analytical chemistry

pubs.acs.org/ac

Article

Quantifying Precision Loss in Targeted Metabolomics Based on Mass Spectrometry and Nonmatching Internal Standards

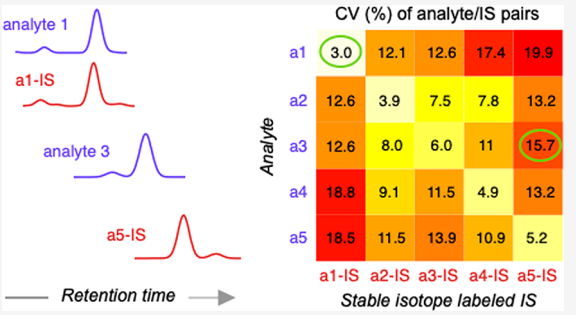
Arve Ulvik,* Adrian McCann, Øivind Midttun, Klaus Meyer, Keith M. Godfrey, and Per M. Ueland

Cite This: *Anal. Chem.* 2021, 93, 7616−7624

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In mass spectrometry, reliable quantification requires correction for variations in ionization efficiency between samples. The preferred method is the addition of a stable isotope-labeled internal standard (SIL-IS). In targeted metabolomics, a dedicated SIL-IS for each metabolite of interest may not always be realized due to high cost or limited availability. We recently completed the analysis of more than 70 biomarkers, each with a matching SIL-IS, across four mass spectrometry-based platforms (one GC−MS/MS and three LC−MS/MS). Using data from calibrator and quality control samples added to 60 96-well trays (analytical runs), we calculated analytical precision (CV) retrospectively. The use of integrated peak areas for all metabolites and internal standards allowed us to calculate precision for all matching analyte (A)/SIL-IS (IS) pairs as well as for all nonmatching A/IS pairs within each platform (total n = 1442). The median between-run precision for matching A/IS across the four platforms was 2.7−5.9%. The median CV for nonmatching A/IS (corresponding to pairing analytes with a non-SIL-IS) was 2.9−10.7 percentage points higher. Across all platforms, CVs for nonmatching A/IS increased with increasing difference in retention time (Spearman's rho of 0.17−0.93). The CV difference for nonmatching vs matching A/IS was often, but not always, smaller when analytes and internal standards were close structural analogs.



INTRODUCTION

The field of metabolomics has seen great expansion in recent years in step with a continual increase in performance of new analytical platforms. Much of the recent development has involved mass spectrometry methods with multiple-stage (e.g., tandem MS) detectors for high selectivity combined with increased ionization efficiencies and improved signal/noise ratios, enabling the detection of low-concentration endogenous metabolites. Metabolomics studies can be divided into targeted and nontargeted analysis, with the latter involving detection of thousands of signals, essentially nonquantitative and with largely deferred identification,[1] although during the past few years, great strides have been made toward developing semitargeted methods with extensive analyte coverage and good quantitative performance.[2−4] In contrast, targeted analysis involves the quantitation of a limited set of preselected analytes.[5]

A recognized limitation of mass spectrometry-based quantitative analysis is the requirement for external or internal calibration to correct for variations in ionization efficiency. Of particular concern is ion suppression/enhancement that is known to vary across samples and with retention time. Currently, the preferred method to correct for ion suppression involves the addition of a stable isotope-labeled internal standard (SIL-IS) where retention times can be expected to be close to or identical to that of the analyte.[5,6] In typical metabolomics settings, analytical runs will involve dozens or even hundreds of analytes. In such cases, the inclusion of a matching SIL-IS for every analyte can be challenging in terms of cost and availability. An alternative is to analyze several analytes against a single or a limited number of internal standards based on structural and/or retention time similarity;[7] if so, steps should be taken to ensure that quantification with the selected internal standards meet some minimum assay acceptance criteria. The analytical performance of different choices of internal standards has been compared in a number of studies,[8−10] but few have undertaken a systematic comparison covering a large range of metabolites in terms of chemical and physical characteristics.

We recently analyzed more than 70 biomarkers across four MS/MS-based analytical platforms as part of a large cohort study. Each biomarker was quantified using a matching (structurally identical) SIL-IS. The samples were analyzed in 96-well plates, with each plate including a fixed number of calibration and quality control samples. Using data for these samples, we were able to calculate within- and between-run

Received: January 11, 2021
Accepted: April 28, 2021
Published: May 20, 2021

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

Table 1. Main Characteristics of Analytical Platforms[a,e]

platform	GC	LC1	LC2	LC3
type of analysis	GC−MS/MS	LC−MS/MS	LC−MS/MS	LC−MS/MS
metabolites analyzed	amino and carboxylic acids	choline derivatives, charged or methylated amino acids	kynurenine pathway metabolites and B-vitamins	lipid-soluble vitamins
pretreatment	dithioerythriol			
protein precipitation	ethanol	trichloroacetic acid	trichloroacetic acid	ethanol
liquid−liquid extraction	isooctane/ chloroform[b]			isooctane/ chloroform[b]
derivatization	methylchloroformate			
column	CP Sil 24 CB	150 × 4.6 mm, 3 μm, phenyl	150 × 4.6 mm, 3.5 μm, C8	50 × 4.6 mm, 2.7 um, C18
mobile phase	helium	acetic acid:methanol	acetic acid:HFBA:acetonitrile	methanol: NH_4-formate
elution	temperature step gradient	isocratic	step gradient	isocratic
MS, ion source	ESI, positive mode	ESI, positive mode	ESI, positive mode	APCI, positive mode
analytes (n)	25	16	24	6
A/IS combinations (n)[c]	625	256	576	36
range of concentrations (μM)[d]	0.15−500	0.5−300	<0.01−70	<0.01−5
analyte peak area ($\log_{10}$) median (range)	4.5 (2.8, 6.7)	5.0 (3.7, 6.5)	5.8 (4.2, 6.9)	4.5 (3.1, 6.1)
IS peak area ($\log_{10}$) median (range)	4.6 (2.7, 6.6)	5.8 (4.3, 6.8)	5.5 (4.1−6.5)	4.9 (3.5−5.9)
RT (min), median (range)	3.7 (2.5, 7.5)	2.2 (2.0, 3.5)	3.6 (2.1, 4.8)	1.7 (1.3, 3.1)

[a]Ranges reported are across n analytes within each platform. [b]Platforms GC and LC3 shared sample processing up to and including this point. [c]Number of A/IS combinations investigated (all possible). [d]Range of typically observed (median) concentrations across n analytes. [e]Abbreviations: ESI, electrospray ionization; APCI, atmospheric pressure chemical ionization; A, analyte; IS, internal standard; RT, retention time; HFBA, heptafluorobutyric acid.

precision retrospectively. We used the integrated peak areas directly to investigate all possible analyte/internal standard (A/IS) pairings within a given analytical platform. Analytes paired with their own specific SIL-IS are referred to as matching A/IS, and all other combinations are nonmatching A/IS. The primary objective of the study was to quantify the difference in analytical precision between nonmatching and matching A/IS. A secondary objective was to assess or quantify the impact of structural similarity and similarity of retention time on analytical precision for nonmatching A/IS.

EXPERIMENTAL SECTION

Source Materials. During the period October 2019 to April 2020, we analyzed 4980 plasma samples as part of the project "Nutritional Intervention Preconception and During Pregnancy to Maintain Healthy Glucose Metabolism and Offspring Health" (NiPPeR).[11] The samples were distributed into 60 96-well plates, each containing 83 project samples, six calibrator samples, and six quality controls. An analytical run was defined as the analysis of one 96-well plate. The calibrator was pooled EDTA plasma from healthy individuals purchased from Innovative Research, Inc. (www.innov-research.com). Quality control 1 (three samples in each run) was pooled serum from Innovative Research, Inc. that were spiked for most analytes to serve as a high concentration control, and quality control 2 (three samples in each run) was serum pooled from healthy personnel (n = 23) at the Bevital laboratory and surrounding research units in Bergen, Norway. For precision calculations, the data from calibrator and quality controls were treated identically; hence, for the purpose of this study, the three sample types (calibrator and quality controls 1 and 2) were labeled source materials 1 to 3 (S1, S2, and S3).

Sample Processing, Chromatography, and Tandem Mass Spectrometry. The analytical platforms are labeled GC, LC1, LC2, and LC3. All sample processing was performed on Hamilton robotic workstations (Bonaduz, Switzerland) equipped with disposable tips and liquid detection. An overview of each platform is shown in Table 1. Briefly, the first step consisted of adding a protein precipitation agent (ethanol or trichloroacetic acid) containing internal standards (for GC, this step was preceded by addition of a reducing agent to liberate thiol compounds) followed by centrifugation. Further processing for platforms GC and LC3 included liquid−liquid extraction to separate water- and fat-soluble compounds. The aqueous phase was subject to chemical derivatization before analysis on GC, and the organic phase was reconstituted in methanol prior to analysis on LC3. For platforms LC1 and LC2, the aqueous phase was applied directly to the LC−MS/MS system.

Further details for GC and LC3,[12] LC1,[13] and LC2[14] were published previously. Most SIL-IS were labeled with deuterium (2−10 substitutions) except for histidine and homoarginine (LC1), which were labeled with ^{15}N, and neopterin and thiamine (LC2), which were labeled with ^{15}N and ^{13}C, respectively (Tables S1−S4).

Statistical Methods. Coefficients of variation (CV %) were calculated, assuming lognormal distributions, as $100 \times (e^{SD(X)} - 1)$ (eq 1),[15] where X is either log(A), log(IS), or log(A/IS). Different concentrations across source materials S1, S2, and S3 were normalized by appropriate correction in regression models. Specifically, normalization of signals (either peak areas or ratios) across a source material and run was obtained by employing a multilevel hierarchical linear regression model with the source material as a fixed effect and run as a random effect and varying intercepts. The residuals from this regression were input as X in eq 1. The same regression model, without correction for source materials, was used to obtain CVs for S1, S2, and S3

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

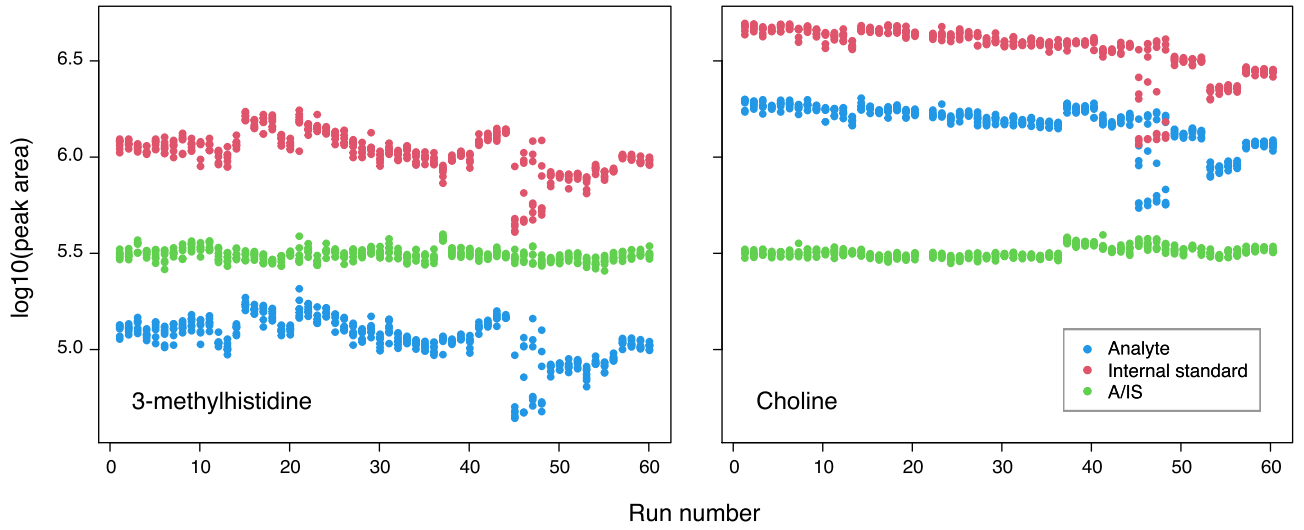


Figure 1. Peak areas (raw values) for analytes and internal standards across 60 runs. Data for two analytes, 3-methylhistidine and choline, on platform LC1 are shown. Corresponding between-run CVs (with adjustment for run) are depicted in Figure 2. The analyte-to-internal standard ratio (A/IS) is included for comparison and is arbitrarily centered at 5.5 on the *y*-axis. Peak areas from source material 1 (pooled EDTA plasma) were used for this analysis.

separately. For some analyses that either assessed or demonstrated the association of CV with peak area, we used data from source material 1 only, as specified. As part of constant quality control monitoring during analysis, a calibrator sample was removed if it fell outside tolerance values established for each analytical run. When such measures were deemed inadequate, the whole run was subject to repeat measurement. In our retrospective analysis of raw data, we aimed at analyzing only primary runs and we used all quality control samples for our main CV calculations. To mimic the data cleaning performed during analysis of project samples, we removed primary runs that displayed unusually high within-run CVs (>3.5 × median within-run CV). By this criterion, 14 runs across 71 analytes (0.3% of the data), closely matching the number and identity of the runs that were reanalyzed during the course of the project, were removed before calculation of CVs. The associations of between-run CV with retention time, peak area, and absolute difference in peak area (matching A/IS) and the association of CVs with difference in retention time between analytes and all nonmatching IS were assessed by Spearman's correlation. R version 4.0.3 was used for all statistical calculations with package "lme4" for multilevel, hierarchical, regression.

RESULTS AND DISCUSSION

Main Characteristics of Analytical Platforms. The number of analytes included on platforms GC, LC1, LC2, and LC3 was 25, 16, 24, and 6, respectively. The number of A/IS pairings investigated was the square of these numbers (Table 1). The metabolites analyzed on platforms GC and LC1 were mostly of intermediate to high concentrations (>1 μM), whereas metabolites measured on platforms LC2 and LC3 largely covered the nanomolar range. Within each platform, the range of typically observed concentrations of metabolites exceeded 500-fold (Table 1). The concentration of each internal standard is given in Tables S1−S4 in the Supporting Information.

CV Calculations Based On Uncorrected Peak Areas. We calculated within- and between-run CVs for all analytes (A) and internal standards (IS) separately based on peak areas and compared the results to the corresponding matching A/IS. Some details of these calculations are shown in Figure 1 (using data from source material 1). Notably, the detector response, measured as peak areas, was either stable for many consecutive runs or demonstrated a gradual decline or sudden shifts over time. Explanations for such phenomena have been discussed in detail previously.[16] Furthermore, while the variation around mean peak area was fairly constant across consecutive runs, there were also instances of large variation for some runs. In Figure 1, this applies to runs 45−48 for the two analytes 3-methylhistidine and choline. Notably, because changes in peak areas tended to occur in parallel for analytes and matching IS, such variations were largely eliminated for the ratio A/IS, as demonstrated for the two analytes in Figure 1. Within- and between-run CVs for all analytes across the four analytical platforms are provided in Tables S1−S4 in the Supporting Information, and between-run CVs based on source material 1 are presented in Figure 2. A notable finding for all platforms was similar CVs for analytes and their matching IS (Tables S1−S4 and Figure 2). Significantly, within-run CVs for uncorrected peak areas ranged from <1 to >100% for the 60 runs with closely corresponding results for analyte and matching IS (Tables S1−S4). This demonstrated that the factors that affected the variability in peak area for analytes or IS were specific to each compounds chemical structure, a finding that, by itself, underscores the importance of using, as far as possible, structurally identical internal standards. Tables S1−S8 and Figure 2 also demonstrate that CVs on platforms GC, LC1, and LC3 were always lower for matching A/IS than for uncorrected peak areas. A few exceptions were found on platform LC2 were matching A/IS for the low-abundance metabolites thiamin monophosphate, flavin adenin mononucleotide, and anthranilic acid demonstrated higher CVs than their uncorrected IS. We attribute this phenomenon to known stability issues for these analytes in serum/plasma.

Taken together, these analyses demonstrate that within-run CVs, even for uncorrected peak areas, can be low as has been noted by others,[17,18] and in many cases, performance can be maintained for several consecutive runs. However, it is also apparent that analytical drift, either gradual or abrupt, may

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

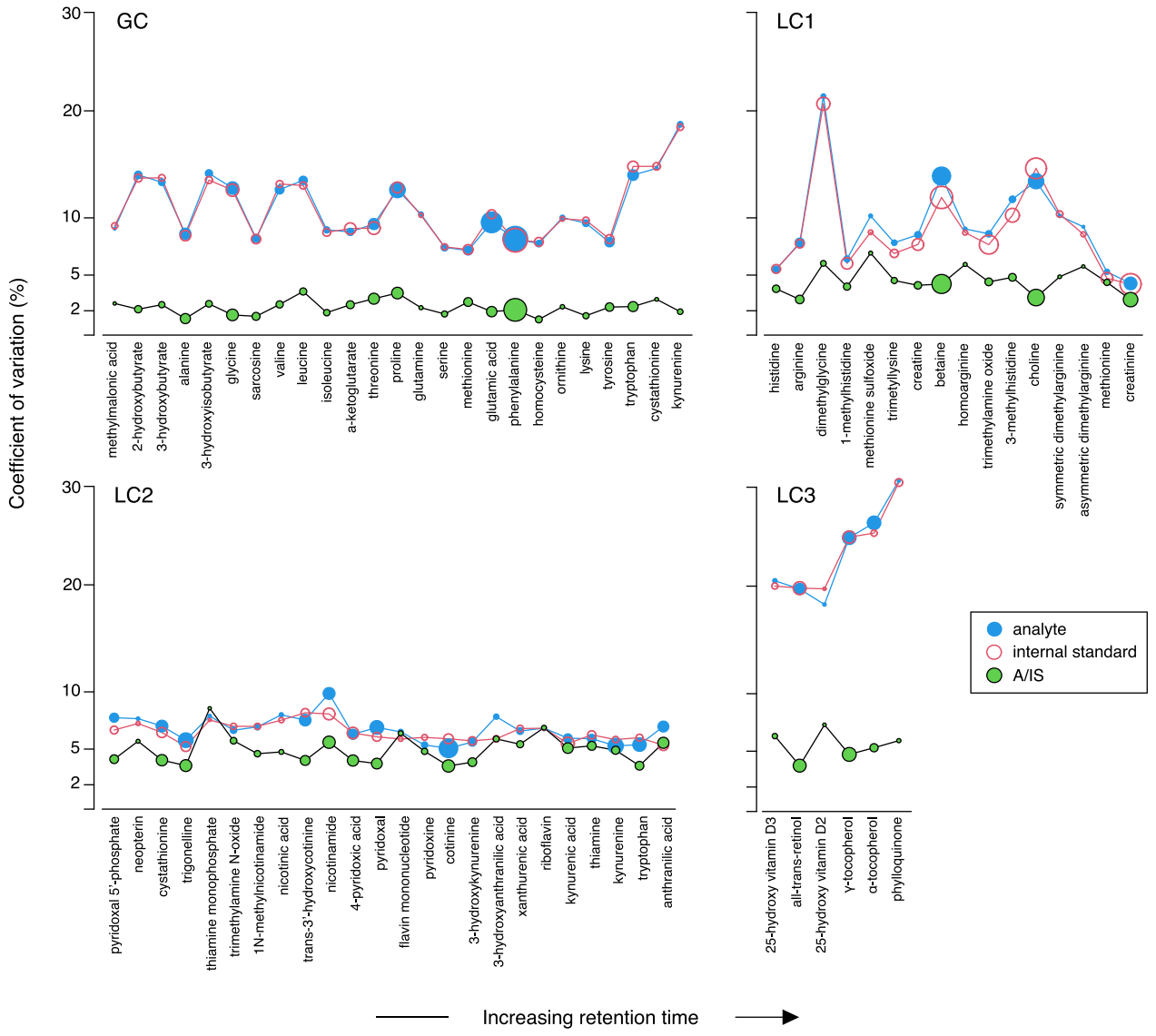


Figure 2. Between-run CVs for analytes and internal standards separately. Results are compared to CVs for the corresponding matching A/IS. For analytes and IS, the size of each symbol reflects the mean peak areas. For A/IS, the size of each symbol reflects the harmonic mean (emphasizing the smaller value) peak area for analytes and IS. The symbols, when interpreted as spheres, represent accurate relative areas within, but not across, platforms. All calculations were based on peak areas from source material 1 (pooled EDTA plasma).

happen even when investigated in a single homogeneous source material using the same analytical equipment and operated by the same personnel on a day to day basis. Thus, as commented previously,[1] accurate correction and normalization of analyte signals become increasingly important when analyzing large sample numbers that necessitate many runs and/or analyses over extended time periods.

Within-Run CVs for Matching A/IS. Within-run CVs for matching A/IS were calculated based on six S1, three S2, and three S3 samples for each run (96-well plate) using a regression model adjusted for the source material. Figure 3 shows the data across all 60 runs for the best and poorest performing analyte on each platform. Generally, we observed little to no trend in terms of decreasing or increasing within-run CVs during the course of the project (7 months) and, importantly, a narrow range of within-run CVs for matching A/IS compared to uncorrected peak areas (Tables S1−S8).

Between-Run CVs for Matching and Nonmatching A/IS. Between-run CVs were calculated using a regression model adjusted for source materials and run and for each source individually (S1, S2, or S3) by adjusting for run. CVs based on all data were generally higher than CVs for any individual source (Table 2). The use of pooled samples for evaluation of precision, in our case, any one of the S1, S2 or S3, has been criticized, as it will not capture the actual variation in ion suppression, or relative matrix effect, that may exist across individual patient/participant samples,[19] However, Matuszewski also suggested that an SIL-IS was able to correct for relative matrix effects within acceptance criteria.[19] Another study found that individual samples regularly exhibited less ion suppression than pooled samples.[16] To cover all possibilities but also to maximize external validity, we chose to report our main results using the combined data of all three pooled serum/plasma sources.

The median CV for matching A/IS (using all data) on each platform varied from 2.7−5.9%, and the median difference in CV between nonmatching and matching A/IS varied from 2.9 to 10.7 percentage points (Table 2). Figure 4 shows the

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

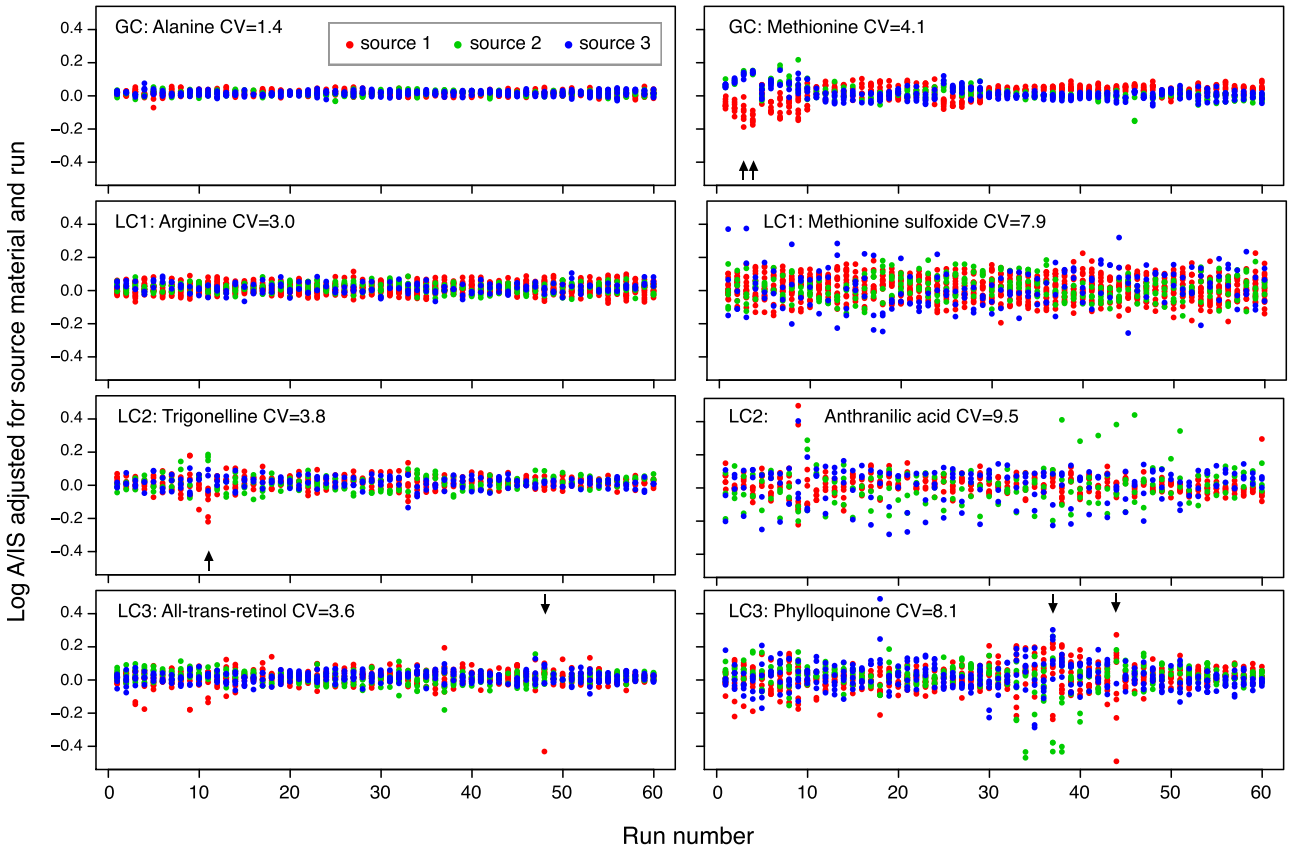


Figure 3. Variation across 60 runs for selected analytes (matching A/IS). The analytes with the lowest (left) and highest (right) between-run CVs within each platform are shown. Data for S1, S2, and S3 are plotted in that order with color-coding as indicated. Arrows indicate runs that had a within-run CV > 3.5 times the median across 60 runs. S1; pooled EDTA plasma; S2; pooled serum spiked for most analytes, S3; pooled serum.

Table 2. Summary of Between-Run Precision[a,d]

		matching A/IS				nonmatching A/IS	
platform	analytes (n)	S1	S2	S3	All	All[b]	CV difference[c]
GC	25	2.3	1.9	2.2	2.7	11.3	8.6
		(1.3, 3.6)	(1.0, 3.8)	(1.2, 4.0)	(1.4, 4.1)	(2.7, 25.8)	(−0.3, 23.6)
LC1	16	4.4	3.7	3.5	4.4	11.5	7.0
		(2.9, 6.9)	(2.3, 6.1)	(2.2, 10.2)	(3.0, 7.9)	(4.3, 31.4)	(0.5, 27.9)
LC2	24	4.9	4.3	4.6	5.5	8.5	2.9
		(3.6, 8.5)	(2.9, 11.8)	(3.1, 9.6)	(3.8, 11.6)	(4.7, 14.2)	(−1.1, 7.6)
LC3	6	5.6	5.5	5.0	5.9	16.0	10.7
		(3.8, 7.3)	(3.0, 7.2)	(2.6, 8.1)	(3.6, 8.1)	(7.6, 28.3)	(1.8, 23.0)

[a]Numbers are median (range) between-run CVs (%) calculated for all data (All) or for individual sources (S1−S3) as indicated. Ranges are across n analytes. [b]Ranges are across all nonmatching A/IS within a platform. [c]Median (range) difference (nonmatching A/IS−corresponding matching A/IS) across all nonmatching A/IS within a platform. [d]Abbreviations: S1, pooled EDTA plasma; S2, pooled serum spiked for most analytes; S3, pooled serum.

detailed results for seven selected analytes within each platform. The analytes were selected based on high structural similarity and/or similar retention time. Left panels show retention times for analytes and their matching internal standards, middle panels show CVs for A/IS pairings, and right panels show CVs versus difference in retention time for analytes versus internal standards. Most, but not all, deuterium-labeled internal standards eluted slightly earlier than their matching analyte. A detailed explanation of this isotope effect can be found in ref 20. As anticipated, no difference in retention time was found for ^{15}N- or ^{13}C-labeled SIL-IS, including ^{15}N-labeled homoarginine (hArg) and ^{13}C-labeled thiamin (Thi) (Figure 4). The variability in retention time measured across 60 runs was greater for the late eluting analytes on all platforms except LC2, possibly related to the step-gradient elution employed on that platform. A feature common to all platforms, but especially notable on LC3, was a marked symmetry around the downward sloping diagonal (containing the matching A/IS, Figure 4, middle panels), and around zero difference in retention time (right panels). This was due to similar precision for "inverse" A/IS pairings: generally, for two analytes, A_1 and A_2, the pairings A_1/A_2-IS and A_2/A_1-IS tended to demonstrate similar CVs.

On GC, the three structural analogs, 2-hydroxybutyrate (aHB), 3-hydroxybutyrate (bHB), and 3-hydroxyisobutyrate (HIB), and the two analogs, leucine and isoleucine, displayed

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

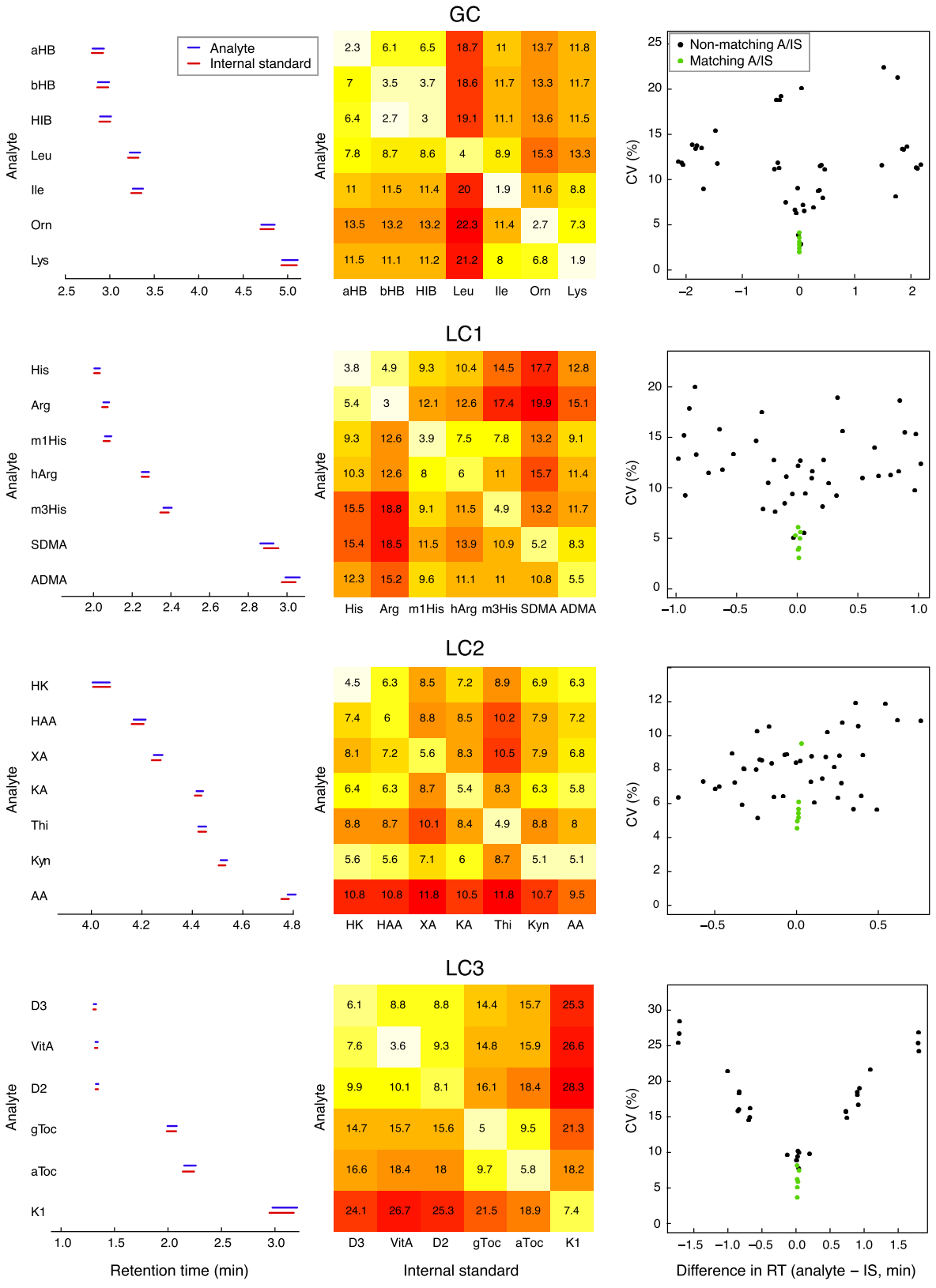


Figure 4. Results for selected analytes on platforms GC through LC3 Left: The mean ± 1 SD retention time (across 60 runs) is shown for analytes and their matching SIL-IS. Middle: CVs for all A/IS pairings. Right: CVs according to difference in retention time (analyte−internal standard). Abbreviations: A, analyte; IS, internal standard; aHB, 2-hydroxybutyrate; bHB, 3-hydroxybutyrate; HIB, 3-hydroxyisobutyrate; Leu, leucine; Ile, isoleucine; Orn, ornithine; Lys, lysine; His, histidine; m1His, 1-methylhistidine; hArg, homoarginine; m3His, 3-methylhistidine; SDMA, symmetric dimethylarginine; ADMA, asymmetric dimethylarginine; HK, 3-hydroxykynurenine; HAA, 3-hydroxyanthranilic acid; XA, xanthurenic acid; KA, kynurenic acid; Thi, thiamine; Kyn, kynurenine; AA, anthranilic acid; D3, 25-hydroxyvitamin D3; VitA, all-*trans*-retinol; D2, 25-hydroxyvitamin D2; gToc, gamma-tocopherol; aToc, alpha-tocopherol; K1, phylloquinone.

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

Table 3. Precision versus Retention Time and Peak Area[a,f]

	nonmatching A/IS		matching A/IS					
	Δ retention time[b]		retention time[c]		peak area[d]		Δ peak area[e]	
platform	rho	*p*-value	rho	*p*-value	rho	*p*-value	rho	*p*-value
GC	0.54	<0.001	−0.08	0.7	−0.15	< 0.001	0.14	<0.001
LC1	0.17	<0.001	0.01	1.0	−0.57	< 0.001	0.20	<0.001
LC2	0.24	<0.001	0.05	0.8	−0.37	< 0.001	0.02	0.6
LC3	0.93	<0.001	−0.09	0.9	−0.48	< 0.001	0.00	0.9

[a]Results of Spearman's correlation analyses. [b]CV vs the absolute difference in retention time between analytes and all nonmatching IS. [c]CV vs the mean retention time for analytes and matching IS. [d]CV vs the harmonic mean (emphasizing the smaller value) peak area of each matching A/IS pair. [e]CV vs the absolute difference in peak between analytes and IS. The two latter correlations were performed with mutual adjustment. [f]Abbreviations: A, analyte; IS, internal standard.

almost identical retention times, whereas ornithine and lysine, which differ by one methylene group, displayed non-overlapping retention times (Figure 4). The median CV difference for nonmatching A/IS within the three groups of structural analogs was 4.0 compared to 8.6 across all analytes on GC.

On LC1, the histidine analogs histidine (His), 1-methyl-histidine (m1His), and 3-methyl-histidine (m3His) displayed non-overlapping retention times. Within this group, CV differences varied from 3.9 to 10.7. The retention times for arginine (Arg) and homoarginine (hArg) were also non-overlapping, and both eluted among the histidines. We found histidine-IS to be the best substitute for arginine-IS and 1-methyl-histidine-IS was the best substitute for homoarginine-IS. The median CV difference for the two groups of structural analogs described above and for symmetric and asymmetric dimethylarginines (SDMA and ADMA) was 5.5 compared to 7.0 across all analytes on LC1.

LC2 contained several groups of similar analytes; however, most of them differed by a functional group or a phosphate group expected to influence both retention time and ionization characteristics.[6] In Figure 4, we included six analytes along the kynurenine pathway of tryptophan degradation and thiamine, which eluted in the middle of this group. All the included analytes had non-overlapping retention times except for kynurenic acid (KA) and thiamine (Thi). As shown, CV differences were low for many nonmatching A/IS, e.g., for KA, four alternative IS had CV differences ≤1 percentage point. Again, there was no obvious pattern to what substitutions might entail the smallest CV difference (Figure 4).

LC3 contained two pairs of structural analogs: vitamins D3 and D2 and gamma- and alpha-tocopherol. Vitamin A (all-*trans*-retinol) coeluted with vitamins D3 and D2, and combinations of D3 and D2 with vitamin A-IS carried modest CV differences of 2.7 and 2.0, respectively, with corresponding inverse combinations less favorable (Figure 4). The median CV difference for the two groups of structural analogs was 3.3 compared to 10.7 across all analytes on LC3.

Analytical Precision According to Retention Time and Peak Area. As indicated in Figure 4 (right panels), there was an association between CV and difference in retention time (analyte−internal standard) on all platforms. We quantified these relations as the correlation of CVs versus the absolute value of retention time differences (Table 3). Corrections of ion suppression effects are believed to depend on the closeness in retention time of analytes and IS.[21] The results for platforms LC1 and LC2 (Spearman's rho values of 0.17 and 0.24, respectively) indicated rather moderate effects, as also found in a previous report.[9] In contrast, we found comparatively strong correlations on platforms GC and LC3 (Spearman's rho values of 0.54 and 0.93, respectively). Notably, these platforms shared a liquid−liquid extraction (LLE) step during sample preparation, and the GC method also comprised a final LLE to extract the derivatized products. As reviewed by Wieling, an important reason for inclusion of a (processed) internal standard in chromatography was (and remains) to correct for phase transfer during sample processing.[22] We therefore speculate that small differences in the partition between liquid phases, in particular, for the fat-soluble analytes on LC3, may have impacted precision in proportion to the difference in retention time between analytes and IS. Deuterium-labeled internal standards often do not coelute exactly with the analyte, especially in LC−MS/MS applications, and are, for this reason, generally considered nonperfect internal standards.[20,23,24] Consequently, higher CVs for LC−MS/MS, compared to GC−MS/MS platforms, might be related to differences in the extent of correction for matrix effects. For platforms LC1 and LC2, this may have reduced the association of CV with differences in retention time as well as overall CV difference for nonmatching A/IS.

We also examined the relation between CV and retention time per se for matching A/IS pairs and found essentially no association across all platforms (Table 3). Note that the bases for this and the following analyses are also depicted graphically in Figure 2. There was a negative association of CV with peak area (meaning that larger areas gave rise to smaller CVs) for all four platforms. This underscores the importance of accurate peak integration for precision measures. Of note, the precision for matching A/IS on LC2 was comparable to those on LC1 despite the 500-fold difference in typical metabolite concentrations. We attribute this to a better signal/noise ratio of the LC−MS/MS system employed on LC2 compared to LC1, resulting in comparable peak areas (Table 1). Some researchers have found increased ion suppression with increasing concentration difference between analytes and internal standards.[25−27] In our data, we found moderately higher CVs for analytes with a larger difference in peak area for analytes and IS within platforms GC and LC1 but no significant association within platforms LC2 and LC3 (Table 3).

Benefits of Dedicated SIL-IS beyond Correction for Ion Suppression. We excluded a few analytes on platform GC that were found to be unstable at some stages of the sample preparation. One such example was histidine. Because histidine and histidine-IS degraded at the same rate, the between-run CV for this analyte was low (3.1%). For the same reason, histidine-IS paired with other analytes on GC produced CVs in excess of 20% in all cases. It should be noted, though, that different stability of analytes and internal standards have

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

been reported, and in such cases, correction by the IS led to increased CV.[6]

Isotope dilution GC−MS (as used on GC) has been termed a "nearly matrix-independent reference method",[5] and as expected, the precision for matching A/IS on this platform was the highest among the four analytical platforms. Still, the median CV difference for nonmatching IS on GC was the second highest. As discussed above, explanations for this may be more related to chromatographic effects during sample preparation than to differences in ion suppression.

The inclusion of a dedicated SIL-IS for all analytes has some practical benefits that are worth mentioning: In chromatographic settings that involve dozens of analytes eluting during a short time span, correct peak identification is essential, especially for low abundance analytes. An SIL-IS eluting at the same time, or within a short, constant, interval facilitates accurate peak identification, and hence both automatically and manually controlled peak integration.

Parameters Not Addressed in This Study. Ionization by atmospheric chemical ionization (APCI) is known to be less prone to ion suppression than electrospray ionization (ESI).[5] We were unable to compare APCI and ESI directly; however, we noted that CVs for matching A/IS were the highest on the platform that utilized APCI (LC3). Possible explanations could include the challenge of obtaining chromatographic separation of matrix components from the low abundance, fat-soluble, metabolites analyzed on this platform but also the effect of peak area integration on CVs. Across all platforms, we employed sample preparation that mainly consisted of solvent extraction/deproteinization by ethanol or acid. Additional purification, e.g., by LLE or solid phase extraction (SPE), may remove more phospholipids and therefore reduce ion suppression.[21,28,29] We employed a shared LLE step on platforms GC and LC3 mainly for the purpose of more efficient utilization of precious source materials. As discussed above, there was no suggestion or any way to verify a benefit of LLE in terms of improving precision (for matching A/IS) on LC3. A recognized drawback of SPE that also applies to LLE is that it will often fail to remove substances that coelute with the analyte during chromatography.[21,30] In some cases, it may even lead to increased ion suppression.[31] Several previous studies noted that a lack of SIL-IS might also impair accuracy, linearity, and lower limit of quantitation (LLOQ).[6,9,17,19] Due to the post-hoc nature of this study, we were unable to evaluate such measures.

CONCLUSIONS

In this study, we quantified the difference between using a nonmatching A/IS pair compared to a matching A/IS pair on analytical precision in four MS-based analytical platforms. Median differences ranged from small to moderate (3−11 percentage points) with wide variation (−1 to 28 percentage points) within and across platforms. In cases characterized by high structural similarity or closeness of retention time between analytes and internal standards, differences tended to be smaller. Larger CV differences for A/IS pairs with increasing difference in retention time was found but mainly in the two platforms that included a liquid−liquid extraction step during sample preparation. The latter finding is a reminder of the importance of close similarity of analytes and IS to ensure similar behavior during all steps of sample preparation. Although not specifically addressed in this study, it should be noted that inclusion of dedicated SIL-IS may also positively impact other quality measures including accuracy, linearity, and LLOQ as well as correct peak identification in multianalyte quantitative mass spectrometry applications.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.1c00119.

(Tables S1−S4) CVs for analytes and internal standards based on uncorrected peak areas on platforms GC, LC1, LC2, and LC3 and (Tables S5−S8) CVs for matching and nonmatching A/IS on platforms GC, LC1, LC2, and LC3 (PDF)

AUTHOR INFORMATION

Corresponding Author

Arve Ulvik − *Bevital, 5021 Bergen, Norway*; orcid.org/0000-0003-1574-9690; Email: arve.ulvik@bevital.no

Authors

Adrian McCann − *Bevital, 5021 Bergen, Norway*

Øivind Midttun − *Bevital, 5021 Bergen, Norway*

Klaus Meyer − *Bevital, 5021 Bergen, Norway*

Keith M. Godfrey − *MRC Lifecourse Epidemiology Unit and NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton SO17 1BJ, United Kingdom*

Per M. Ueland − *Bevital, 5021 Bergen, Norway*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c00119

Author Contributions

Ø.M. and A.U. conceived and developed the main ideas for the statistical analyses. A.U. performed statistical analyses and wrote the manuscript. A.M., Ø.M., K.M., K.M.G., and P.M.U. reviewed the manuscript for important intellectual content.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Gry Kvalheim, Marit Krokeide, Randi Heimdal, and Ove Aarseth for pertinent information during the analysis and write-up of the manuscript.

REFERENCES

(1) Gertsman, I.; Barshop, B. A. *J. Inherited Metab. Dis.* **2018**, *41*, 355−366.

(2) Shi, X.; Wang, S.; Jasbi, P.; Turner, C.; Hrovat, J.; Wei, Y.; Liu, J.; Gu, H. *Anal. Chem.* **2019**, *91*, 13737−13745.

(3) Zha, H.; Cai, Y.; Yin, Y.; Wang, Z.; Li, K.; Zhu, Z.-J. *Anal. Chem.* **2018**, *90*, 4062−4070.

(4) Zhong, F.; Xu, M.; Zhu, J. *J. Chromatogr., B* **2019**, *1120*, 80−88.

(5) Vogeser, M.; Seger, C. *Clin. Chem.* **2010**, *56*, 1234−1244.

(6) Stokvis, E.; Rosing, H.; Beijnen, J. H. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 401−407.

(7) Townsend, M. K.; Clish, C. B.; Kraft, P.; Wu, C.; Souza, A. L.; Deik, A. A.; Tworoger, S. S.; Wolpin, B. M. *Clin. Chem.* **2013**, *59*, 1657−1667.

(8) Bornø, A.; van Hall, G. *J. Chromatogr., B* **2014**, *951-952*, 69−77.

(9) Krautbauer, S.; Büchler, C.; Liebisch, G. *Anal. Chem.* **2016**, *88*, 10957−10961.

(10) Lanckmans, K.; Sarre, S.; Smolders, I.; Michotte, Y. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1187−1195.

(11) Godfrey, K. M.; Cutfield, W.; Chan, S. Y.; Baker, P. N.; Chong, Y. S.; NiPPeR, S. G. *Trials* **2017**, *18*, 131.
(12) Midttun, Ø.; McCann, A.; Aarseth, O.; Krokeide, M.; Kvalheim, G.; Meyer, K.; Ueland, P. M. *Anal. Chem.* **2016**, *88*, 10427−10436.
(13) Midttun, Ø.; Kvalheim, G.; Ueland, P. M. *Anal. Bioanal. Chem.* **2013**, *405*, 2009−2017.
(14) Midttun, Ø.; Hustad, S.; Ueland, P. M. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1371−1379.
(15) Kirkwood, T. B. L. *Biometrics* **1979**, *35*, 908−909.
(16) Tan, A.; Hussain, S.; Musuku, A.; Massé, R. *J. Chromatogr., B* **2009**, *877*, 3201−3209.
(17) Ask, K. S.; Lid, M.; Øiestad, E. L.; Pedersen-Bjergaard, S.; Gjelstad, A. *J. Chromatogr., A.* **2019**, *1602*, 117−123.
(18) Tokuoka, S. M.; Yasumoto, A.; Kita, Y.; Shimizu, T.; Yatomi, Y.; Oda, Y. *Rapid Commun. Mass Spectrom.* **2020**, *34*, No. e8814.
(19) Matuszewski, B. K. *J. Chromatogr., B* **2006**, *830*, 293−300.
(20) Berg, T.; Karlsen, M.; Øiestad, Å. M.; Johansen, J. E.; Liu, H.; Strand, D. H. *J. Chromatogr., A.* **2014**, *1344*, 83−90.
(21) Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M. *Talanta* **2013**, *115*, 104−122.
(22) Wieling, J. *Chromatographia* **2002**, *55*, S107−S113.
(23) Ciccimaro, E.; Blair, I. A. *Bioanalysis* **2010**, *2*, 311−341.
(24) Wang, S.; Cyronak, M.; Yang, E. *J. Pharm. Biomed. Anal.* **2007**, *43*, 701−707.
(25) Liang, H. R.; Foltz, R. L.; Meng, M.; Bennett, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2815−2821.
(26) Remane, D.; Wissenbach, D. K.; Meyer, M. R.; Maurer, H. H. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 859−867.
(27) Sojo, L. E.; Lum, G.; Chee, P. *Analyst* **2003**, *128*, 51−54.
(28) Sitnikov, D. G.; Monnin, C. S.; Vuckovic, D. *Sci. Rep.* **2016**, *6*, 38885.
(29) Tulipani, S.; Llorach, R.; Urpi-Sarda, M.; Andres-Lacueva, C. *Anal. Chem.* **2013**, *85*, 341−348.
(30) Hewavitharana, A. K. *J. Chromatogr., A* **2011**, *1218*, 359−361.
(31) Gosetti, F.; Mazzucco, E.; Zampieri, D.; Gennaro, M. C. *J. Chromatogr., A* **2010**, *1217*, 3929−3937.

https://doi.org/10.1021/acs.analchem.1c00119
Anal. Chem. 2021, 93, 7616−7624

Supporting information for:

Quantifying precision loss in targeted metabolomics based on mass spectrometry and nonmatching internal standards

Authors: Arve Ulvik, Adrian McCann, Øivind Midttun, Klaus Meyer, Keith M Godfrey[1], Per M Ueland

Bevital, laboratoriebygget, 9 etg. Jonas Lies veg 87, 5021 Bergen, Norway

[1] MRC Lifecourse Epidemiology Unit and NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust

List of contents:

Table S1. CVs for analytes and internal standards based on uncorrected peak areas on platform GC

			CV (%)			
			within-run[c]		between-run	
Analyte	IS[a]	conc. (µM)[b]	analyte	internal standard	analyte	internal standard
methylmalonic acid	$^2H_3$	1.2	6.8 (3.9, 31)	6.7 (3.8, 34)	8.7	8.7
2-Hydroxybutyrate	$^2H_3$	50	6.3 (3.2, 103)	5.8 (2.4, 103)	12.3	12.0
3-Hydroxybutyrate	$^2H_2$	70	5.9 (3.2, 105)	6.1 (3.4, 105)	11.9	12.1
alanine	$^2H_3$	182	5.4 (2.7, 26)	5.4 (2.7, 27)	10.3	10.2
3-Hydroxyisobutyrate	$^2H_5$	32	6.6 (3.8, 103)	5.9 (3.4, 104)	12.3	11.7
glycine	$^2H_2$	90	5.5 (1.8, 101)	5.1 (2.2, 99)	11.0	10.9
sarcosine	$^2H_3$	2.7	6.1 (2.8, 24)	5.8 (2.3, 25)	8.0	7.9
valine	$^2H_2$	9.5	5.2 (1.9, 99)	6.1 (2.6, 98)	11.9	12.7
leucine	$^2H_{10}$	50	6.0 (3.5, 104)	5.2 (2.3, 106)	12.2	19.3
isoleucine	$^2H_{10}$	61	6.2 (3.3 32)	6.0 (3.2, 32)	8.5	8.3
a-Ketoglutarate	$^2H_6$	64	6.9 (2.9, 34)	7.2 (3.8, 34)	9.4	9.6
threonine	$^2H_2$	87	6.6 (3.3, 25)	5.8 (2.4, 28)	9.5	8.6
proline	$^2H_3$	48	5.2 (2.3, 100)	5.3 (2.5, 101)	11.1	11.2
glutamine	$^2H_5$	239	9.0 (3.2, 42)	9.0 (3.9, 42)	11.6	11.7
serine	$^2H_3$	68	5.1 (2.2, 26)	5.2 (1.7, 27)	7.1	7.1
methionine	$^2H_4$	23	5.6 (3.4, 24)	5,7 (2.8, 24)	9.2	8.4

glutamic acid	$^{2}H_5$	32	8.2 (3.8, 28)	8.9 (5.0, 27)	10.7	11.1
phenylalanine	$^{2}H_8$	31	5.2 (1.7, 34)	5.0 (2.1 33.4)	8.3	8.1
homocysteine	$^{2}H_4$	30	5.0 (2.0, 26)	5.5 (2.6 25.4)	8.0	7.7
ornithine	$^{2}H_3$	42	7.9 (3.2, 31)	7.6 (3.6, 30)	10.7	10.3
lysine	$^{2}H_9$	40	7.3 (3.0, 31)	7.6 (3.3, 32)	9.2	9.6
tyrosine	$^{2}H_2$	230	6.0 (2.6, 28)	5.8 (2.7, 29)	7.7	7.7
tryptophan	$^{2}H_5$	36	12.1 (4.3, 31)	13.1 (4.7, 35)	14.1	15.6
cystathionine	$^{2}H_4$	2.3	13.7 (4.8, 35)	13.8 (4.2, 35)	16.9	16.5
kynurenine	$^{2}H_3$	1.8	15.0 (4.4, 75)	14.0 (4.4 76)	19.0	18.6

[a]Number and type of stable isotope substitutions for internal standard. [b]concentration of internal standard. [c]median (range) across 60 runs. Abbreviations IS, internal standard.

Table S2. CVs for analytes and internal standards based on uncorrected peak areas on platform LC1

			CV (%)			
			within-run[c]		between-run	
Analyte	IS[a]	conc. (µM)[b]	analyte	internal standard	analyte	internal standard
histidine	$^2H_3$	66	3.6 (1.9, 18.5)	3.8 (2.0, 16.4)	6.0	6.3
arginine	$^2H_3$	30	3.2 (1.8, 33.3)	3.3 (1.6, 33)	8.4	8.4
dimethylglycine	$^2H_2$	5	6.0 (3.4, 118)	3.9 (2.0, 114)	22.5	22
1-methylhistidine	$^2H_3$	20	4.9 (2.3, 19.3)	4.4 (2.2, 19.3)	7.0	6.7
methionine sulfoxide	$^2H_5$	5	8.4 (4.6, 28)	7.2 (3.9, 24)	11.6	9.5
trimethyllysine	$^2H_2$	90	4.8 (2.5, 27)	4.1 (2.1, 25)	8.4	7.3
creatine	$^2H_3$	50.6	5.3 (3.1 28)	5.5 (2.3, 25)	9.0	8.3
betaine	$^2H_2$	20	3.9 (1.9, 63)	3.3 (1.6, 52)	14.5	12.3
homoarginine	$^2H_{10}$	1.25	7.5 (3.8, 19.1)	7.3 (3.6, 21)	9.2	8.9
trimethylamine N-oxide	$^2H_{10}$	35.6	4.4 (2.8, 31)	3.6 (1.4, 29)	9.0	8.2
3-methylhistidine	$^2H_6$	14.8	7.1 (4.0, 47)	6.3 (2.9, 41)	13.2	12.0
choline	$^2H_2$	10	3.6 (1.7, 60)	3.3 (1.9, 67)	14.5	15.5
symmetric dimethylarginine	$^2H_3$	2	11.6 (6.5, 30)	12.6 (3.7, 45)	14.8	16.5
asymmetric dimethylarginine	$^2H_5$	1	10.5 (3.7, 20)	10.3 (4.8, 19.5)	12.1	12.0
methionine	$^2H_3$	70	4.1 (2.3, 23)	4.0 (1.7, 16.9)	5.1	4.9
creatinine	$^2H_4$	94	4.0 (1.3, 9.9)	3.8 (0.7, 283)	4.4	19.4

[a]Number and type of stable isotope substitutions for internal standard. [b]concentration of internal standard. [c]median (range) across 60 runs. Abbreviations IS, internal standard.

Table S3. CVs for analytes and internal standards based on uncorrected peak areas on platform LC2

			CV (%)			
			within-run[c]		between-run	
Analyte	IS[a]	conc. (nM)[b]	analyte	internal standard	analyte	internal standard
pyridoxal 5'-phosphate	$^{2}H_{2}$	100	5.7 (2.6, 23)	5.4 (1.5, 18.7)	10.3	8.8
neopterin	$^{15}N_{1}$	25	6.4 (3.3, 19.9)	6.1 (3.2, 18.8)	8.7	8.3
cystathionine	$^{2}H_{4}$	200	6.1 (2.6, 15.4)	6.1 (2.1, 16.0)	8.5	8.6
trigonelline	$^{2}H_{3}$	1000	4.1 (2.2, 14.5)	4.4 (1.4, 15.3)	6.6	6.7
thiamine monophosphate	$^{2}H_{3}$	10	7.6 (3.3, 21)	6.6 (3.4, 19.2)	12.2	8.1
trimethylamine N-oxide	$^{2}H_{9}$	10	5.3 (2.0, 21)	5.9 (1.5, 22)	7.6	8.6
1N-methylnicotinamide	$^{2}H_{4}$	250	5.5 (2.2, 21)	5.9 (1.5, 22	9.7	8.6
nicotinic acid	$^{2}H_{4}$	100	6.7 (3.3, 20)	6.1 (2.3, 22)	9.7	8.7
trans-3'-hydroxycotinine	$^{2}H_{3}$	100	6.3 (1.9, 27)	6.6 (2.6, 30)	9.4	10.3
nicotinamide	$^{2}H_{4}$	250	7.4 (2.4, 27)	6.6 (2.9, 29)	11.1	9.7
4-pyridoxic acid	$^{2}H_{2}$	100	4.9 (1.9, 19.2)	5.1 (2.5 19.8)	7.1	8.3
pyridoxal	$^{2}H_{3}$	25	5.1 (2.7, 20)	4.7 (1.7, 24)	7.2	7.3
flavin mononucleotide	$^{2}H_{7}$	30	6.1 (2.6, 22)	5.1 (1.2, 14.8)	9.1	6.4
pyridoxin	$^{2}H_{4}$	20	4.3 (1.5, 17)	5.5 (1.7, 19.4)	6.3	6.9
cotinine	$^{2}H_{3}$	200	4.3 (1.4, 16.1)	5.0 (1.5, 19.0)	6.2	6.7
3-hydroxykynurenine	$^{2}H_{2}$	50	4.3 (1.4, 18.2)	4.8 (1.4, 18.0)	7.3	6.7

3-hydroxyanthranilic acid	$^{2}H_{2}$	50	6.1 (2.3, 27)	4.7 (1.6, 18.7)	8.8	7.1
xanthurenic acid	$^{2}H_{4}$	30	5.9 (1.7, 17.6)	5.8 (1.6, 24)	8.2	8.1
riboflavin	$^{2}H_{8}$	30	5.7 (2.7, 17.6)	5.8 (3.0, 19.1)	7.5	7.4
kynurenic acid	$^{2}H_{5}$	100	4.8 (1.5, 14.6)	4.5 (1.7, 14.5)	6.4	6.3
thiamine	$^{13}C_{4}$	10	5.0 (2.1, 18.4)	5.1 (1.1, 18.3)	9.2	9
kynurenine	$^{2}H_{4}$	200	4.5 (1.8, 13.6)	4.9 (1.6, 14.8)	5.7	6.5
tryptophan	$^{2}H_{5}$	2000	4.6 (1.2, 15.1)	4.3 (2.0, 18.9)	6.2	6.5
anthranilic acid	$^{2}H_{4}$	25	6.1 (2.3, 33.1)	4.4 (1.4, 18.1)	10.0	6.2

[a]Number and type of stable isotope substitutions for internal standard. [b]concentration of internal standard. [c]median (range) across 60 runs. Abbreviations IS, internal standard.

Table S4. CVs for analytes and internal standards based on uncorrected peak areas on platform LC3

			CV (%)			
			within-run[c]		between-run	
Analyte	IS[a]	conc. (nM)[b]	analyte	internal standard	analyte	internal standard
25-hydroxy vitamin D3	$^2H_6$	112	13.6 (6.4, 51)	15.0 (5.5, 55)	19.0	18.6
all-trans-retinol	$^2H_6$	1850	15.4 (6.7, 70)	15.9 (6.8, 69)	18.6	18.7
25-hydroxy vitamin D2	$^2H_6$	147	14.6 (5.3, 50)	14.8 (6.0, 56)	17.5	18.3
γ-tocopherol	$^2H_9$	4500	15.2 (5.2, 141)	15.2 (6.1, 137)	24.0	24.1
α-tocopherol	$^2H_9$	21000	15.5 (6.4, 136)	14.8 (8.3, 128)	26.1	25.2
phylloquinone	$^2H_4$	3.6	18.4 (6.9, 136)	18.7 (7.2, 153)	31.2	33.7

[a]Number and type of stable isotope substitutions for internal standard. [b]concentration of internal standard. [c]median (range) across 60 runs. Abbreviations IS, internal standard.

Table S5. CVs for matching and nonmatching A/IS on platform GC[a]

	matching A/IS					nonmatching A/IS	
	within-run	between-run					
Analyte	All[b]	S1	S2	S3	All	All[c]	CV difference[d]
methylmalonic acid	2.4 (1.6, 4.5)	2.6	1.7	3.1	2.8	9.7 (5.3, 19.8)	6.8 (2.5, 17.0)
2-hydroxybutyrate	1.7 (0.9, 3.9)	2.1	1.9	1.8	2.3	11.0 (6.0, 17.7)	8.8 (3.7, 15.4)
3-hydroxybutyrate	1.9 (0.7, 8.8)	2.5	2.2	3.8	3.4	11.8 (3.7, 17.7)	8.4 (0.3, 14.4)
alanine	1.2 (0.7, 2.3)	1.4	1	1.2	1.4	10.5 (8.0, 21)	9.0 (6.5, 19.4)
3-hydroxyisobutyrate	2.4 (1.3, 5.8)	2.6	2.3	2.6	2.9	11.6 (2.7, 18.1)	8.7 (-0.3, 15.2)
glycine	1.3 (0.6, 3.6)	1.7	1.2	1.5	1.7	10.7 (5.0, 18.1)	8.9 (3.3, 16.3)
sarcosine	1.3 (0.6, 4.1)	1.5	1.3	1.5	1.6	9.2 (5.4, 19.9)	7.6 (3.7, 18.3)
valine	2.2 (0.8, 5.6)	2.5	2.5	2.7	2.7	12.6 (6.8, 18.9)	9.9 (4.1, 16.2)
leucine	3.4 (1.4, 7.0)	3.6	3.8	4	4	11.9 (7.3, 19.3)	7.9 (3.3, 15.3)
isoleucine	1.5 (0.5, 3.5)	1.8	1.4	1.6	1.8	10.4 (4.9, 19.2)	8.5 (3.1, 17.3)
a-Ketoglutarate	2.3 (1.1, 4.9)	2.5	1.9	3.1	3.6	10.8 (7.0, 20.2)	7.2 (3.5, 16.6)
threonine	2.4 (1.1, 6.3)	3.0	2.9	2.8	3.3	10.9 (8.3,22.0)	7.7 (5.1, 18.3)
proline	2.8 (1.1, 8.4)	3.5	3	3.5	3.7	11.3 (5.7, 18.5)	7.6 (2.0, 14.4)
glutamine	2.1 (1.1, 4.9)	2.3	2.5	2	2.7	13.3 (8.0, 24)	10.6 (5.2, 21)
serine	1.3 (0.7, 3.5)	1.7	1.2	1.2	1.7	8.3 (5.2, 19.9)	6.6 (3.5, 18.1)
methionine	2.2 (0.7, 6.6)	2.7	2.2	2.1	4.1	8.4 (6.3, 17.7)	4.4 (2.3, 13.6)

glutamic acid	1.7 (0.9, 3.4)	1.9	1.8	1.5	2.7	10.7 (7.4, 21)	8.0 (4.7, 18.3)
phenylalanine	1.5 (0.8, 7.0)	2.1	2.4	2.5	3	11.0 (5.7, 18.7)	8.1 (2.8, 15.7)
homocysteine	1.4 (0.7, 3.3)	1.3	1.8	1.6	2.7	9.3 (5.6, 19.4)	6.7 (3.0, 16.7)
ornithine	2.2 (1.2, 4.9)	2.3	2.3	2.2	2.7	11.2 (7.3, 22)	8.5 (4.6, 18.8)
lysine	1.4 (0.4, 3.7)	1.6	1.5	1.5	1.9	8.4 (5.4, 20)	6.5 (3.5, 18.4)
tyrosine	2.1 (1.2, 6.4)	2.3	2.3	2.7	2.8	8.7 (4.5, 19.4)	6.0 (1.7, 16.7)
tryptophan	1.6 (0.5, 5.7)	2.3	2.3	2.7	2.6	12.6 (8.4, 23)	10.0 (5.8, 21)
cystathionine	2.3 (1.2, 4.6)	2.9	1.8	2.1	2.9	15.7 (9.0, 25)	12.8 (6.1, 22)
kynurenine	1.9 (1.0, 3.5)	1.9	1.9	2.2	2.3	18.1 (8.7, 26)	15.8 (6.4, 24)

[a]Numbers are CV (%) based on all data (All) or specific source as indicated. [b]Median (range) across 60 sets. [c]Median (range) across all nonmatching A/IS. [d]Median (range) difference nonmatching A/IS vs. matching A/IS across all nonmatching A/IS. Abbreviations A, analyte, IS, internal standard. S1, pooled EDTA plasma. S2, pooled serum spiked for most analytes. S3, pooled serum.

Table S6. CVs for matching and nonmatching A/IS on platform LC1[a]

	matching A/IS					nonmatching A/IS	
	within-run	between-run					
Analyte	All[b]	S1	S2	S3	All	All[c]	CV difference[d]
histidine	3.5 (1.9, 5.8)	3.8	3.3	3.2	3.8	12.5 (4.9, 27)	4.4 (2.9, 5.9)
arginine	2.6 (1.5, 4.4)	2.9	2.3	2.2	3.0	15.1 (5.4, 31)	2.6 (1.5, 4.1)
dimethylglycine	5.5 (3.8, 7.7)	6.0	4.5	4.7	5.9	19.6 (10.3, 31)	4.7 (3.9, 7.6)
1-methylhistidine	3.7 (2.5, 5.7)	4.0	3.3	3.2	3.9	9.1 (5.1, 20)	2.9 (1.4, 5.1)
methionine sulfoxide	7.1 (5.3, 11.4)	6.9	6.0	10.5	7.9	12.1 (9.2, 23)	0.9 (-0.2, 2.6)
trimethyllysine	4.0 (2.6, 6.8)	4.5	3.5	3.8	4.4	9.9 (4.8, 20)	0.3 (-1.1, 2.7)
creatine	3.9 (2.7, 5.1)	4.1	3.3	3.2	4.0	9.3 (5.5, 21)	3.5 (2.6. 6.0)
betaine	3.8 (2.2, 9.4)	4.2	4.5	3.6	4.5	11.8 (5.2, 24)	3.5 (2.6, 6.0)
homoarginine	5.8 (3.1, 8.6)	5.9	6.1	5.9	6.0	11.0 (8.0, 21)	3.6 (0.4, 7.3)
trimethylamine N-oxide	4.0 (2.9, 5.2)	4.4	2.9	3.4	4.1	9.9 (6.0, 21)	3.3 (1.5, 5.9)
3-methylhistidine	4.4 (2.9, 7.0)	4.8	3.9	3.9	4.9	11.5 (8.0, 23)	1.7 (0.1, 3.6)
choline	2.8 (1.7, 5.4)	3.1	2.6	2.6	3.2	11.0 (5.3, 23)	2.7 (0.5, 4.4)
symmetric dimethylarginine	4.8 (3.1, 6.8)	4.9	5.1	4.6	5.2	13.9 (8.3, 24	3.9 (2.4, 5.8)
asymmetric dimethylarginine	5.2 (3.7, 8.0)	5.7	4.3	4.6	5.5	11.3 (9.6, 21)	2.1 (-0.5, 3.3)
methionine	4.1 (2.8, 7.0)	4.4	4.4	3.2	4.5	11.4 (7.2, 24)	3.0 (0.5, 4.6)
creatinine	2.6 (1.2, 5.3)	2.9	2.6	2.5	3.1	9.0 (4.3, 22)	5.9 (1.2, 18.7

[a]Numbers are CV (%) based on all data (All) or specific source as indicated. [b]Median (range) across 60 sets. [c]Median (range) across all nonmatching A/IS. [d]Median (range) difference nonmatching A/IS vs. matching A/IS across all nonmatching A/IS. Abbreviations A, analyte, IS, internal standard. S1, pooled EDTA plasma. S2, pooled serum spiked for most analytes. S3, pooled serum.

Table S7. CVs for matching and nonmatching A/IS on platform LC2[a]

	matching A/IS					nonmatching A/IS	
	within-run	between-run					
Analyte	All[b]	S1	S2	S3	All	All[c]	CV difference[d]
pyridoxal 5'-phosphate	3.5 (1.2, 12.3)	4.1	4.0	4.5	5.7	10.1 (8.6, 11.5)	4.4 (2.9, 5.9)
neopterin	5.3 (2.8, 10.7)	5.7	4.3	6.5	6.6	9.2 (8.1, 10.7)	2.6 (1.5, 4.1)
cystathionine	3.4 (1.4, 8.5)	4.1	3.6	4.1	4.5	9.3 (8.4, 12.1)	4.7 (3.9, 7.6)
trigonelline	2.9 (1.4, 8.6)	3.6	2.9	3.1	3.8	6.7 (5.2, 8.8)	2.9 (1.4, 5.1)
thiamine monophosphate	7.1 (2.8, 29)	8.5	8.9	7.8	11.6	12.5 (11.4, 14.2)	0.9 (-0.2, 2.6)
trimethylamine N-oxide	4.2 (2.3, 21)	5.7	4.9	5.8	6.9	7.2 (5.8, 9.6)	0.3 (-1.1, 2.7)
n-methylnicotinamide	4.0 (1.6, 9.4)	4.6	4.3	3.7	5.5	9.0 (5.5, 12.5)	3.5 (2.6. 6.0)
nicotinic acid	4.0 (2.3, 7.9)	4.8	3.7	4.1	5.4	8.9 (8.0, 11.4)	3.5 (2.6, 6.0)
trans-3'-hydroxycotinine	3.6 (1.4, 7.6)	4.0	3.4	3.5	4.5	8.2 (4.9, 11.8)	3.6 (0.4, 7.3)
nicotinamide	4.2 (1.6, 15.7)	5.6	4.0	5.2	6.1	9.4 (7.6, 12.0)	3.3 (1.5, 5.9)
4-pyridoxic acid	3.6 (1.4, 8.7)	4.0	4.4	4.0	5.8	7.5 (6.0, 9.5)	1.7 (0.1, 3.6)
pyridoxal	3.5 (2.0, 7.3)	3.8	3.9	3.7	5.1	7.8 (5.6, 9.4)	2.7 (0.5, 4.4)
flavin mononucleotide	5.7 (3.0, 18.3)	6.4	7.4	7.3	8.1	12.0 (10.5, 14.0)	3.9 (2.4, 5.8)
pyridoxin	4.5 (2.2, 7.6)	4.8	4.4	4.7	5.2	7.3 (4.7, 8.5)	2.1 (-0.5, 3.3)
cotinine	3.0 (1.2, 7.4)	3.6	3.4	3.8	4.2	7.2 (4.8, 8.8)	3.0 (0.5, 4.6)
3-hydroxykynurenine	3.2 (1.3, 7.8)	3.9	4.3	3.4	4.5	7.8 (6.3, 9.7)	3.4 (1.8, 5.2)

3-hydroxyanthranilic acid	4.6 (1.9, 13.2)	5.9	5.1	5.1	6.0	8.8 (6.8, 10.2)	2.8 (0.8, 4.1)
xanthurenic acid	4.8 (1.7, 10.1)	5.4	4.8	4.7	5.7	8.5 (6.8, 11.3)	2.8 (1.1, 5.6)
riboflavin	6.0 (2.2, 11.5)	6.8	5.8	5.6	7.1	8.3 (7.1, 10.9)	1.2 (0.0, 3.8)
kynurenic acid	4.4 (1.8, 13.9)	5.1	4.6	5.5	5.4	7.4 (5.8, 9.8)	2.0 (0.4, 4.4)
thiamine	3.5 (1.8, 9.9)	5.3	3.5	3.9	4.9	9.6 (8.0, 11.8)	4.7 (3.1, 6.9)
kynurenine	3.6 (1.9, 13.6)	4.9	4.2	4.9	5.1	7.1 (5.1, 8.7)	2.0 (-0.1, 3.6)
tryptophan	3.1 (1.5, 6.8)	3.6	3.7	3.2	3.9	7.4 (5.1, 9.0)	3.5 (1.2, 5.2)
anthranilic acid	7.4 (2.5, 24)	5.5	11.8	9.6	9.5	11.4 (10.1, 13.3)	1.9 (0.6, 3.8)

[a]Numbers are CV (%) based on all data (All) or specific source as indicated. [b]Median (range) across 60 sets. [c]Median (range) across all nonmatching A/IS. [d]Median (range) difference nonmatching A/IS vs. matching A/IS across all nonmatching A/IS. Abbreviations A, analyte, IS, internal standard. S1, pooled EDTA plasma. S2, pooled serum spiked for most analytes. S3, pooled serum. Table S4a. Platform LC3

Table S8. CVs for matching and nonmatching A/IS on platform LC3[a]

	matching A/IS					nonmatching A/IS	
	within-run	between-run					
Analyte	All[b]	S1	S2	S3	All	All[c]	CV difference[d]
25-hydroxy vitamin D3	5.1 (3.2, 13.5)	6.3	5.1	5.4	6.1	14.4 (8.8, 25.3)	8.3 (2.7, 19.2)
all-trans-retinol	2.9 (1.6, 7.2)	3.8	3.0	2.6	3.6	14.8 (7.6, 26.6)	11.3 (4.1, 23.0)
25-hydroxy vitamin D2	6.6 (1.9, 22)	7.3	7.2	8.1	8.1	16.1 (9.9, 28.3))	8.0 (1.8, 20.2)
γ-tocopherol	3.7 (1.4, 10.8)	4.7	4.5	4.1	5.0	15.6 (9.5, 21.3)	10.6 (4.6, 16.3)
α-tocopherol	4.3 (2.0, 13.4)	5.3	5.8	4.6	5.8	18.0 (9.7, 18.4)	12.2 (3.9, 12.6)
phylloquinone	4.8 (2.1, 14.8)	5.9	6.8	6.6	7.4	24.1 (18.9, 26.7)	16.7 (11.5, 19.4)

[a]Numbers are CV (%) based on all data (All) or specific source as indicated. [b]Median (range) across 60 sets. [c]Median (range) across all nonmatching A/IS. [d]Median (range) difference nonmatching A/IS vs. matching A/IS across all nonmatching A/IS. Abbreviations A, analyte, IS, internal standard. S1, pooled EDTA plasma. S2, pooled serum spiked for most analytes. S3, pooled serum.