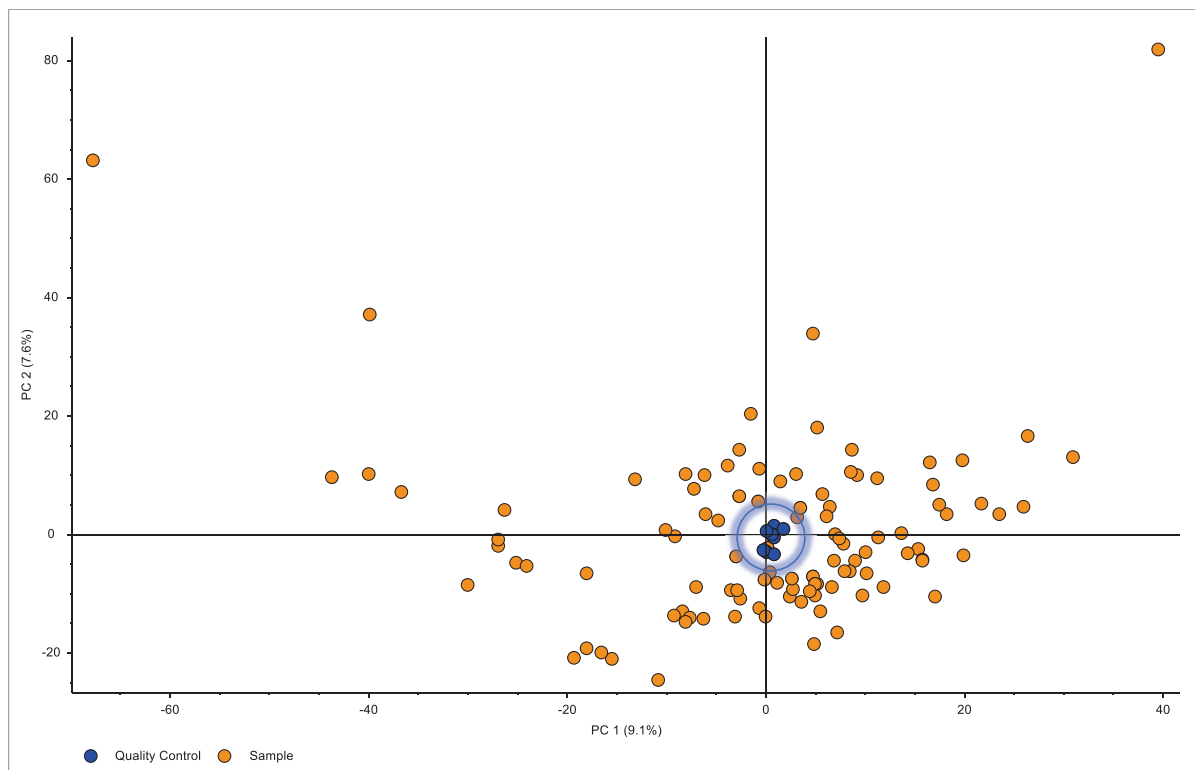
Figure 13. PCA plot generated from an untargeted analysis using QC samples to assess data quality. QC samples (blue circles) are highly correlated, indicated by tight clustering, and are clustered at the center of all experimental samples (orange circles) indicating reproducible and unbiased samples analysis.

In some instances, it may not be possible to generate a pooled QC sample due to a lack of sample material or when working with longitudinal studies samples collected over a course of several months or years. In this case, the use of surrogate matrices for endogenous metabolites is one approach.[13] A surrogate matrix is chosen to have close composition to the samples to be analyzed and be completely analyte-free, which may be a synthetic mixture (e.g., bovine serum albumin [BSA] in phosphate-buffered saline). Another solution is to employ a reference material that closely matches the matrix of the study samples.[1] Furthermore, in large longitudinal studies involving thousands of subjects is a sub-pooled sample prepared from the available samples at the time of starting the analysis (e.g., n > 500), which is then used throughout the study. Those subjects, however, should be sufficiently randomized such that the pooled sample reasonably represents the entire study population.

4.3. Long-term reference (LTR) materials

Additional QC recommendations include the use of Long-Term Reference (LTR) materials. LTR materials monitor measurement quality and assay performance over time. LTR samples typically comprise a large, pooled sample from a suitably relevant population to ensure an adequate supply over longitudinal and other long-term studies. These materials can be commercially available or sourced locally. LTR materials track historical trends to establish expectant LC-MS results for instrument performance; enable correction for batch-to-batch variation with a study, and permit comparison between laboratories for large multi-institution studies. Figure 14 illustrates the use of a yeast extract LTR material in lieu of a pooled QC showing excellent quantitative reproducibility of 15 amino acid metabolites across 3 different sample batches.

5. Implementation of QA and QC in metabolomics: a case study

In this section, LC-MS instrument performance was monitored during a phenotypical mouse study analysis utilizing internal standards coupled with pooled QC samples to provide detailed QC information.

5.1. Experimental design

This case study comprised a total of 40 plasma samples from C57BL/6 mice, 20 males, and 20 females, that were put on a normal or high-fat diet for 4 weeks. The study samples were randomized, and 10 µL of each plasma sample was then combined into a pooled QC sample. Each sample was split into two aliquots before being extracted for polar metabolites and lipids analyses. Sample preparation for polar metabolites consisted of protein precipitation using cold methanol (3:1) containing internal standards, whereas lipids were extracted using isopropanol (3:1) containing internal standards.[14] Chromatographic separation was achieved using a Thermo Scientific™ Hypersil GOLD™ reversed-phase C18 column for polar metabolites and a Thermo Scientific™ Accucore™ C30 column for lipids analysis. LC-MS data were obtained using a Thermo Scientific™ Vanquish™ Horizon UHPLC system coupled to an Orbitrap Exploris 240 mass spectrometer operated at 120,000 resolution using internal calibration (Thermo Scientific™ EASY-IC™ ion source). Pooled QC samples were injected every 10 samples and were also employed for unknown annotation using Thermo Scientific™ AcquireX™ intelligent acquisition.
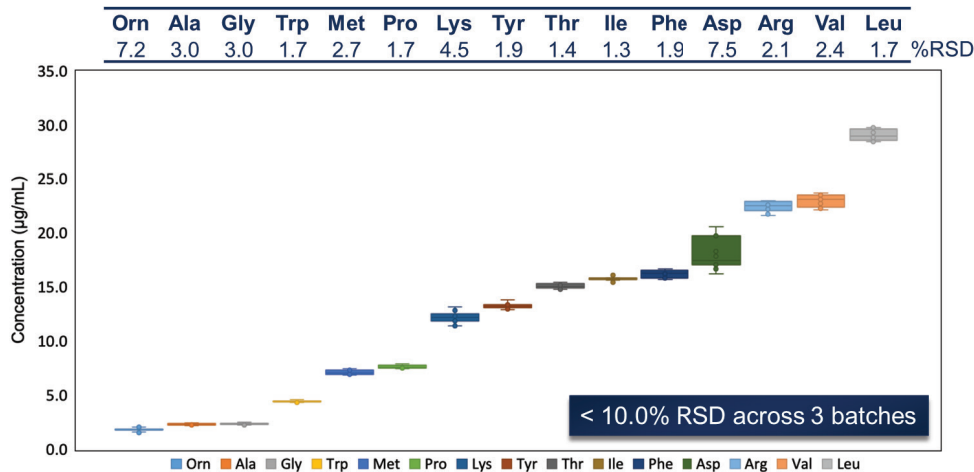
Figure 14. Box-and-whisker plot showing the reproducibility of 15 amino acids (< 10% CV) in yeast extract injected every 11 samples during 3 sample batches analyzed over the course of 9 days with an Orbitrap Exploris 240 MS
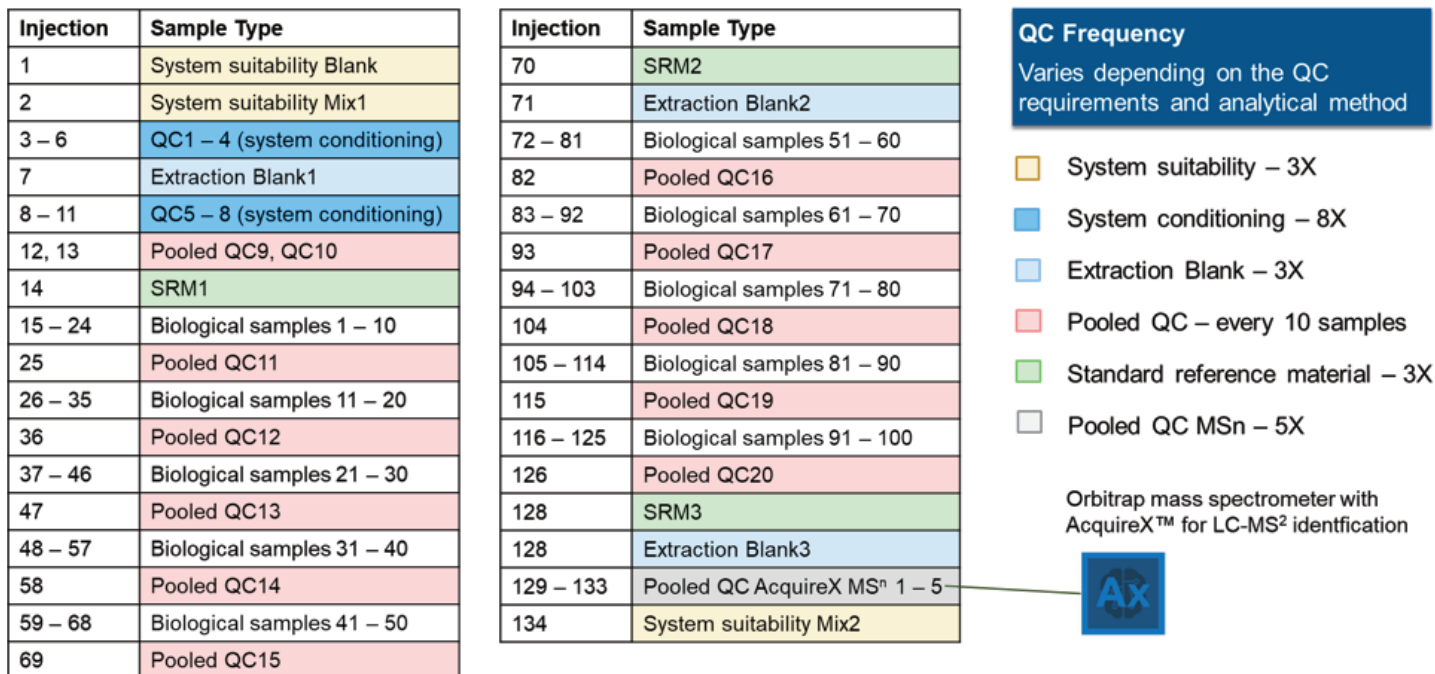
Injection	Sample Type
1	System suitability Blank
2	System suitability Mix1
3 – 6	QC1 – 4 (system conditioning)
7	Extraction Blank1
8 – 11	QC5 – 8 (system conditioning)
12, 13	Pooled QC9, QC10
14	SRM1
15 – 24	Biological samples 1 – 10
25	Pooled QC11
26 – 35	Biological samples 11 – 20
36	Pooled QC12
37 – 46	Biological samples 21 – 30
47	Pooled QC13
48 – 57	Biological samples 31 – 40
58	Pooled QC14
59 – 68	Biological samples 41 – 50
69	Pooled QC15

Injection	Sample Type
70	SRM2
71	Extraction Blank2
72 – 81	Biological samples 51 – 60
82	Pooled QC16
83 – 92	Biological samples 61 – 70
93	Pooled QC17
94 – 103	Biological samples 71 – 80
104	Pooled QC18
105 – 114	Biological samples 81 – 90
115	Pooled QC19
116 – 125	Biological samples 91 – 100
126	Pooled QC20
128	SRM3
128	Extraction Blank3
129 – 133	Pooled QC AcquireX MS^n 1 – 5
134	System suitability Mix2

QC Frequency
Varies depending on the QC requirements and analytical method

- System suitability – 3X
- System conditioning – 8X
- Extraction Blank – 3X
- Pooled QC – every 10 samples
- Standard reference material – 3X
- Pooled QC MSn – 5X

Orbitrap mass spectrometer with AcquireX™ for LC-MS^2 identfication

**Figure 15. Sample sequence including QA/QC**

5.2. Sample sequence

Sample order in analytical batches is important since the injection sequence prioritizes the information needed first for establishing system suitability and then ensures the LC-MS system is conditioned prior to sample analysis. The frequency of pooled QC sample injections, extraction blanks, SRM/LTR, and QC samples for MS/MS unknowns' annotation is determined based on the study size and requirements.

Figure 15 illustrates a sample sequence in a batch of 100 biological samples. Initially, blanks and standards were injected to establish LC-MS system suitability. Next, the entire LC-MS system was conditioned with pooled QCs. After assessing system suitability criteria, analysis of biological samples commenced after running extraction blank and SRMs. QC samples were run at a constant interval (every 10–15 samples) with the frequency determined by study size and analytical requirements. At the end of the batch, pooled QC samples were repeatedly run five times using AcquireX intelligent data acquisition for iterative dd-MSn compound identification and the system suitability mixture was re-analyzed to establish the final instrument status.

5.3. Data analysis

For the analysis of polar metabolites, D8-phenylalanine and D8-valine internal standards were added during the extraction process to evaluate sample preparation and instrument technical variance. The internal standards were monitored for mass accuracy and peak area reproducibility (Figure 16A) using Thermo Scientific™ TraceFinder™ software. Mass measurement accuracy was sub-1 ppm, and the peak area CV was 2.4% (D8-Phe) and 5.2% (D8-Val).

Avanti SPLASH Lipidomix standards were added during lipid extraction to evaluate the overall technical variance for non-polar metabolites. The internal standards were monitored for mass accuracy and peak area reproducibility (Figure 16B). The peak area plotted for 6 of SIL lipid standards remained constant throughout the study with less than 5% CV for most of the lipids. All internal standards were measured with sub-ppm mass accuracy in every injection.
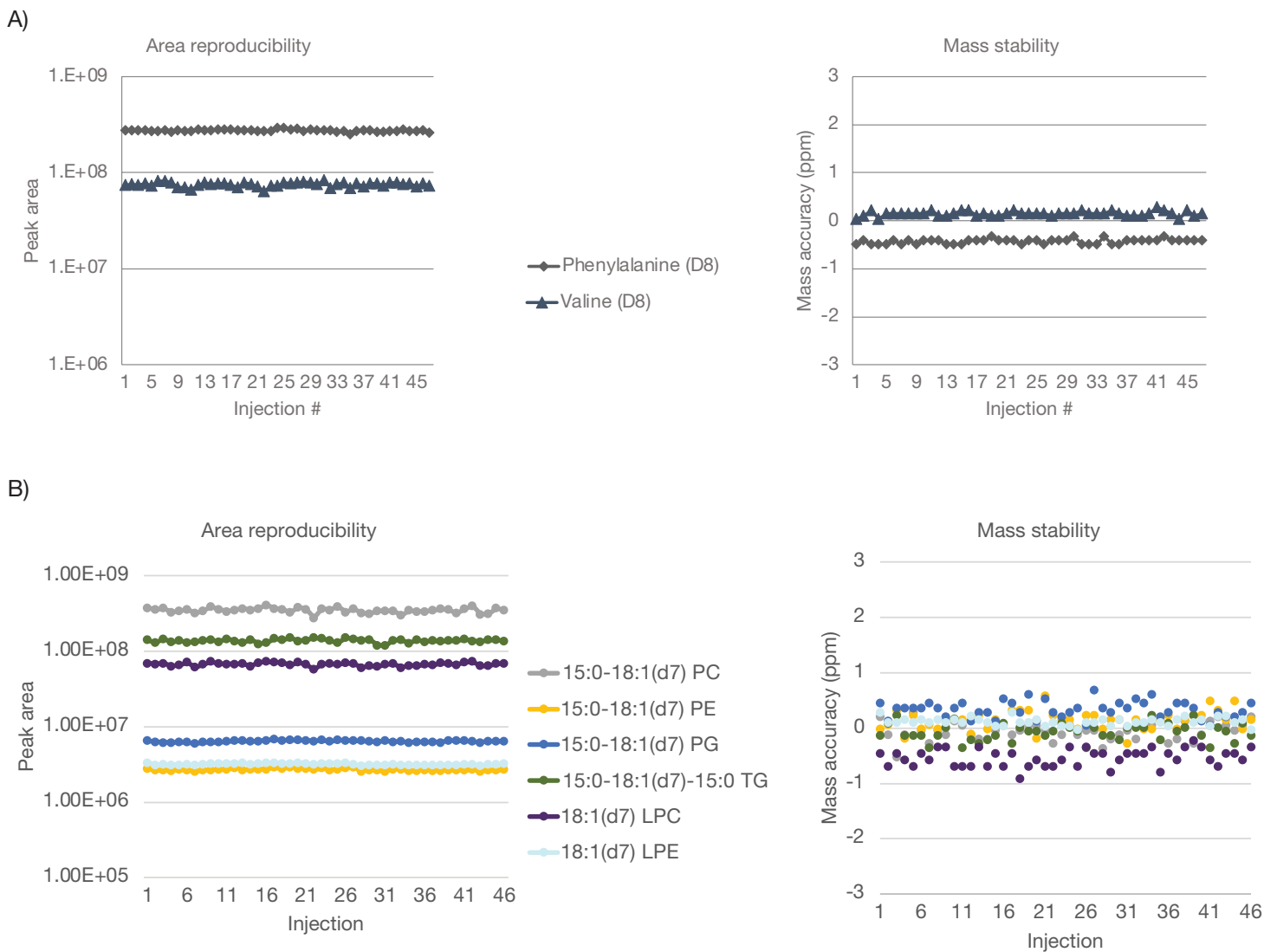
Figure 16. Internal standards in the study and pooled QC samples: mass accuracy and peak area reproducibility, 47 injections over 15 hours. (A) polar metabolites and (B) lipids

The study data were processed with Thermo Scientific™ Compound Discoverer™ software, and PCA reveals that there is a clear separation between mice fed a normal diet and mice fed with a high-fat diet (Figures 17A, polar metabolites, and 17B, non-polar lipids). The pooled QC samples (black) tightly cluster in the middle of the PCA plot confirming low technical variance in this study. Therefore, ensuring confidence observed changes/variations in metabolites, as shown in the volcano plots, that are downregulated (green) or upregulated (red) by the high-fat diet.
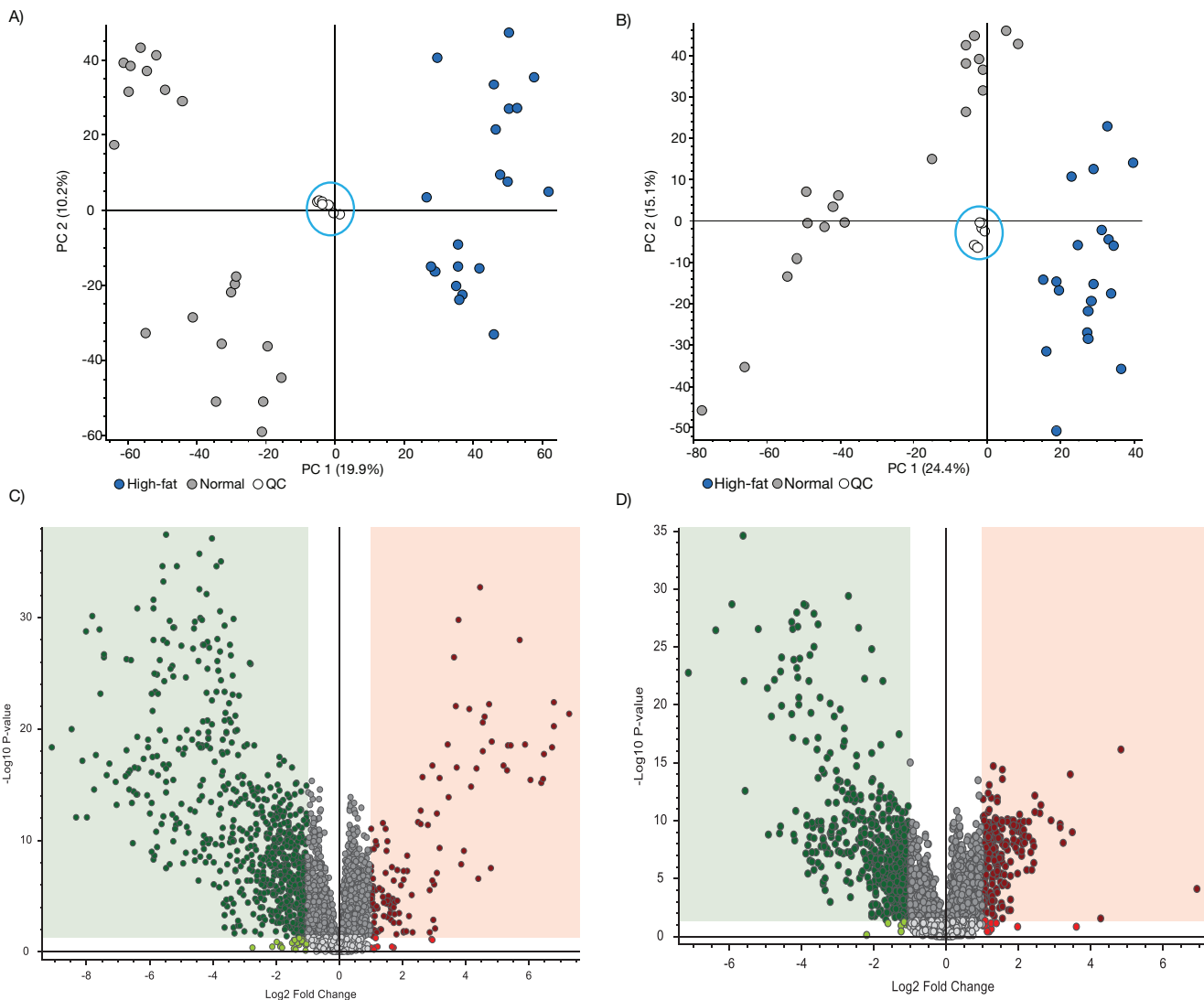
Figure 17. QC samples were utilized in a diet-induced mouse study to evaluate quantitative reproducibility. (A) PCA plot from the analysis of polar metabolites shows pooled QC samples (white circles) tightly clustered at the center among the experimental samples, mice fed normal (gray) and high-fat diets (blue), confirming low technical variance. (B) PCA plot from the analysis of lipid species shows pooled QC samples (white circles) also tightly clustered at the center among the experimental samples confirming low technical variance in this dataset. Volcano plots present changes/ variations in metabolites and lipids that are downregulated (green) or upregulated (red) by the high-fat diet, (C) and (D), respectively.

6. Summary/conclusions

Incorporating robust QA/QC strategies such as the examples given in this technical note will help improve the overall quality and reproducibility of reported datasets for untargeted metabolomics and save time and valuable samples by minimizing the need for repeat analysis. Broadhurst et al.[5] also suggest that studies should be published alongside comprehensive QC reporting to increase confidence in the data. The metabolomics community is actively working towards consensus for standardized QA and QC strategies for untargeted metabolomics, sharing best practices through mQACC.

Additional work yet to be done includes designing more universal standard mixtures suitable for use with different chromatographic (RP and HILIC) and ionization methods (positive and negative ion

response in ESI LC-MS). In addition, DOE (design of experiment) needs to be done to optimize relevant processes for sample collection, sample preparation, metabolomics workflow, QA/ QC, and data analysis to provide the desired higher confidence in results while maximizing biological significance. The field of metabolomics is starting to move beyond the traditional separate discovery untargeted and targeted quantitation runs to single injection, simultaneous targeted and untargeted discovery metabolomics[15] where stable labeled internal standards are used for QA/QC, unambiguous identification, and absolute quantification. This is another push toward maximizing the biological significance of the data by understanding the differences in metabolite concentration and allows for the comparison of metabolites within a sample in contrast to relative quantitation where patterns are only valid for single metabolites across samples.

Glossary

Analyte – a molecule that is annotated, identified, and quantified during metabolomics analysis

Batch – samples that are prepared at one time and analyzed in a single instrument run

Blank – a system blank, typically pure water, that is carried through the entire sample preparation method

Calibration Standards – standards used to calibrate peak areas vs. concentration for one or more analytes

Internal Standards – non-endogenous or labeled standards for normalizing extraction and/or injection variation

Long-term (LT) QC – reference material (pooled human plasma) used to compare historical data tracking

Lipidomics – untargeted lipidomics endeavors to comprehensively identify and quantify endogenous lipids in biological systems

Metabolomics – untargeted metabolomics endeavors to comprehensively identify and quantify endogenous metabolites in biological systems

Metabolic signature – a panel of affected endogenous, relevant metabolites that change in concentration in disease vs. normal states

Pathways – overview of biochemical reactions in cells; first compiled in 1965 by Dr. Gerhard Michal, ed. Roche Biochemical Pathways **http://biochemical-pathways.com/#/map/1**

Pooled QC – a sample consisting of equal amounts of extract from each biological sample

Quality Assurance – a procedure, performed prior to sample analysis, to test that the analytical system is suitable to obtain the required data quality

Quality Control – a procedure, performed during or after sample analysis, to test that the analytical data are reliable and reproducible

QA/QC Criteria – for a given test procedure, criteria define what range of results are acceptable and unacceptable

Reference Material – a material containing representative metabolites in a particular biological matrix

Reference Standards – metabolites synthesized, purified, and characterized by multiple analytical methods

System Suitability mixture – a mixture of metabolites designed to test system suitability before and after sample analysis

References

1. Dunn, W.B., et al. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans, Review, *Bioanalysis* **2012**, 4(18), 2249–2264. doi: 10.4155/bio.12.204 https://www.future-science.com/doi/10.4155/bio.12.204

2. Viant, M.R., et al. Use cases, best practice and reporting standards for metabolomics in regulatory toxicology, *Nature Communications* **2019**, 10, 3041. doi: 10.1038/s41467-019-10900-y https://www.nature.com/articles/s41467-019-10900-y

3. Dunn, W.B., et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, *Nature Protocols* **2011**, 6, 1060–1083. doi: 10.1038/nprot.2011.335 https://pubmed.ncbi.nlm.nih.gov/21720319/

4. Dunn, W.B., et al. Molecular phenotyping of a UK population: defining the human serum metabolome, *Metabolomics* **2015,** 11, 9–26. doi: 10.1007/s11306-014-0707-1 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4289517/

5. Broadhurst, D., et al. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies, *Metabolomics* **2018**. 14, 72. doi: 10.1007/s11306-018-1367-3 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5960010/pdf/11306_2018_Article_1367.pdf

6. Bowden, J.A., et al. Harmonizing lipidomics: NIST interlaboratory comparison exercise for lipidomics using SRM 1950–Metabolites in Frozen Human Plasma, *J. Lipid Research* **2017**, 58(12) 2275–2288. doi: 10.1194/jlr.M079012 https://pubmed.ncbi.nlm.nih.gov/28986437/

7. Dunn, W.B., et al. Quality assurance and quality control processes: summary of metabolomics community questionnaire, *Metabolomics* **2017**, 13, 50. doi.org/10.1007/s11306-017-1188-9 https://link.springer.com/article/10.1007%2Fs11306-017-1188-9

8. Beger, R.D., et al. Towards quality assurance and quality control in untargeted metabolomics studies, *Metabolomics* **2019,** 15(1), 4. doi: 10.1007/s11306-018-1460-7 https://www.ncbi.nlm.nih.gov/pubmed/30830465

9. Kirwan J.A., et al. Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management, *Metabolomics* **2022**, 18, 70. https://link.springer.com/article/10.1007/s11306-022-01926-3

10. Lipidomics Standards Initiative Consortium., Liebisch, G., Ahrends, R. et al. Lipidomics needs more standardization, *Nature Metabolism* **2019,** 1, 745–747. doi: 10.1038/s42255-019-0094-z https://pubmed.ncbi.nlm.nih.gov/32694765/

11. De Jesús, V.R., et al. The newborn screening quality assurance program at the Centers for Disease Control and Prevention: thirty-five year experience assuring newborn screening laboratory quality, *International J. Neonatal Screening* **2016,** 1(1), 13–26. doi.org/10.3390/ijns1010013 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4545740/

12. Chase, D.H., Kalas, T.A. and Naylor, E.W. The application of tandem mass spectrometry to neonatal screening for inherited disorders of intermediary metabolism, *Annual Review of Genomics and Human Genetics* 2002, 3, 17–45. doi: 10.1146/annurev.genom.3.022502.103213 https://pubmed.ncbi.nlm.nih.gov/12142359/

13. Jones, B.R., Schultz, G.A., Eckstein, J.A. and Ackermann, B.L. Surrogate matrix and surrogate analyte approaches for definitive quantitation of endogenous biomolecules, *Bioanalysis* **2012,** 4(19), 2343–2356. https://www.future-science.com/doi/full/10.4155/bio.12.200

14. Hackbusch, S., Deshpande, Rahul., Amer, B., and Souza, A., Thermo Fisher Scientific Application Note 000696: Reproducible measurement and confident identification of the plasma lipidome from obese mice using the Orbitrap Exploris 240 mass spectrometer, **2022.** https://thermofishercmd.highspot.com/items/6243d915c3a59ca8f85f16a5?lfrm=srp.0

15. Amer B., Combining the Best of Both Worlds with Semi-Targeted Metabolomics. *American Pharmaceutical Review.* June 2022. https://www.americanpharmaceuticalreview.com/Featured-Articles/587068-Combining-the-Best-of-Both-Worlds-with-Semi-Targeted-Metabolomics/

Learn more at **thermofisher.com/metabolomics**

thermo scientific