Journal of Mass Spectrometry and Advances in the Clinical Lab 33 (2024) 22–30

Contents lists available at ScienceDirect

Journal of Mass Spectrometry and Advances in the Clinical Lab

journal homepage: www.sciencedirect.com/journal/journal-of-mass-spectrometry-and-advances-in-the-clinical-lab

Research Article

Impact of internal standard selection on measurement results for long chain fatty acids in blood

John M. Goodwin VII *, Heather C. Kuiper, Barrett Brister, Hubert W. Vesper

Division of Laboratory Sciences, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA, USA

ARTICLE INFO

Keywords:
Internal standards
Isotope dilution
Gas chromatography
Mass spectrometry
Fatty acids

ABSTRACT

Introduction: Internal standards correct for measurement variation due to sample loss. Isotope labeled analytes are ideal internal standards for the measurement of fatty acids in human plasma but are not always readily available. For this reason, quantification of multiple analytes at once is most often done using only a single or few internal standards. The magnitude of the impact this has on method accuracy and precision is not well studied for gas chromatography-mass spectrometry systems.

Objective: This study aims to estimate bias and changes in uncertainty associated with using alternative fatty acid isotopologue internal standards for the estimation of similar or dissimilar long chain fatty acids.

Method: Using a previously reported method for the quantification of 27 fatty acids in human plasma using 18 internal standards we obtained estimates of bias and uncertainty at up to three levels of fatty acid concentration.

Results: With some notable exceptions, method accuracy remained relatively stable when using an alternative internal standard (Median Relative Absolute Percent Bias: 1.76%, Median Spike-Recovery Absolute Percent Bias: 8.82%), with larger changes in method precision (Median Increase in Variance: 141%). Additionally, the degree of difference between analyte and internal standard structure was related to the magnitude of bias and uncertainty of the measurement.

Conclusion: The data presented here show that the choice of internal standard used to estimate fatty acid concentration can affect the accuracy and reliability of measurement results and, therefore, needs to be assessed carefully when developing analytical methods for the measurement of fatty acid profiles.

Disclaimer: The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry. Use of trade names is for identification only and does not imply endorsement by the Centers for Disease Control and Prevention, the Public Health Service, and the US Department of Health and Human Services.

Introduction

Fatty acids (FA) are a broad class of bioactive molecule that play a wide variety of roles in the body and influence disease [1,2]. For these reasons, FA profiles have been used to assess aspects of human health [3-5]. Gas chromatography-mass spectrometry (GC–MS) is one of the most popular techniques for obtaining FA profiles due to its advantages in both selectivity and sensitivity [6]. However, GC–MS analyses of FAs in human samples requires significant sample preparation, which increases the risk of sample loss.

The use of an internal standard (IS) is a widely adopted technique to control for sample loss during FA analysis [7-10]. Briefly, this is accomplished by spiking the sample with a known concentration of IS material prior to sample preparation, and using the ratio of FA-to-IS instrument responses to estimate FA concentration. However, chemical species differ in their ability to control for sample dependent analyte loss.

ISs improve method precision only if the IS instrument response is correlated to that of the analyte to a degree that overcomes the increase in error associated with taking their quotient [11]. ISs improve method accuracy only if the proportions of analyte and IS concentration to instruments response remain constant in the calibration and unknown sample materials. For this reason, ISs perform better if they have chemical and physical properties that are similar to, but are measurable

* Corresponding author.
E-mail address: ocw6@cdc.gov (J.M. Goodwin).

https://doi.org/10.1016/j.jmsacl.2024.07.002
Received 18 December 2023; Received in revised form 11 July 2024; Accepted 24 July 2024
Available online 5 August 2024

apart from, the analyte of interest. Additionally, ISs with a natural abundance in the sample matrix confound the estimation of analyte concentration and perform more poorly as controls.

Common ISs for the quantification of FAs in human plasma include FA isotopologues and odd-chain FAs, and there are advantages and disadvantages to using either. Odd-chain FAs are easily obtained and are expected to have similar properties to FAs of similar length. However, there is evidence that odd-chain fatty acids are naturally abundant in human tissue and may not be suitable as controls in this sample matrix (Supplemental Fig. S1) [12]. FA isotopologues also have similar chemical and physical properties to their FA counterparts but can be labelled to the extent that there is no natural abundance in human tissue [13-15]. Additionally, LC-MS studies have found isotopologues to perform better than chemical analogue ISs [16,17]. However, preparing or purchasing isotopologues for each FA in a FA profile is often impractical. For this reason, many FA profiling methods choose to use a single IS to control for multiple FAs [7,8,18-20].

To the authors' knowledge, there is no data estimating the analytical cost of doing so in tandem with GC–MS. As machine learning and analytical techniques improve, the pace of research becomes more reliant on the reusability of method data for large cross-study analyses [21]. Method standardization and the reduction of method bias and imprecision is critical. Our study aims to inform IS selection in FA quantification by GC–MS by assessing changes in method accuracy, precision, and ruggedness associated with using a selection of 1 to 18 FA isotopologue ISs for the estimation of 27 FAs in human plasma. Furthermore, we seek to objectively find a relationship between method performance and FA/IS pair structures.

Materials and methods

Chemicals and Blood Samples

FA methyl esters (FAMEs) were purchased from Nu-Check Prep (Elysian, MN). Stable isotope-labeled FAs and FAME ISs (Table 1) were obtained from Sigma-Aldrich (St. Louis, MO), IsoSciences (King of Prussia, PA), Cayman Chemical (Ann Arbor, MI), and Cambridge Isotope Laboratories (Andover, MA). A GLC-674 Reference Standard Mixture was purchased from Nu-Check Prep (Elysian, MN) for spike-and-recovery accuracy assessments. All other reagents used were of analytical grade, and were purchased from Fisher Scientific (Pittsburgh, PA).

Quality control (QC) samples at low, medium, and high FA concentrations were prepared from plasma units obtained from Bioreclamation, Inc. (Westbury, NY). Individual plasma samples were also obtained from Bioreclamation, Inc. The company has IRB approval (IRB No. 20161665) to collect blood and obtains informed consent from donors. The CDC's use of the blood and urine is consistent with the IRB approval and donor consent. No personal identifiers were provided to the CDC, and the CDC's participation did not constitute human subject research.

Table 1
FA/IS pairs from Kuiper et al [10].

Analyte	Formula	Internal Standard
Myristic Acid	C14:0	D_{27}-C14:0
Myristoleic Acid	C14:1n-5	D_{27}-C14:0
Palmitic Acid	C16:0	$^{13}C_{16}$-C16:0
Palmitoleic Acid	C16:1n-7	$^{13}C_{16}$-C16:1n-7
Palmitelaidic Acid	C16:1n-7 t	$^{13}C_{5}$-C16:1n-7 t
Stearic Acid	C18:0	D_{35}-C18:0
Oleic Acid	C18:1n-9	$^{13}C_{18}$-C18:1n-9
Elaidic Acid	C18:1n-9 t	$^{13}C_{5}$-C18:1n-9 t
Vaccenic Acid	C18:1n-7	$^{13}C_{5}$-C18:1n-7
Trans-Vaccenic Acid	C18:1n-7 t	$^{13}C_{5}$-C18:1n-7 t
Linoleic Acid	C18:2n-6,9	$^{13}C_{18}$-C18:2n-6,9
Linoelaidic Acid	C18:2n-6 t,9t	$^{13}C_{5}$-C18:2n-6 t,9t
Alpha-Linolenic Acid	C18:3n-3,6,9	D_{14}-C18:3n-3,6,9
Gamma-Linolenic Acid	C18:3n-6,9,12	D_{14}-C18:3n-3,6,9
Arachidic Acid	C20:0	D_{39}-C20:0
11-Eicosenoic Acid	C20:1n-9	D_{39}-C20:0
11,14-Eicosadienoic Acid	C20:2n-6,9	D_{39}-C20:0
Homogamma-Linolenic Acid	C20:3n-6,9,12	D_{8}-C20:4n-6,9,12,15
Arachidonic Acid	C20:4n-6,9,12,15	D_{8}-C20:4n-6,9,12,15
Eicosapentaenoic Acid	C20:5n-3,6,9,12,15	D_{5}-C20:5n-3,6,9,12,15
Behenic Acid	C22:0	D_{43}-C22:0
Adrenic Acid	C22:4n-6,9,12,15	D_{5}-C22:6n-3,6,9,12,15,18
Docosapentaenoic Acid	C22:5n-3,6,9,12,15	D_{5}-C22:6n-3,6,9,12,15,18
Docosapentaenoic Acid	C22:5n-6,9,12,15,18	D_{5}-C22:6n-3,6,9,12,15,18
Docosahexaenoic Acid	C22:6n-3,6,9,12,15,18	D_{5}-C22:6n-3,6,9,12,15,18
Lignoceric Acid	C24:0	D_{47}-C24:0
Nervonic Acid	C24:1n-9	D_{47}-C24:0

Sample Preparation and GC–MS Analysis

Calibrators, QC samples, QC samples spiked with GLC-674, and individual plasma samples were processed following a previously published procedure [10,22]. In brief, samples were combined with ISs, hydrolyzed first with HCl, and then hydrolyzed again with NaOH. Total FAs were extracted from the hydrolysis solution with hexane. The isolated FAs were then derivatized with pentafluorobenzyl-bromide and analyzed on a 7890/5975C GC/MSD from Agilent Technologies in selected ion monitoring mode using negative chemical ionization with methane as the reagent gas. Separation was carried out on an Agilent Select FAME 200 m x 250 μm x 0.25 μm column using hydrogen as the carrier gas. This method (referred to from here on as "Kuiper et al.") determines the concentration of 27 FAs using 18 ISs (Table 1). Validation data and example chromatograms for this method have been published in the past [10].

Data analysis

Samples were measured in batches. Each batch measured three levels of QC material in duplicate, and five levels of calibrator material in singlicate. QC sample data and related calibrator data from 534 independent batches (1,068 measurements per QC level) collected over five years were combined into a "QC data set". Another 166 measurements from plasma samples from individual donors were combined into a "population data set". Finally, 12 batches of low and medium level QC samples were spiked with 0, 0.2, 2, 4, and 8 mg/mL of GLC-674 material, measured in triplicate per batch, and the data was combined into a "spike-recovery data set" for the estimation of true bias.

For each data set, instrument response ratios for all 486 possible combinations of FA and IS from Kuiper et al. were determined for each calibrator and sample. By batch, calibrator ratios associated with a specific FA/IS pair were used to estimate the FA concentrations of unknowns from the respective FA/IS ratios. The resulting concentration estimates were grouped by FA, IS, and sample. Basic statistics were calculated by group. Groups estimated in the QC and population data sets measuring the same FA from the same sample and using an FA/IS pair alternative to that described in Kuiper et al. were compared to those from Kuiper et al.

The percent bias relative to the group mean sample concentration, excluding outliers, was used as a measure of accuracy relative to the FA/IS pairs from Kuiper et al. Significant biases ($\alpha = 0.05$) relative to Kuiper et al. were identified using a series of paired T-tests. The Holm-Bonferroni method was used to compensate for familywise error.

The percent recovery of spike material was used as a measure of true accuracy for each FA/IS pair. The slope of the linear relationship between the known concentrations of spike material added to each sample and the measured sample concentration was used to estimate the percent recovery of spike material. Simple least squares regression was used to fit the data to a linear model. The confidence intervals for the slope estimate were estimated via bootstrapping. Confidence intervals

were adjusted for multiple comparisons using Bonferroni alpha correction.

Precision was ascertained by assessing the ratio of each group's variance over that of the associated group from Kuiper et al. Significant changes in measurement variance from that of Kuiper et al. were determined using a series of Levene's tests ($\alpha = 0.05$). Although the variables compared were not independent, the Levene's test was used here, regardless of decreased power, as no suitable alternative test for variance comparison between two dependent variables was available. The Holm-Bonferonni method was used to compensate for familywise error.

Ruggedness was measured as the proportion of extreme outliers to total data points by group. Extreme outliers were identified as those FA concentration estimates greater than three times the group interquartile range from the upper or lower quartiles. Chi-square tests were used to identify significant changes in method ruggedness ($\alpha = 0.05$) between each alternative FA/IS pairs and the associated pairs from Kuiper et al. The Holm-Bonferroni method was used to compensate for familywise error.

The accuracy and precision measurements from the population data set were obtained using the same methods used to measure the QC data set. Extreme outliers were not removed, and ruggedness was not assessed from the Population data set as FA data among individuals do not have symmetric distributions. To identify a potential structural basis for the IS-related changes in method performance, the difference in degrees of unsaturation and difference in acyl-chain length between FA and IS, as well as their associated absolute values, were assessed. This was done using a series of Mood's tests to understand the impact of simplified FA structural differences on accuracy, precision, and ruggedness measurements. The effects of nuisance variables were removed from the test using group-dependent medians in the Mood's test instead of the general median. A Pearson's correlation weighted by the total number of observations in each category was used to describe any general linear relationships between the proportion of data above the group-dependent median and the structural similarity variables. All data analysis was performed using R Version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results and discussion

Effects of IS Selection on Accuracy

Paired t-tests confirmed a statistically significant bias relative to Kuiper et al ($p < 0.05$) in 55.2 % of QC FA estimates using an alternative IS (Fig. 1 and Supplemental Table S1) and 66.7 % of population estimates (Fig. 2 and Supplemental Table S1). In the QC data, the median absolute value of relative percent biases was 1.76 %, and the values ranged from −39.6 % to 18.1 % among groups. In the population data, the median absolute relative percent bias was 1.74 % and ranged from −13.2 % to 8.67 %. In the spike-recovery data, percent recoveries were significantly different than 100 % ($p < 0.05$) in 60.9 % of FA/IS pairs, and percent recoveries ranged from 77.0 % to 238 % (Fig. 3). The median absolute percent difference from 100 % recovery was 8.82 %. These data suggest that IS selection from a pool of structurally similar candidates generally produces small to moderate bias; however, high bias is observed in certain cases. The risk of high bias among very similar ISs is especially notable, given the common practice of using only a single or a few ISs for the measurement of FA profiles.

FA/IS relative bias estimates from the QC and population data agree. Estimates from the population data set were moderately correlated to those estimated from the QC data set at each QC level (Pearson correlations: Low, 0.78; Medium, 0.69; High, 0.67) demonstrating general agreement between the two separate analyses with and without population level variability. Relative percent bias estimates derived from the spike-recovery data agree less strongly with the population and QC data sets (Pearson correlations: Low, 0.25; Medium, 0.25; High, 0.41; Population, 0.40). This is at least partly due to the smaller variability associated with the paired analysis of the QC and population accuracy data versus the unpaired regression analysis of the Spike-recovery data, and the difference in the sample number leveraged in the estimation of relative bias between the data sets (QC n = 1,068+, Population n = 166, Spike-Recovery n = 144).

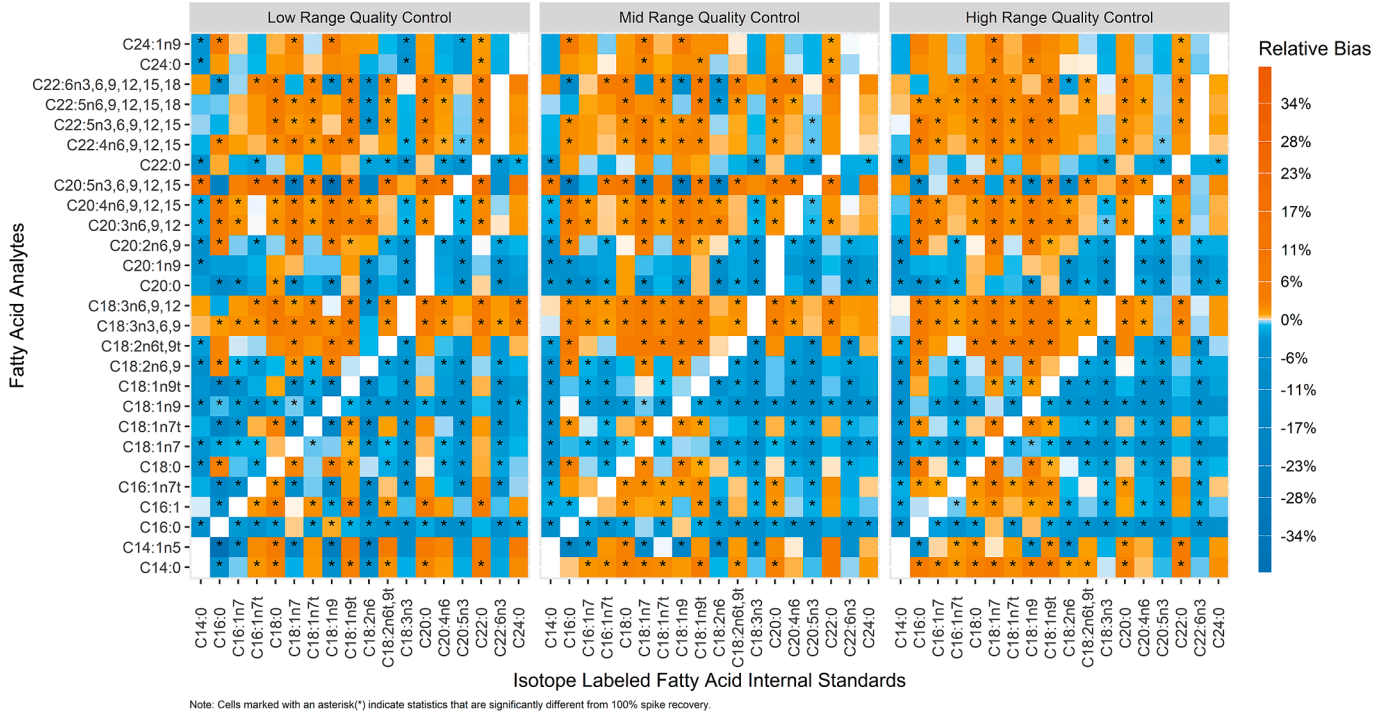

Fig. 1. Heat map showing the absolute bias for the low, medium, and high QC materials using all analyte-internal standard combinations for the quantitation of 27 FA with 18 isotopically labeled internal standards.

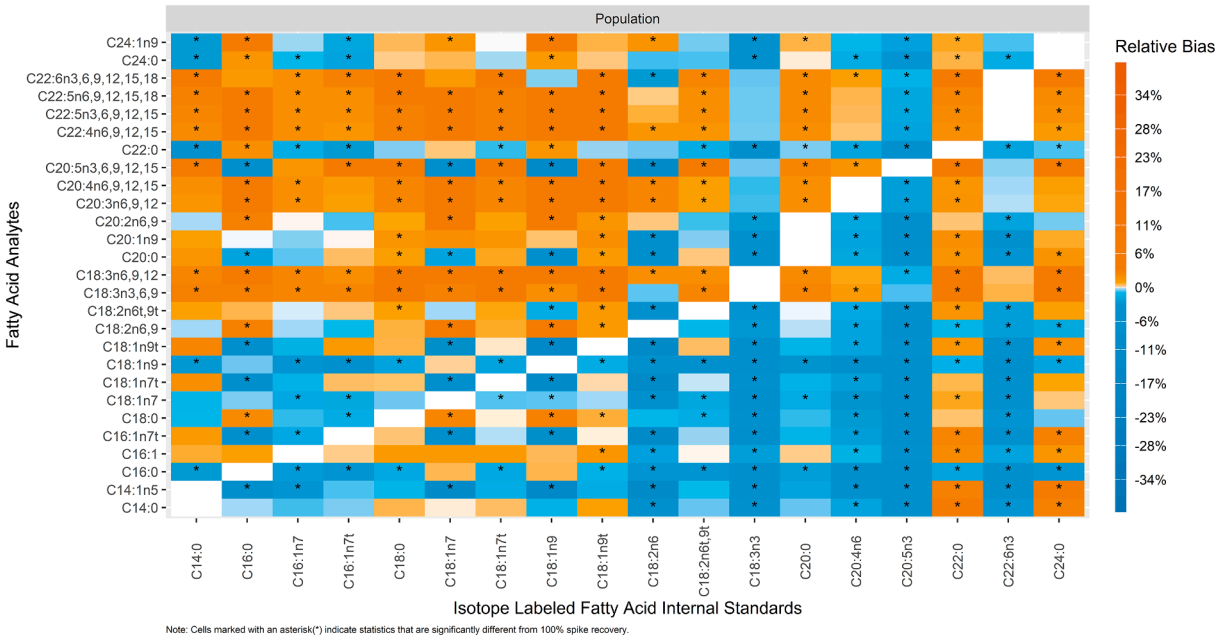

Fig. 2. Heat map showing the absolute bias estimates for the population samples using all analyte-internal standard combinations for the quantitation of 27 FA with 18 isotopically labeled internal standards.

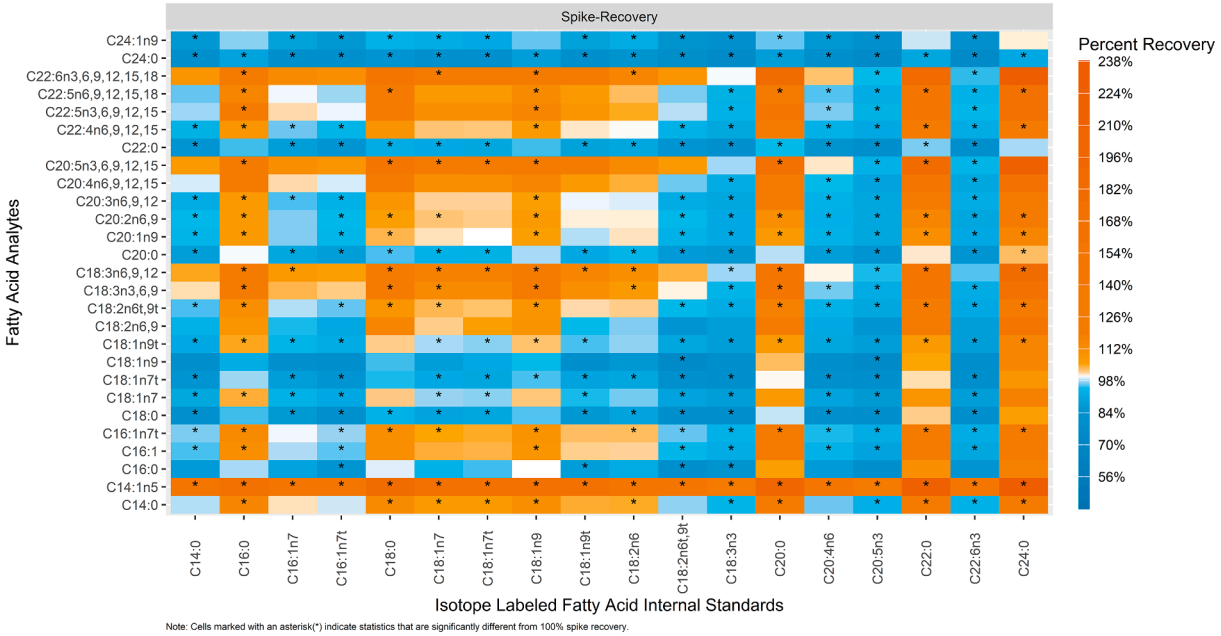

Fig. 3. Heat map showing the percent recovery estimates for the spike-recovery samples using all analyte-internal standard combinations for the quantitation of 27 FA with 18 isotopically labeled internal standards.

In 23 of the 27 FAs, there was at least one alternative FA/IS pair with percent recoveries closer to 100 % than those from Kuiper et al. However, in only 10 of the 23 FAs was the percent recovery statistically different than those from Kuiper et al ($p < 0.05$) (Table 2). For 6 of the 10 FAs that were statistically different than those from Kuiper et al, a corresponding FA isotopologue IS was among the pool of available 18 ISs from Kuiper et al, and these FA/IS pairs may be possible alternatives to those from Kuiper et al. Interestingly, for 4 of the 10 FAs, a corresponding isotopologue IS was available among the pool of 18 IS from Kuiper et al. All four are in the highly crowded C18 region of the chromatogram which suggests that co-eluting factors and/or matrix effects from the human plasma sample matrix are confounding accurate measurements. Perhaps these alternative ISs compensate for these effects better than the isotopologue IS. Or, it may be that some method aspect other than IS choice is producing bias, and the alternative IS counteracts this by producing a bias in the opposite direction. Limitations in the method design and the method's intended purpose prevent total resolution of all species in the C18:0 region; however, further spike-recovery experiments without the complex sample matrix may assess if co-eluting elements are the cause. As the intended purpose of the method is the measurement of FAs in human plasma, alternative FA/IS pairs identified in Table 2 represent attractive IS alternatives for improved method accuracy and interesting examples of alternative ISs outperforming analyte isotopologues.

Table 2
A list of alternative FA/IS pairs with spike recoveries closer to 100 % and significantly different ($p < 0.05$) from those FA/IS pairs recommended in Kuiper et al.

Analyte	Internal Standard	Percent Spike Recovery (95 % Confidence Interval)	Percent Recovery Closer to 100 % than Kuiper et al
C18:3n3,6,9	$^{13}C_{16}$-C16:1n-7	103.8 % (98.1 % – 111.2 %)	−1.5 %
C20:2n6,9	$^{13}C_{5}$-C18:1n-9 t	100.5 % (98.7 % – 103 %)	−11.1 %
C20:2n6,9	$^{13}C_{18}$-C18:2n-6,9	100.6 % (95.7 % – 104.1 %)	−11 %
C20:2n6,9	$^{13}C_{5}$-C18:1n-7 t	101.9 % (100.1 % – 103.7 %)	−9.7 %
C20:2n6,9	$^{13}C_{16}$-C16:1n-7	97.2 % (93.3 % – 100.9 %)	−8.8 %
C20:2n6,9	$^{13}C_{5}$-C16:1n-7 t	95 % (91.5 % – 97.7 %)	−6.6 %
C20:2n6,9	$^{13}C_{5}$-C18:2n-6 t,9t	94.9 % (91.7 % – 97.3 %)	−6.5 %
C20:2n6,9	D_{27}-C14:0	94.9 % (91.6 % – 98.5 %)	−6.5 %
C20:2n6,9	D_{39}-C20:0	93.1 % (87.1 % – 96.9 %)	−4.8 %
C20:2n6,9	D_{14}-C18:3n-3,6,9	91.8 % (84.5 % – 97 %)	−3.4 %
C20:2n6,9	D_{5}-C22:6n-3,6,9,12,15,18	91 % (83.6 % – 97.3 %)	−2.6 %
C20:2n6,9	D_{5}-C20:5n-3,6,9,12,15	90.5 % (82.8 % – 96.5 %)	−2.1 %
C20:1n9	$^{13}C_{5}$-C18:1n-7 t	100 % (96.2 % – 103 %)	−7.6 %
C20:1n9	$^{13}C_{5}$-C18:1n-7	101.1 % (94.5 % – 104 %)	−6.5 %
C20:1n9	$^{13}C_{5}$-C18:1n-9 t	98.4 % (94.5 % – 102 %)	−6%
C20:1n9	$^{13}C_{16}$-C16:1n-7	97.2 % (91.8 % – 101.1 %)	−4.8 %
C20:1n9	$^{13}C_{5}$-C16:1n-7 t	94.8 % (89.8 % – 98.7 %)	−2.5 %
C20:1n9	D_{27}-C14:0	94.3 % (87 % – 98.9 %)	−1.9 %
C20:1n9	D_{39}-C20:0	93.6 % (85.1 % – 98.7 %)	−1.3 %
C20:1n9	$^{13}C_{5}$-C18:2n-6 t,9t	93.5 % (87.8 % – 97 %)	−1.1 %
C20:3n6,9,12	$^{13}C_{18}$-C18:2n-6,9	99.3 % (97.5 % – 101.3 %)	−5.7 %
C20:3n6,9,12	$^{13}C_{5}$-C18:1n-7	101.7 % (96.6 % – 109.9 %)	−4.7 %
C20:3n6,9,12	$^{13}C_{18}$-C18:1n-9	106.2 % (100.1 % – 113 %)	−0.2 %
C22:4n6,9,12,15	$^{13}C_{18}$-C18:2n-6,9	100.1 % (98.5 % – 102.1 %)	−7.2 %
C22:4n6,9,12,15	$^{13}C_{5}$-C18:1n-7	102.6 % (98.1 % – 111.5 %)	−4.7 %
C22:4n6,9,12,15	$^{13}C_{18}$-C18:1n-9	107 % (101.9 % – 115.6 %)	−0.2 %
C18:2n6t,9t	$^{13}C_{5}$-C18:1n-9 t	101.3 % (97.8 % – 105.4 %)	−3.7 %
C18:2n6t,9t	$^{13}C_{16}$-C16:1n-7	98.5 % (96.5 % – 100.9 %)	−3.5 %
C18:2n6t,9t	$^{13}C_{18}$-C18:2n-6,9	101.8 % (99.3 % – 103.9 %)	−3.2 %
C18:2n6t,9t	$^{13}C_{5}$-C18:1n-7 t	102.8 % (99.8 % – 107.2 %)	−2.3 %
C18:2n6t,9t	$^{13}C_{5}$-C18:1n-7	103.6 % (100.8 % – 107.3 %)	−1.5 %
C18:1n9t	D_{35}-C18:0	102 % (98.2 % – 108.8 %)	−2.1 %
C18:1n9t	$^{13}C_{18}$-C18:1n-9	102.7 % (100.7 % – 104.8 %)	−1.4 %
C18:1n7t	$^{13}C_{16}$-C16:0	97.9 % (94.7 % – 100.6 %)	−6.9 %
C18:1n7t	$^{13}C_{18}$-C18:1n-9	95.9 % (94 % – 98.3 %)	−4.9 %
C18:1n7t	D_{47}-C24:0	108.6 % (94.5 % – 137.1 %)	−0.3 %
C22:5n6,9,12,15,18	$^{13}C_{18}$-C18:2n-6,9	103.3 % (99.6 % – 108.3 %)	−2.6 %
C22:5n3,6,9,12,15	$^{13}C_{18}$-C18:2n-6,9	104.8 % (99.9 % – 112.7 %)	−0.5 %

While IS studies for LC-MS and GC–MS methods are not rare in the literature, very few attempt to find statistical evidence relating IS structure to efficacy [23–25]. The number, and high structural similarity, of FA and IS in this study allows for the use of non-parametric statistics to mathematically demonstrate a relationship between FA/IS structure and method performance. As the ISs are all isotopologues of the FA analytes, the only changing structural variables are the mass, acyl-chain length, and the degrees acyl-chain saturation. As it is assumed that an external calibration curve corrects for isotope effects on analyte and IS response factors, we assume here that mass has little effect on method performance and focus instead on acyl-chain length and degree of saturation.

Mood's analysis indicates FA/IS pairs with different acyl-chain lengths produced biases that were significantly different in the QC and Population data sets, but no significant change in bias related to acyl-chain length was found from the Spike-Recovery data (QC data set p: 0.004, Population data set p: 0.001, Spike-Recovery data set p: > 0.05). In addition, no significant relationship between the difference in FA and IS acyl-chain length and the associated bias was observed from either the Population or QC data sets. Bias from all data sets changed as the difference between FA and IS unsaturation changed ($p < 0.001$). However, FA/IS pairs with greater differences in unsaturation tended to have larger bias (Population: 0.96, QC: 0.95, Spike-Recovery: 0.98). Additionally, FA concentrations estimated with a more unsaturated IS tended to have a more negative bias, while those estimated using a more saturated IS tended to have a more positive bias (Population: 0.81, QC: 0.83, Spike-Recovery: 0.95). As Kuiper et al. uses both external standards and ISs for the quantification of FAs in plasma, it is unlikely that any difference in behavior between the FAs and ISs alone is the cause. As the calibration curve does not include the entire sample matrix, it is likely the bias is the outcome of complex sample matrix interactions not captured in the external standards. More investigation is necessary to ascertain the cause of the relationship between IS saturation and bias, but these findings are consistent with the general understanding that ISs more structurally like the analyte perform better.

Effects of IS Selection on Precision

Generally, using an alternative IS reduced method precision. In the QC data set, using an alternative IS increased method variance by 141 % on median (Fig. 4) and ranged from a decrease of 15.4 % to an increase of 577 %. Variance increased in 97.7 % of observations and was statistically significant ($p < 0.05$) in 85.7 % of observations. In the population data set, variance of the log-transformed alternative IS data increased by 3 % on median and ranged from a decrease of 13.6 % to an increase of 70 % (Fig. 5). While generally in the population data set, alternative IS increased variance, no change was statistically significant due to the large inter-individual variance component and the lack of replicate measurements for individual samples.

A total of seven alternative FA/IS pairs produced measurements with variance smaller than that from Kuiper et al (Table 3), though none of these improvements to precision were found to be statistically significant. Of these, only C18:1n9t/$^{13}C_{5}$-C18:1n-7 t, C20:1n9/D_{35}-C18:0, and C24:1n9/D_{43}-C22:0 produced smaller variances at all three QC levels. Notably, all three have minor differences in saturation and acyl-chain length between FA and IS. The largest increase in variance occurs when measuring the shortest saturated FA in the study (C14:0) using the longest most unsaturated IS (D_{5}-C22:6n-3,6,9,12,15,18) (Fig. 6)

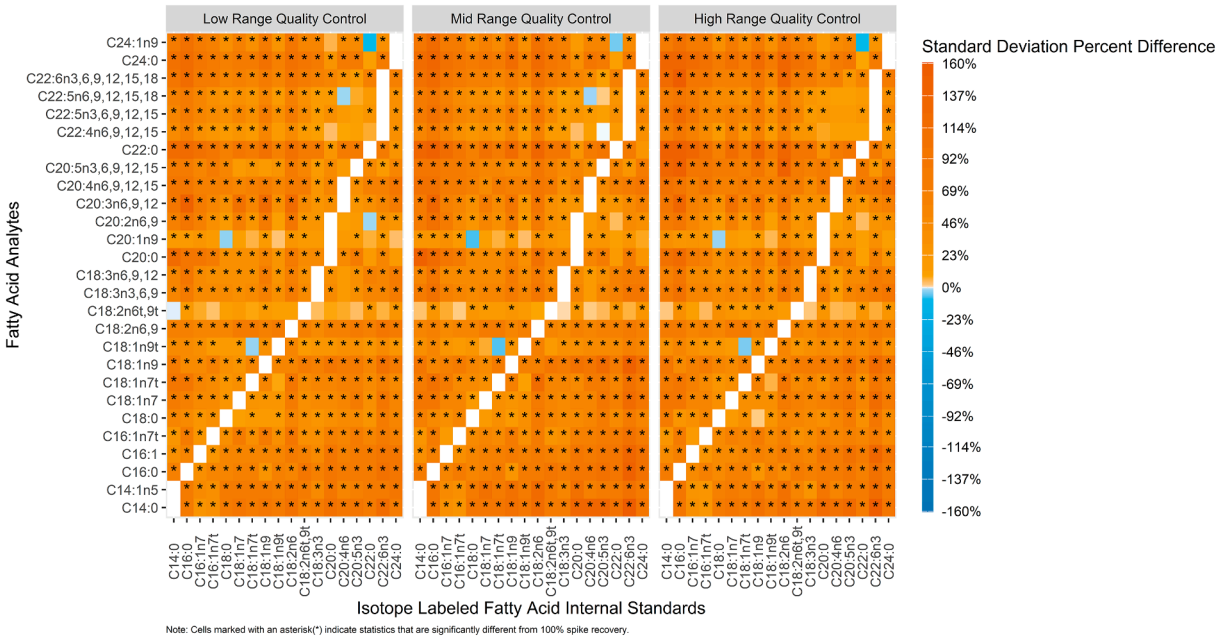

Fig. 4. Heat map showing the variance for the low, medium, and high QC materials using all analyte-internal standard combinations for the quantitation of 27 FA with 18 isotopically labeled internal standards.

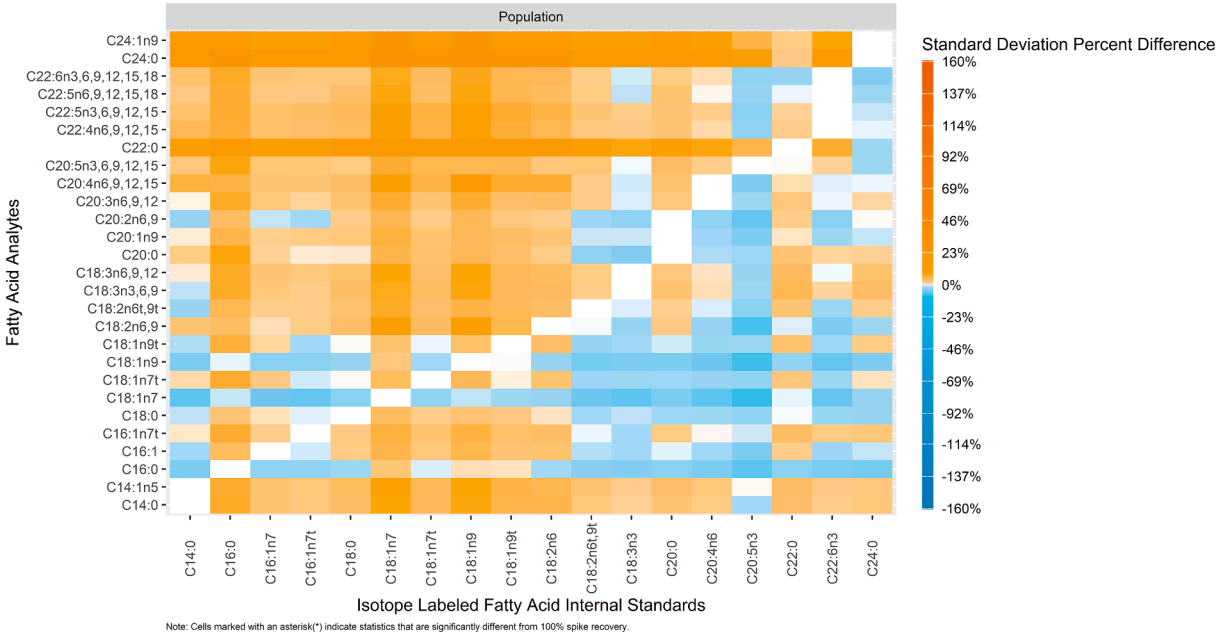

Fig. 5. Heat map showing the variance estimates for the population samples using all analyte-internal standard combinations for the quantitation of 27 FA with 18 isotopically labeled ISs.

suggesting a possible relationship between FA/IS pair structural similarity and method precision. Indeed, Mood's analysis found that sample variance in the QC data set changed significantly with differences in the saturation between FA and IS ($p < 0.001$). Additionally, there is a positive and statistically significant relationship between the absolute difference in FA and IS unsaturation and the sample variance (Weighted Pearson Correlation: 0.86, $p = 0.012$) suggesting that FA/IS pairs with similar number of double bonds tend to have more precise results. Additionally, absolute differences in FA and IS acyl-chain length also coincide with a significant change in sample variance ($p < 0.001$). However, no significant correlation was observed between similarity in FA/IS pair acyl-chain lengths and method variance suggesting differences between FA and IS acyl-chain length are less informative than differences between FA and IS unsaturation. These findings are similar to those from the IS dependent accuracy analysis and suggest that agreement in FA/IS pair saturation is an important IS selection criteria for both method accuracy and precision.

Of the alternative FA/IS pairs that improved accuracy, C20:2n6,9/D39-C20:0, C20:1n9/13C5-C18:1n7, C20:1n9/13C5-C18:1n7t, C20:1n9/13C5-C18:1n9t, C18:2n6t,9t/13C16-C16:1n7, and C18:2n6t,9t/13C18-C18:2n-6,9 showed no statistically significant reduction in precision while also showing a statistically significant improvement in accuracy (Table 2). This provides additional support for these FA/IS pairs as more accurate alternatives to those from Kuiper et al

Table 3
A list of alternative FA/IS pairs with less variance than those from Kuiper et al.

Analyte	Internal Standard	Percent Variance Change from Kuiper et al		
		Low QC	Mid QC	High QC
C22:5n6,9,12,15,18	D_{39}-C20:0	−2.6 %	−2.9 %	20.2 %
C22:4n6,9,12,15	D_5-C20:5n-3,6,9,12,15	17.5 %	−0.2 %	23 %
C20:2n6,9	D_{43}-C22:0	−2.3 %	7.9 %	7.1 %
C20:1n9	D_{35}-C18:0	−4.3 %	−12 %	−3.7 %
C24:1n9	D_{43}-C22:0	−15.4 %	−7.8 %	−15 %
C18:1n12	$^{13}C_5$-C18:1n-9 t	1.7 %	0.8 %	−0.2 %
C18:1n9t	$^{13}C_5$-C18:1n-7 t	−3.9 %	−10.9 %	−7.2 %
C18:2n6t,9t	D_{27}-C14:0	−0.7 %	5.4 %	3.5 %

contingent on ruggedness results.

Case studies from the literature demonstrate that an IS's improvement to accuracy or precision do not necessarily coincide [26], as is evidenced from the other 32 FA/IS pairs that significantly improved accuracy but significantly reduced precision. Within the scope of this method, accuracy as it pertains to IS selection is dependent on the relative magnitude of systematic errors in the FA and IS responses from calibrator and sample, while precision depends on how well the IS mirrors the response of the analyte to sources of random error such as fluctuations in the environment or the instruments used during analysis. Therefore, as mentioned in the accuracy section, systematic bias already present in the method can be counteracted by a less ideal IS that produces an opposing bias resulting in a more accurate but less precise method as the less ideal IS's response does not correlate as well with that of the analyte.

Effects of IS Selection on Ruggedness

In the QC data, on median, 5.83 % of FA/IS pair results were greater than 3 interquartile ranges above the third quartile or less than 3 interquartile ranges below the first quartile (Fig. 7). Depending on IS choice, extreme outliers comprised from 0 % to 12.98 % of FA/IS pair data. Generally, using an alternative IS increased the number of extreme outliers observed in the data. When using an alternative IS, the proportion of outlying data points increased by 50 % on median. Changes to the proportion of outlying data points ranged from a decrease of 100 % to an increase of 1,294 %. Using an alternative IS significantly increased the number of outliers in 49.7 % of FA/IS pairs, but significantly decreased the number of outliers in 17.5 %.

Multiple alternative FA/IS pairs significantly improved ruggedness over Kuiper et al (Table 4). FAs for which an alternative FA/IS pair showed improvement include C14:1n5c, C16:1n7t, C20:2n6c,9c, C20:3n6,9,12, C20:1n9c, C22:4n6,9,12,15, C22:5n3,6,9,12,15, and C22:5n6,9,12,15,18. Of these C22:4n6,9,12,15/D_{27}-C14:0, C22:4n6,9,12,15/$^{13}C_{18}$-C18:2n-6,9, C22:4n6,9,12,15/D_{47}-C24:0, C20:2n6,9/$^{13}C_5$-C18:1n-7, C20:2n6,9/$^{13}C_{18}$-C18:1n-9, C20:3n6,9,12/$^{13}C_{18}$-C18:2n-6,9, C22:5n3,6,9,12,15/D_{14}-C18:3n-3,6,9, and C22:5n6,9,12,15,18/D_{14}-C18:3n-3,6,9 significantly improved ruggedness across all three levels of QC. The associated isotopologues of these FAs are not included in the study IS pool which suggests these alternative FA/IS pairs may be an improvement over those from Kuiper et al.

All six FA/IS pairs that improved accuracy with no significant change in precision also significantly improved, or at least did not significantly reduce, method ruggedness (Table 5). The three FA/IS pairs measuring C20:1n9 significantly improved ruggedness in mid and high levels QC materials and improved ruggedness from the low QC measurements, though the improvement was not statistically significant ($p > 0.05$). The other three FA/IS pairs did not significantly change method ruggedness from Kuiper et al. All six represent good alternatives that will likely improve method accuracy with only small to no changes in variance and ruggedness, apart from those FA/IS for C20:1n9, which dramatically improve method ruggedness over the current FA/IS from Kuiper et al.

QC data also suggests that structural dissimilarities between FA and IS modulate method ruggedness. The proportion of outliers changed significantly as the difference between analyte and IS unsaturation changed ($p \ll 0.001$). There is a strong positive correlation between the absolute difference between FA and IS unsaturation and the proportion of extreme outliers (Pearson correlation: 0.968). This suggests analyte and IS pairs more dissimilar in terms of unsaturation tend to produce larger numbers of outlying measurements. The proportion of outliers also changed significantly with regards to absolute difference in acyl-chain length ($p = 0.039$). However, no significant linear relationship

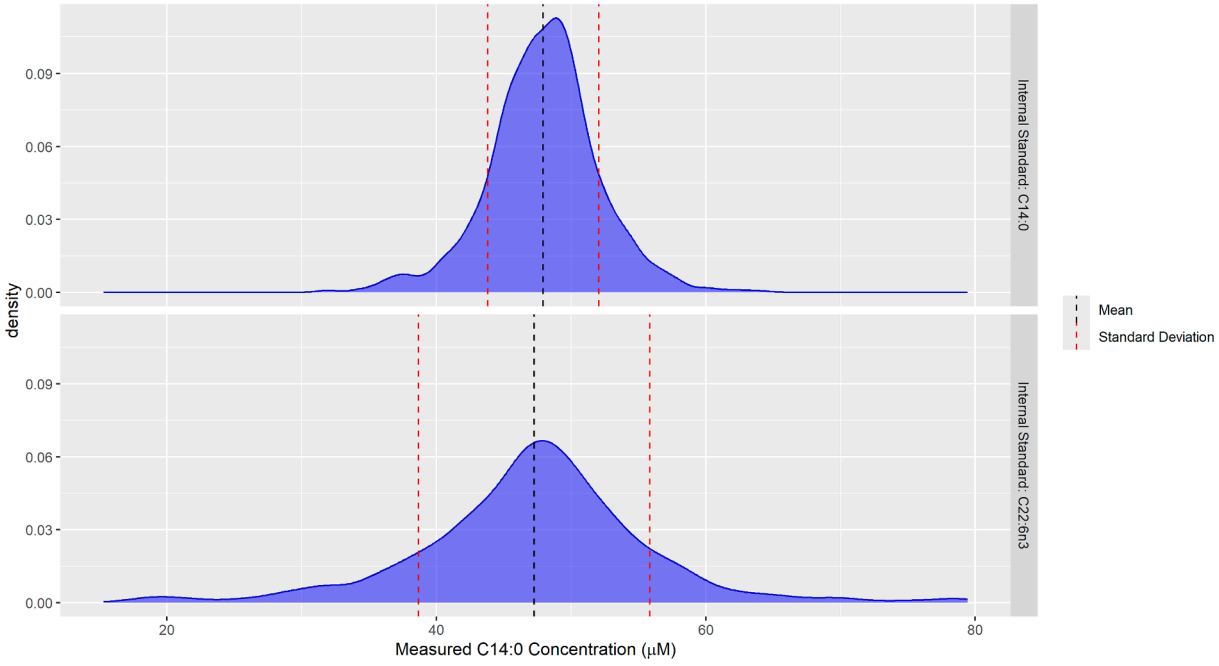

Fig. 6. Density plots demonstrating the change in variance when estimating sample concentration of C14:0 in low QC material using internals standards D_{27}-C14:0 or D_5-C22:6n3.

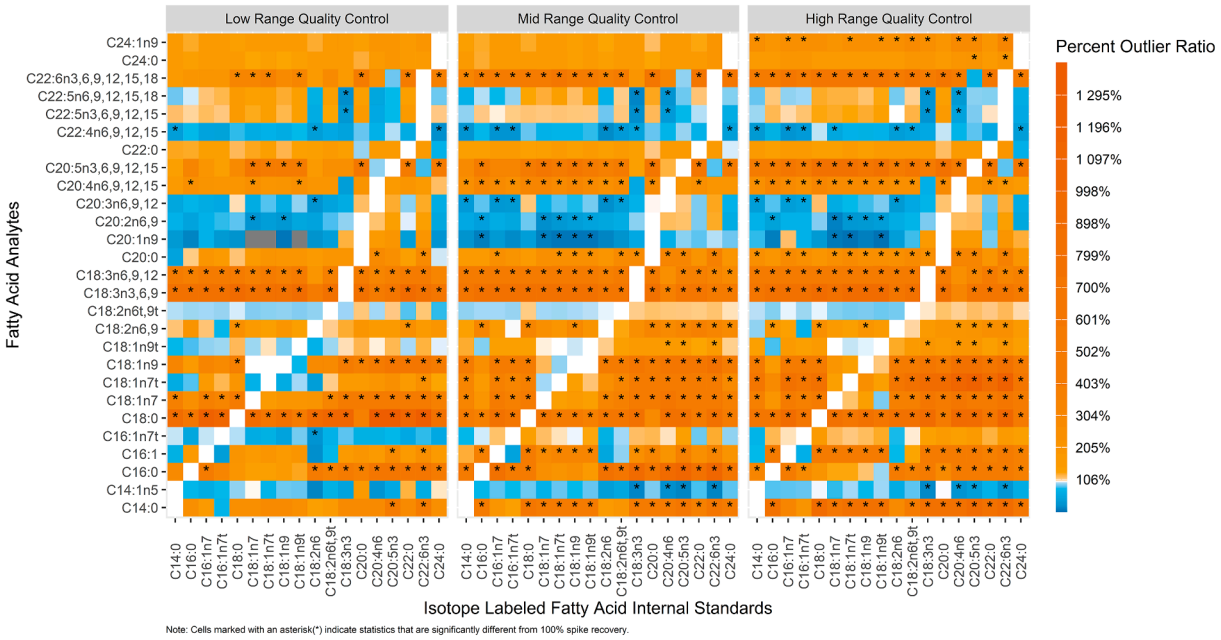

Fig. 7. Heat map showing the ruggedness for the low, medium, and high QC materials using all FA/IS combinations for the quantitation of 27 FA with 18 isotopically labeled ISs.

Table 4
A list of FA/IS pairs that reduced the number of extreme measurements across all three Quality Control levels.

Analyte	Internal Standard	Percent Outlier Reduction		
		QC Low	QC Mid	QC High
C22:5n3,6,9,12,15	D14-C18:3n3,6,9	76.0 %	67.0 %	65.0 %
C22:5n6,9,12,15,18	D14-C18:3n3,6,9	80.0 %	76.0 %	71.0 %
C22:4n6,9,12,15	13C18-C18:2n-6,9	55.0 %	78.0 %	58.0 %
C22:4n6,9,12,15	D47-C24:0	76.0 %	54.0 %	54.0 %
C22:4n6,9,12,15	D27-C14:0	57.0 %	67.0 %	62.0 %
C20:2n6c,9c	13C5-C18:1n-7	68.0 %	75.0 %	84.0 %
C20:2n6c,9c	13C18-C18:1n-9	65.0 %	58.0 %	62.0 %
C20:3n6,9,12	13C18-C18:2n-6,9	70.0 %	77.0 %	62.0 %

Table 5
A list of alternative FA/IS pairs that improve accuracy, do not significantly reduce precision, and do not significantly reduce ruggedness. Values marked with an asterisk (*) represent values significantly different than those FA/IS pairs recommended in Kuiper et al.

Analyte	Internal Standard	Percent Outlier Change from Kuiper et al.		
		Low QC	Mid QC	High QC
C20:2n6,9	D_{39}-C20:0	10.4 %	14.4 %	14.4 %
C20:1n9	$^{13}C_5$-C18:1n-7	−100 %	*-85.2 %	*-88.5 %
C20:1n9	$^{13}C_5$-C18:1n-7 t	−100 %	*-88.9 %	*-92.3 %
C20:1n9	$^{13}C_5$-C18:1n-9 t	−100 %	*-88.9 %	*-92.3 %
C18:2n6t,9t	$^{13}C_{16}$-C16:1n-7	−4.3 %	−4.2 %	−1.3 %
C18:2n6t,9t	$^{13}C_{18}$-C18:2n-6,9	−7.4 %	−1.1 %	−11.2 %

was observed between the difference in FA and IS acyl-chain lengths and percent outliers. As in the accuracy and precision data, the ruggedness analysis suggests that agreement in FA and IS unsaturation is important for good method performance.

It is interesting that the agreement between FA and IS saturation changed significantly with method accuracy, precision, and ruggedness. As sources of systematic error are mitigated by the external calibration curve and the calibrators do not contain the sample matrix, the change in method performance may be due to sample matrix effects. Plasma containing various storage proteins and surfactants is most likely to affect steps prior to and including extraction. As unsaturation has a large effect on FA packing mechanics within and between two immiscible medias, it follows that human plasma may affect extraction efficiencies of both IS and FA analytes and may also explain the large impact of FA and IS differences in unsaturation on method performance. In the literature, the most discussed matrix effects in GC–MS methods are due to the creation or blocking of active sites which may retain or remove analyte or IS [27]. However, this seems less likely in this case due to the significant sample clean-up prior to injection. Less discussed is ion suppression or enhancement as it is believed that ionization in GC–MS is more complete [28]. However, previously discussed observations of reduced method accuracy in the crowded C18 region of the chromatogram are consistent with matrix effects occurring during detection.

Conclusions

We have demonstrated how IS choice can produce significant changes in method bias, variability, and ruggedness for FA analysis by GC–MS. This is especially noteworthy since many FA quantitation methods commonly use only one IS per FA profile. As a result, data generated with different FA/IS pairs may not be comparable across different methods or between laboratories.

CRediT authorship contribution statement

John M. Goodwin: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Heather C. Kuiper:** Resources, Methodology, Investigation, Conceptualization. **Barrett Brister:** Validation, Software, Formal analysis, Writing – review & editing. **Hubert W. Vesper:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Conceptualization, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Emily Mueller, Samantha McGunigale, Alex Doty, Melissa Missinne, Chui Tse, Hayoung Kim, Victoria Kennereley and Na Wei for their support in laboratory measurements as well as Samuel Caudill for his support with statistical assessment.

Disclaimer.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry. Use of trade names is for identification only and does not imply endorsement by the Centers for Disease Control and Prevention, the Public Health Service and the US Department of Health and Human Services.

Use of trade names and commercial sources is for identification only and does not constitute endorsement by the U.S. Department of Health and Human Services, or the U.S. Centers for Disease Control and Prevention (Division of Laboratory Sciences).

Funding Support

This work was funded by the Centers for Disease Control and Prevention, United States.

Ethics Statement.

Quality control (QC) samples were prepared from plasma units obtained from Bioreclamation, Inc. (Westbury, NY). Individual plasma samples were also obtained from Bioreclamation, Inc. The company has IRB approval (IRB No. 20161665) to collect blood and obtains informed consent from donors. The CDC's use of the blood and urine is consistent with the IRB approval and donor consent. No personal identifiers were provided to the CDC, and the CDC's participation did not constitute human subject research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmsacl.2024.07.002.

References

[1] P.C. Calder, Functional Roles of Fatty Acids and Their Effects on Human Health, JPEN. 39 (2015) 18S–32S.
[2] K.L. Fritsche, The science of fatty acids and inflammation, Adv. Nutr. 15 (2015) 293S–301S.
[3] K.H. Jackson, W.S. Harris, Blood Fatty Acid Profiles: New Biomarkers for Cardiometabolic Disease Risk, Curr. Atheroscler. Rep. 20 (2018).
[4] E.M. Song, J.S. Byeon, S.M. Lee, H.J. Yoo, S.J. Kim, S.H. Lee, K. Chang, S. W. Hwang, D.H. Yang, J.Y. Jeong, Fecal Fatty Acid Profiling as a Potential New Screening Biomarker in Patients with Colorectal Cancer, Dig. Dis. Sci. 63 (2018) 1229–1236.
[5] D.C. Wang, C.H. Sun, L.Y. Liu, X.H. Sun, X.W. Jin, W.L. Song, X.Q. Liu, X.L. Wan, Serum fatty acid profiles using GC-MS and multivariate statistical analysis: potential biomarkers of Alzheimer's disease, Neurobiol. Aging. 33 (2012) 1057–1066.
[6] I. Tiuca, K. Nagy, R. Oprean, Recent developments in fatty acids profile determination in biological samples - a review, Rev. Romana. Med. Lab. 23 (2015) 371–384.
[7] S.A. Abdelmagid, S.E. Clarke, D.E. Nielsen, A. Badawi, A. El-Sohemy, D.M. Mutch, D.W.L. Ma, Comprehensive Profiling of Plasma Fatty Acid Concentrations in Young Healthy Canadian Adults, PLoS ONE. 10 (2015) e0116195. https://doi:10.1371/journal.pone.0116195.
[8] S.A. Abdelmagid, D.E. Nielsen, A. Badawi, A. El-Sohemy, D.M. Mutch, W.L. David, Circulating concentrations and relative percent composition of trans fatty acids in healthy Canadian young adults between 2004 and 2010: a cross-sectional study, CMAJ OPEN. 5 (2017). https://doi:10.9778/cmajo.20160048.
[9] E. Kish-Trier, E.L. Schwarz, M. Pasquali, T. Yuzyuk, Quantitation of total fatty acids in plasma and serum by GC-NCI-MS, Clin. Mass Spectrom. 2 (2016) 11–17.
[10] H.C. Kuiper, N. Wei, S.L. McGunigale, H.W. Vesper, Quantitation of trans-fatty acids in human blood via isotope dilution-gas chromatography-negative ionization-mass spectrometry, J. Chromatogr. B 1076 (2018) 35–43.
[11] P. Haefelfinger, Limits of the internal standard technique in chromatography, J. Chromatogr. a. 218 (1981) 73–81.
[12] M. Pfeuffer, A. Jaudszus, Pentadecanoic and Heptadecanoic Acids: Multifaceted Odd-Chain Fatty Acids, Adv. Nutr. 7 (2016) 730–734.
[13] M. Claeys, S.P. Markey, W. Maenhaut, Variance analysis of error in selected ion monitoring assays using various internal standards, Biomed. Mass Spectrom. 4 (1977) 122–128.
[14] A.P. De Leenheer, M.F. Lefevere, W.E. Lambert, E.S. Colinet, Isotope-dilution mass spectrometry in clinical chemistry, Adv. Clin. Chem. 24 (1985) 111–161.
[15] M.G. Giovannini, G. Pieraccini, G. Moneti, Isotope dilution mass spectrometry: definitive methods and reference materials in clinical chemistry, Ann. Ist. Super Sanita. 27 (1991) 401–410.
[16] A. Tan, S. Hussain, A. Musuku, R. Masse, Internal standard response variations during incurred sample analysis by LC-MS/MS: Case by case trouble-shooting, J. Chromatogr. b. 877 (2009) 3201–3209.
[17] M.M. Khamis, D.J. Adamko, A. El-Aneed, Strategies and Challenges in Method Development and Validation for the Absolute Quantification of Endogenous Biomarker Metabolites Using Liquid Chromatography-Tandem Mass Spectrometry, Mass Spectrom Rev (2019).
[18] A.L. Mihai, M. Negoita, N. Belc, Evaluation of fatty acid profile of oils/fats y GC-MS through two quantification approaches, Rom. Biotechnol. Lett. 24 (2019) 973–985.
[19] W.M.N. Ratnayake, M.R. L'Abbe, S. Farnworth, L. Dumais, C. Gagnon, B. Lampi, V. Casey, D. Mohottalage, I. Rondeau, L. Underhill, M. Vigneault, W. Lillycrop, M. Meleta, L.Y. Wong, T. Ng, Y. Gao, K. Kwong, S. Chalouh, P. Pantazopoulos, H. Gunaratna, A. Rahardja, R. Blagden, V. Roscoe, T. Krakalovich, G. Neumann, G. A. Lombaert, Trans fatty acids: Current contents in Canadian foods and estimated intake levels for the Canadian population, J AOAC. 92 (2009) 1258–1276.
[20] J.M. Fernandez-Real, J. Vendrell, W. Ricart, Circulating Adiponectin and Plasma Fatty Acid Profile, Clin. Chem. 51 (2005) 603–609.
[21] L. Kumuthini, M. Zass, Z. Chaouch, V. Gill, Z. Ras, D. Mungloo-Dilmohamud, A. Sathan, F. Ghoorah, C. Fadlelmola, J.V. Fields, F. Horn, M. Radouani, E. R. Konopko, S.B. Chimusa, Data standardization in the omics field, Genomic Data Sharing. (2023) 137–155, https://doi.org/10.1016/B978-0-12-819803-2.00008-0.
[22] H.W. Vesper, S.P. Caudill, H.C. Kuiper, Q. Yang, N. Ahluwalia, D.A. Lacher, J. L. Pirkle, Plasma trans-fatty acid concentrations in fasting adults declined from NHANES 1999–2000 to 2009–2010, AJCN. 105 (2017) 1063–1069.
[23] M.H. Hiatt, Internal standards: A source of analytical bias for volatile organic analyte determinations, J. Chromatogr. a. 1218 (2011) 498–503.
[24] K.A. Smith, S.D. Merrigan, K.L. Johnson-Davis, Selecting a Structural Analog as an Internal Standard for the Quantification of 6-Methylmercaptopurine by LC-MS/MS, JALM. (2018) 384–396.
[25] M. Wang, C. Wang, X. Han, Selection of internal standards for accurate quantification of complex lipid species in biological extracts by electrospray ionization mass spectrometry – What, how, and why, Mass Spectrom Rev. 36 (2017) 693–714.
[26] A. Tan, S. Hussain, A. Musuku, R. Massé, Internal standard response variations during incurred sample analysis by LC–MS/MS: Case by case trouble-shooting, J. Chromatogr. b. 877 (2009) 3201–3209.
[27] K. Mastovska, S.J. Lehotay, M. Anastassiades, Combination of Analyte Protectants to Overcome Matrix Effects in Routine GC Analysis of Pesticide Residues in Food Matrixes, Anal. Chem. 77 (2005) 8129–8137.
[28] T.M. Annesley, Ion Suppression in Mass Spectrometry, Clin. Chem. 49 (2003) 1041–1044.