



A normalized signal calibration with a long-term reference improves the robustness of RPLC-MRM/MS lipidomics in plasma

Bohdan Babiy¹ · Rebeca Busto^{3,2} · Óscar Pastor^{1,3}

Received: 11 March 2021 / Accepted: 19 April 2021 / Published online: 28 April 2021
© Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Improving the reliability of quantification in lipidomic analyses is crucial for its successful application in the discovery of new biomarkers or in clinical practice. In this study, we propose a workflow to improve the accuracy and precision of lipidomic results issued by the laboratory. Lipid species from 11 classes were analyzed by a targeted RPLC-MRM/MS method. The peak areas of species were used to estimate concentrations by an internal standard calibration approach (IS-calibration) and by an alternative normalization signal calibration schema (NS-calibration). The latter uses a long-term reference plasma material as a matrix-matched external calibrator whose accuracy was compared to the NIST SRM-1950 mean consensus values reported by the Interlaboratory Lipidomics Comparison Exercise. The bias of lipid concentrations showed a good accuracy for 69 of 89 quantified lipids. The quantitation of species by the NS-calibration schema improved the within- and between-batch reproducibility in quality control samples, in comparison to the usual IS-calibration approach. Moreover, the NS-calibration workflow improved the robustness of the lipidomics measurements reducing the between-batch variability (relative standard deviation <10% for 95% of lipid species) in real conditions tested throughout the analysis of 120 plasma samples. In addition, we provide a free access web tool to obtain the concentration of lipid species by the two previously mentioned quantitative approaches, providing an easy follow-up of quality control tasks related to lipidomics.

Keywords Lipidomics · Mass spectrometry · Quality assurance · Molar concentration · Software · NIST SRM-1950

Introduction

Advances in mass spectrometry (MS) applied to the field of lipidomics allow the identification of hundreds of lipid molecules in plasma [1], but obtaining their concentration precisely and accurately in analytical terms is still a challenge. Improving the reliability of lipidomics quantification is crucial, whether it is focused on the discovery of new biomarkers, or its application

in actual clinical practice [2]. This is because the relevant information of any lipid analysis employed every day in clinical practice, take for example the analyses of total cholesterol and triglycerides, is related to the monitorization of their absolute concentration in response to the disease or a therapeutic intervention [3]. This poses a real challenge as there is no universally accepted *lipidomic approach*, and many methodological steps, data pipelines, and technical choices are required in the analysis of lipid species by MS [4, 5]. In general, the objectives of identifying a large number of lipids collide with those of obtaining a precise and exact quantification of the species. In addition, there are many additional choices taken by laboratories before translating lipid signals produced by MS into lipid concentrations [6].

The accuracy of an analytical procedure expresses the closeness of the determined value to the value which is accepted either as a conventional true value or an accepted reference value [7, 8]. The accuracy of the quantification of lipid species in MS lipidomics depends essentially on three factors: (1) the consideration of different instrumental responses of the lipid species, (2) accounting for the matrix effects acting on

✉ Óscar Pastor
oscar.pastor@salud.madrid.org

¹ Servicio de Bioquímica Clínica, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid, Spain

² Servicio de Bioquímica-Investigación, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid, Spain

³ CIBER de Fisiología de la Obesidad y Nutrición (CIBEROBN), ISCIII, Madrid, Spain

the lipid species in individual samples, and (3) choice of the type of calibration [9]. The instrumental response for every lipid species in MS is determined by its lipid class and other structural elements, such as bond types (ester vs. ether), number of double bonds, and length of the acyl chain [10, 11]. Matrix effects are caused by the alteration of ionization efficiency of target analytes in the presence of co-eluting compounds in the same matrix. They can be observed either as a loss in response (ion suppression) or as an increase in response (ion enhancement). The use of stable isotopically labelled compounds as internal standards (IS) and calibration with matrix-matched standards are the best choices to cope with matrix effects, as they are widely employed in liquid chromatography (LC)-MS/MS clinical chemistry assays [12]. With regard to the type of calibration, the use of an external multi-level calibration within the working range of the target compound, given by the lower limit of quantification (LLOQ) and the upper limit of quantification (ULOQ), is the best choice. Single-point calibrations are accepted, but additional experiments to prove the linear dynamic range of the method are necessary [6].

From the above, it is clear that the highest metrological order for reporting absolute concentrations is isotope dilution MS, established by matrix-matched multilevel external calibration with internal standardization. Unfortunately, this is only possible when both reference standards for each target lipid and a matching IS in co-ionization with the species are available. Therefore, although providing absolute concentrations is an achievable analytical goal when there are standards (synthetically equivalent and deuterated), their use for each lipid is not feasible due to their unavailability, high price, or practicability; thus, compromise is necessary for the quantification of lipids. This has led to the implementation of alternative quantification strategies (surrogate quantification) with the aim of reducing the overall number of standards, measurements, and costs involved. Surrogate quantification can be accomplished either by using structurally similar standards, or by using isotopically labelled IS or non-endogenous IS. The most common practice is to add a mixture of IS, one for each class of lipids, sometimes called the “*one-internal standard calibration*” approach (IS-calibration), and use their signals to calculate the concentrations of lipid species of the same class [13]. This strategy of quantification drastically reduced the number of necessary standards, but the accuracy of quantification was compromised. The application of response factors based on mathematical models can correct the differences in the ionization between the surrogate and the target lipid species and improve the accuracy [11, 14], but these corrections also require multiple standards and careful validation, and there is no definitive proof of the stability of factors over time or between MS equipment.

Besides, surrogate quantification by the IS-calibration approach not only compromises absolute quantification but also

affects the intralaboratory reproducibility. This is because the IS-calibration approach is highly dependent on the stability of the IS-mix, as it is used to calibrate, to make up for the losses of the analytes during sample preparation, and to correct the presence of matrix effects on individual species. Unfortunately, it is not easy to guarantee the stability of the lipid IS-mix dissolved in highly volatile organic solvents, which leads to poor reproducibility of the results between different analytical batches, even within the same laboratory. This problem is clearly noticeable when lipidomics is applied in large clinical studies that involve the analysis of many specimens in many batches within the same laboratory and in interlaboratory comparisons. For example, in a recent validation exercise of a ready-to-use commercial kit, the intralaboratory within-batch variability reported by the 14 participants ranged from 5.6% up to a maximum of 41% for lipid species of several classes [15]. In the same study, the reported interlaboratory variability ranged from 11.3% up to a maximum of 306%. The same accuracy problems have been shown in the Interlaboratory Lipidomics Comparison Exercise (ILCE) [16] in which 31 laboratories worldwide, with different MS-based methods, found significant disparities in the lipid concentrations for the same plasma reference (National Institute of Standards and Technology Standard Reference Material-1950 (NIST SRM-1950)). Therefore, the difficulties in lipidomic workflows to cope with absolute quantification translate into two major problems: a lack of intralaboratory reproducibility and a poor interlaboratory comparison.

Motivated by these drawbacks, the lipidomic community has produced recommendations focused on improving harmonization of the results issued by the different laboratories covering different aspects related to the analytical process [17]. In particular, to improve intra- and interlaboratory reproducibility, it is encouraged to perform a validation of the lipidomic workflow adapting published bioanalytical guidelines [18, 19]. This validation should cover aspects such as the within- and between-batch reproducibility of the measurement, the limit of quantification, and linearity. Hence, the inclusion of a quality control (QC) strategy with the inclusion of QC samples at different concentrations and the use of a long-term reference (LTR) to graduate or adjust the deviation of the measurement procedure from the expected values are advised. The LTR closely resembles the function of an external calibrator, but the fundamental difference is that its values are not absolute concentrations. Indeed, a recent study has shown the benefits of using a common LTR to reduce interlaboratory variability in the quantification of 75 species belonging to several lipid classes, which was carried out in two laboratories using different LC methodologies (HILIC and RP) and direct infusion [20]. We believe that the benefits of this strategy can be extended to improve the intralaboratory variability.

Ensuring quantitatively reliable results in lipidomics is difficult not only due to the dispersion of technology and workflows but also due to the need for specialized software

tools to trace errors in the quantification of a large number of species and samples. There is a software tool that compares the results produced by a laboratory with those of the NIST-ILCE consensus [13, 21], but there is no free and easy-to-use tool that provides an easy follow-up of QC tasks and visualizes the between- and within-batch stability of the lipidomic measurements over time.

The main objective of this study was to demonstrate that the use of an external calibrator in combination with the addition of an IS for each class improves the intralaboratory reproducibility of lipidomics performed by RPLC-MRM/MS in our laboratory. As an external calibrator, we use a LTR that contains the estimated concentrations of the lipid species of interest, whose accuracy values were first compared to those provided by the NIST SRM-1950 consensus values. Subsequently, the LTR is used as a calibration material in successive batches of sample analyses to perform matrix-matched single-point *normalized signal calibration* (NS-calibration) [22]. We demonstrate that the use of this strategy reduced the between-batch variability, both under controlled conditions in the validation phase and in real test conditions performed on 120 patient samples. In addition, we provide a free access web tool, which delivers the concentration of lipid species by the two alternative quantitative approaches, IS-calibration and NS-calibration, and visualizes the results of the QC in lipidomics.

Materials and methods

Materials

Cholesterol, cholesterol-d7 (d7-FC), cholesteryl stearate (CE 18:0), cholesteryl-d7 palmitate (d7-CE 16:0), D-glucosyl- β -1,1'-N-dodecanoyl-D-erythro-sphingosine (HexCer 30:1 [d18:1/12:0]), N-dodecanoyl-D-erythro-sphingosylphosphorylcholine (SM 30:1 [d18:1/12:0]), N-dodecanoyl-D-erythro-sphinganylphosphorylcholine (dhSM 30:0 [d18:0/12:0]), 1-heptadecanoyl-2-hydroxy-sn-glycero-phosphocholine (LPC 17:0), 1-stearoyl-2-hydroxy-sn-glycero-phosphocholine (LPC 18:0), 1,2-dipalmitoleoil-sn-glycero-3-phosphoethanolamine (PE 32:2 [16:1/16:1]), 1,2-dimyristoleoil-sn-glycero-3-phosphocholine (PC 28:2 [14:1/14:1]), PC 32:1 [18:1/14:0], PC 34:1 [16:0/18:1], PC 36:4 [16:0/20:4], and PC 38:6 [18:0/22:6] were acquired from Avanti Polar Lipids (Alabaster, AL, USA). N-heptadecanoyl-D-erythro-dihydrosphingosine (dhCer 35:0 [d18:0/17:0]), N-tetracosanoyl-D-erythro-sphingosine (Cer 42:1-d7 [d18:1/24:0]), N-stearoyl-D-erythro-sphingosine (Cer 36:1 [d18:1/18:0] and deuterated N-C18:0-d3), and N-stearoyl-D-erythro-dihydrosphingosine (dhCer 36:0 [d18:0/18:0] and deuterated N-C18:0-d3) were from Matreya LLC

(PA, USA), and 1,3-diolein-2-decanoyl sn-glycerol (TG 46:2 (18:1/10:0/18:1)) and triolein (TG 54:3 [18:1/18:1/18:1]) were from Larodan (Monroe, USA). Only LC-MS quality or HPLC grade solvents were used, acetonitrile and isopropanol (VWR, Merck); acetone, dichloromethane, and chloroform (Sigma-Aldrich). Ammonium formate and other reagents were from Sigma-Aldrich.

Preanalytics

Patient samples were collected in EDTA-K3 tubes. The plasma was immediately separated from erythrocytes by centrifugation (1900 g for 15 min, 4 °C), aliquoted, and stored at -80 °C until processing. QC pools (QC-low, QC-medium, and QC-high) and the LTR external calibrator were prepared by pooling residual human plasma after routine clinical diagnostics. Total cholesterol (total-COL) and total triacylglycerols (total-TG) were measured by enzymatic assays on an Abbott-Architech c16000 analyzer. They were aliquoted in 50- μ L plastic Eppendorf tubes and immediately frozen at -80 °C. As expected, when the QCs were analyzed, it was found that different lipid class concentrations in these QC samples were correlated with the concentrations of total-COL and total-TG, as in general, the different lipid species circulate together, bound to plasma lipoproteins [23]. The final estimated concentration values of lipid classes are shown in Supplementary Information (ESM) Table S1. The concentration of individual lipid species can be downloaded from the web application (see below for more details).

The NIST SRM-1950 plasma was delivered frozen in 1-mL tubes (Sigma-Aldrich) and stored at -80 °C. Frozen samples were always slowly thawed on ice for 1 h and carefully vortexed before proceeding with the extraction.

Patient samples included in the evaluation of robustness were part of the study: "Predictive factors of histological lesion in patients with non-alcoholic fatty liver disease." The study received approval from the clinical research ethical board of Hospital Ramón y Cajal (EC 276-14).

Extraction of lipids from plasma

Lipids were extracted following the method described by Folch et al., with minor modifications [24]. Briefly, 50 μ L of plasma was mixed vigorously for 2 h with 450 μ L of 0.88% NaCl, 2 mL of chloroform:methanol (2:1 vol/vol), and 10 μ L of an IS cocktail (IS-mix) prepared in chloroform:methanol (2:1 vol/vol) composed of PC 28:2 (14:1/14:1), PE 32:2 (16:1/16:1), LPC 17:0, dhCer 35:0 (d18:0/17:0), Cer 42:1-d7 (d18:1-d7/24:0), HexCer 30:1 (d18:1/12:0), SM 30:1 (d18:1/12:0), dhSM 30:0 (d18:0/12:0), TG 46:2 (18:1/10:0/18:1), d7-CE 16:0, and d7-FC. The concentrations of the IS-mix are given in ESM Table S2. Positive pressure plastic tips were used for addition

of the IS-mix, as we observed that they improved precision with organic solvent additions. After centrifugation (1900 g, 5 min, 4 °C), the lower phase was obtained using a glass Pasteur pipette. The upper phase was re-extracted again with 1.5 mL of chloroform, shaken for 30 s, and centrifuged (1900 g, 5 min, 4 °C), and the lower phase was mixed with the previously collected lower phase. The extract was gently dried under a nitrogen gas current and immediately reconstituted for LC-MS analysis.

LC-MS/MS analysis

The dry lipid extracts were dissolved in 250 µL of acetonitrile:isopropanol (1:1, vol:vol), sonicated for 10 min, and then transferred to an injection vial. To fit peak area signals within the linear dynamic range for all lipid species, it was necessary to perform three separate injections in the LC-MS system. The LC system was an Eksigent UltraLC-100. Five microliters was injected for the analysis of Cer, dihydroceramide (dhCer), hexosylceramide (HexCer), and dihydroSM (dhSM) species, whereas 2 µL of 1:5 diluted lipid extract was injected for the analysis of PC, lysoPC (LPC), PE, SM, and TG. The analysis of free cholesterol (FC) and cholesterol ester (CE) species was carried out by injecting 5 µL from the 1:5 diluted extract. Lipids were separated on a Kinetex C18 column (100 × 2.1 mm, 1.7 µm; Phenomenex, Macclesfield, UK) maintained at 55 °C and a binary gradient of phase A (60% acetonitrile in water, 10 mM ammonium formate) in B (90% isopropanol in acetonitrile, 10 mM ammonium formate) from 60% A up to 100% B for 12 min, with an additional 8 min of re-equilibration to the initial conditions. A flow rate of 0.4 mL/min was applied. The detection of lipid species was carried out following a targeted approach setting MRM transitions for each lipid species at their retention times (ESM Table S3). Detection and analysis of lipid classes were carried out on a QTrap 4000 (AB-Sciex LLP, Framingham, MA, USA) with Analyst 1.6.2 software. Nitrogen was employed as the drying gas at a temperature of 500 °C, the curtain gas was set at 30 psi, and the ion source gas was set at 50 psi. The detection was set in ESI positive mode for all lipid classes with the exception of CE and FC, which were analyzed using the atmospheric pressure chemical ionization (APCI) source in the positive ion mode.

Quantification of lipid species

The concentration of lipid species was obtained from the peak areas using two calibration strategies that have been excellently reviewed [22]. The first, denoted as *IS-calibration*, is based on the addition of one IS per lipid class. The concentration of each individual lipid species in a particular sample was estimated by multiplying the surrogate signal-concentration ratio (calibration factor, CF) of the class by the corresponding area

of the target species (Eq. 1). The accuracy of the quantification by IS calibration relies on the ability of the IS to act as a surrogate for the lipid species of the class.

Internal standard calibration.

$$\text{Conc}_i^m = \text{CF}_{IS}^m \cdot \text{Area}_i^m \quad (1)$$

where,

$$\text{Calibration factor; } \text{CF}_{IS}^m = \frac{\text{Conc}_{IS}^m}{\text{Area}_{IS}^m}$$

i lipid species from class
m sample
IS internal standard of class
CAL calibrator
Conc concentration.

When the CF changes from one lipid species to another within the same class, the application of the *IS-calibration* requires normalization of the measured analytical signals by determining a normalization signal factor (NS) for each species in relation to a reference compound. This calibration methodology is called *NS-calibration*.

Internal standard normalization.

$$\text{Conc}_i^m = \text{CF}_{IS}^m \cdot \text{Area}_i^m \cdot \text{NS}_i^{\text{CAL}} \quad (2)$$

where,

$$\text{Normalization signal factor; } \text{NS}_i^{\text{CAL}} = \left(\frac{\text{Area}_{IS}^{\text{CAL}} \cdot \text{Ref. Conc}_i^{\text{CAL}}}{\text{Area}_i^{\text{CAL}} \cdot \text{Conc}_{IS}^{\text{CAL}}} \right)$$

Ref. Conc_{*i*}^{CAL} Reference concentration of species *i*, in the CAL.

Note that when NS_{*i*}^{CAL}=1, then Eqs. 1 and 2 are identical.

The NS values of the analytes depend on the individual species. This is a methodology of double calibration, including external-internal, which is applied in two steps. In the first step, a single external standard measurement determines the normalization factors of the lipid species in relation to the IS (Eq. 2). This is done by adding the IS to a reference material (calibrator) in which the concentrations of lipid species are known. Secondly, the concentration of each lipid species is calculated by multiplying the IS concentration by the estimated NS.

Following the definition of the *type of quantification* proposed by the Lipidomics Standard Initiative (LSI) [9], the quantifications performed by both the *IS-calibration* and *NS-calibration* are estimated concentrations, which are derived from the use of a non-matching IS (= other lipid class than analyte or no co-ionization of analyte and IS) and provide a *level 3* type of quantification. Although the *NS-calibration* considers the species-specific analytical response, there is no co-ionization of the IS and the species of interest. Therefore, *NS-calibration* does not fit the proposed definition of level 1 type of quantification.

Quantification by external standard calibration with internal standardization

Concentrations of a selected subset of lipid species were estimated from multilevel non-matrix-matched external standards with internal standardization curves. These calibration curves were prepared from serial dilution of commercially available lipid standards. The concentrations used are given in ESM Table S4. The IS-mix added in standard curves had the same concentration as that used to obtain the concentration in plasma samples, with the exception of dhCer 36:0 and Cer 36:0, which were normalized by adding 10 μ L of isotopically labelled dhCer 36:0-d3 (d18:0/18:0-d3) and Cer 36:1-d3 (d18:1/18:0-d3) at final concentrations of 1 and 2.5 nmol/mL, respectively.

Processing of LC-MS chromatograms and statistical analyses

Raw LC-MS/MS chromatogram files in Sciex format (.wiff) were imported into Skyline software (ver. 19.1.0.193) [25], and the areas corresponding to lipid peaks were integrated. Isotopic type I correction was applied over peak areas, as explained previously [26]. The corrected isotopic areas were transformed into concentrations of lipid species by the IS- and NS-calibration approaches using self-programmed scripts in R (version 3.0.2).

Normalized coverage equivalent

In order to check the bias of SRM-1950 concentrations for lipid species measured using our methodology and the consensus values reported in the NIST-ILCE exercise (16), we used the normalized coverage equivalents (also known as Z-score), as proposed by Ulmer et al. [21]. The dots represent the average values obtained for each lipid species, and the whiskers represent their normalized standard deviation (SDnorm). This type of representation has the advantage of normalizing very different concentrations of lipids within the same plot, providing an easier look at values beyond the acceptance limits.

Normalized coverage equivalent.

$$Z\text{-score} = \frac{\text{Consensus mean} - \text{Measured mean}}{\text{sdu}} \quad (3)$$

Normalized standard deviation.

$$SD\text{norm} = \frac{\text{sd}}{\text{sdu}} \quad (4)$$

The upper and lower confidence limits (UCL and LCL, respectively) were defined by three times the standard deviation uncertainty (sdu) values given by the consensus.

Process control analysis

In order to trace the stability of the IS-mix added for each class between and within different batches of analysis, we constructed Shewhart's control charts [27]. The area of the IS for each sample was compared to the mean IS area (Eq. 5) within the same batch, and the confidence limits were defined in Eq. 6:

Mean IS area.

$$\overline{\text{IS}} = \frac{\sum_{m=1}^n \text{Area}_{\text{IS}}^m}{N} \quad (5)$$

Confidence limits.

$$\text{Confidence Limits} = \overline{\text{IS}} \pm 3 \cdot \frac{\overline{\text{mR}}}{1.128} \quad (6)$$

where the $\overline{\text{mR}}$ (mean IS moving range, Eq. 7) in a batch is calculated from the individual areas of IS in each sample ($\text{Area}_{\text{IS}}^m$) as follows:

Mean IS area moving range.

$$\overline{\text{mR}} = \frac{\sum_{m=1}^{n-1} |\text{Area}_{\text{IS}}^m - \text{Area}_{\text{IS}}^{m+1}|}{n-1} \quad (7)$$

n number of samples in each batch of analyses
 m sample

During our process control analysis, the batch (extraction and injection), the run order (order in which the samples were injected into the LC-MS system), and the sample type (QC, LTR, or sample) were registered. A typical sequence, corresponding to the evaluation of robustness, is given in ESM Table S5. The Shewhart's control charts given in ESM Figs. S3 and S4 show the order in which the batches, samples, and injections were carried out.

Data availability and web application

The raw area data of individual samples, QCs, calibration values, and lipid species concentrations are available on the web page <https://clipidomics.com/QCTool>. Select the *demonstration mode* to access data from evaluation of the within- and between-batch precision (QC pool evaluation) and the robustness evaluation of the workflow (robustness evaluation). All the results and datasets, and a detailed *User's Manual* describing the functionalities of the application can be downloaded from the same web page (<https://clipidomics.com/QCTool/Manual.pdf>).

Results

Accuracy of the LC-MS/MS methodology against the SRM-1950 material

As lipids are commonly present in the matrix sample, the accuracy of measurement poses a challenge as the assays cannot distinguish between an exogenously added lipid standard and the endogenous counterpart. To confirm the accuracy, we compared the concentrations obtained by our methodology using the IS-calibration approach (Eq. 1) with those provided by the NIST-ILCE for the SRM-1950 reference material [16]. It should be noted that the NIST-ILCE consensus mean values were reported using the sum composition annotations (i.e., PE 36:4); therefore, the areas of isobaric species of the same lipid class were added before applying the quantification. We did not include PC and PE ether-linked lipid species (plasmeyl ethers or plasmeyl ether-linked phospholipid species) in our targeted method as the analysis of these species introduces an important layer of uncertainty (i.e., PC O-38:3, PC O-38:2, and PC 37:3 are isobaric species), not fully reachable using our instrumentation without sacrificing the quantitative performance of the approach [28]. In the same way, our analysis only covers oleic acid (FA 18:1)-containing TG species, as we only followed MRM transitions of TG species corresponding to a neutral loss scan (NLS) at $m/z = 299$ ($M^+ \rightarrow [M+H-299]^+$) (ESM Table S3). To improve the accuracy, we added a specific IS for the quantification of dhCer and dhSM, using a lower concentration than used for Cer and SM, because the magnitude of plasma concentrations of these species is much lower than their unsaturated counterparts. We also had to change Cer 37:1 by deuterated (Cer 42:1-d7) for plasma Cer quantification as we observed unreliable recovery of the former probably due to the presence of endogenous Cer 37:1 in plasma (data not shown).

Of 213 lipids from nine lipid classes (eleven classes if sphinganine-derived dhSM and dhCer are considered separate lipid classes) with established consensus values reported for the SRM-1950, we quantified 108 lipid species (Fig. 1a). Ninety species detected using our method did not match

species in the NIST-ILCE. For example, from eleven Cer species, only eight have consensus values reported in the NIST-ILCE SRM-1950 (Fig. 1b).

When we compared the bias of lipid concentrations to the NIST-ILCE values, using the normalized coverage equivalent, we found good accuracy for 69 of the 89 lipids quantified (Fig. 1c and ESM Table S6). However, a major bias was observed between the consensus reported values for the HexCer and LPC classes. To demonstrate the leverage of IS selection in the accuracy of values, we decided to re-analyze SRM-1950 samples adding IS-HexCer 30:1 instead of IS-HexCer 33:1, thus, obtaining a good agreement with the consensus means (ESM Fig. S1). Therefore, we choose IS-HexCer 30:1 for subsequent quantifications.

From this comparison, it is straightforward to obtain a set of correction factors to translate the concentration values for each lipid species to the consensus values (given in EMS Table S6). We also quantified lipid species in a self-prepared LTR plasma (LTR-calibrator) whose values will be used as a matrix-matched single-point external calibrator for further comparisons between the IS-calibration and the proposed NS-calibration.

It is important to stress that the use of the SRM-1950 consensus values as an absolute reference to check the accuracy of our results should be done cautiously, as their concentrations are derived from results that varied largely between laboratories, and not on actual absolute molar concentrations. Therefore, to test how close our values were to the true values, we performed an additional multilevel external standard calibration with internal standardization (ESM Fig. S2). In general, there was good agreement between the results obtained for the selected lipid species between the external calibration and the single-point IS-calibration approach (ESM Table S7). In addition, the results of these lipid species were, in general, close to those previously estimated by other authors using different extraction and MS settings [15, 29].

Linear dynamic range

The LLOQ and ULOQ define the linear range of the assay, which is the range of concentrations where the analyte signals are directly proportional to the concentration. Typically, the linear range is obtained by diluting samples of high concentration in an acceptable matrix diluent, free of the analyte. In lipidomics, the problem arises from the lack of a lipid-free biological matrix. An experimental alternative is to use the IS as a surrogate analyte and then vary its concentration in the sample matrix to generate a standard curve as performed by Wolrab et al. [29]. While a reasonable alternative, this can give rise to questions about the differences in the chemical noise that may be present at the m/z values and retention time for the surrogate analyte versus the target analytes. To overcome this limitation, we prepared a 20-fold concentrated extract (from a plasma pool) without adding IS. The extract was successively diluted to 1:500 in the injection

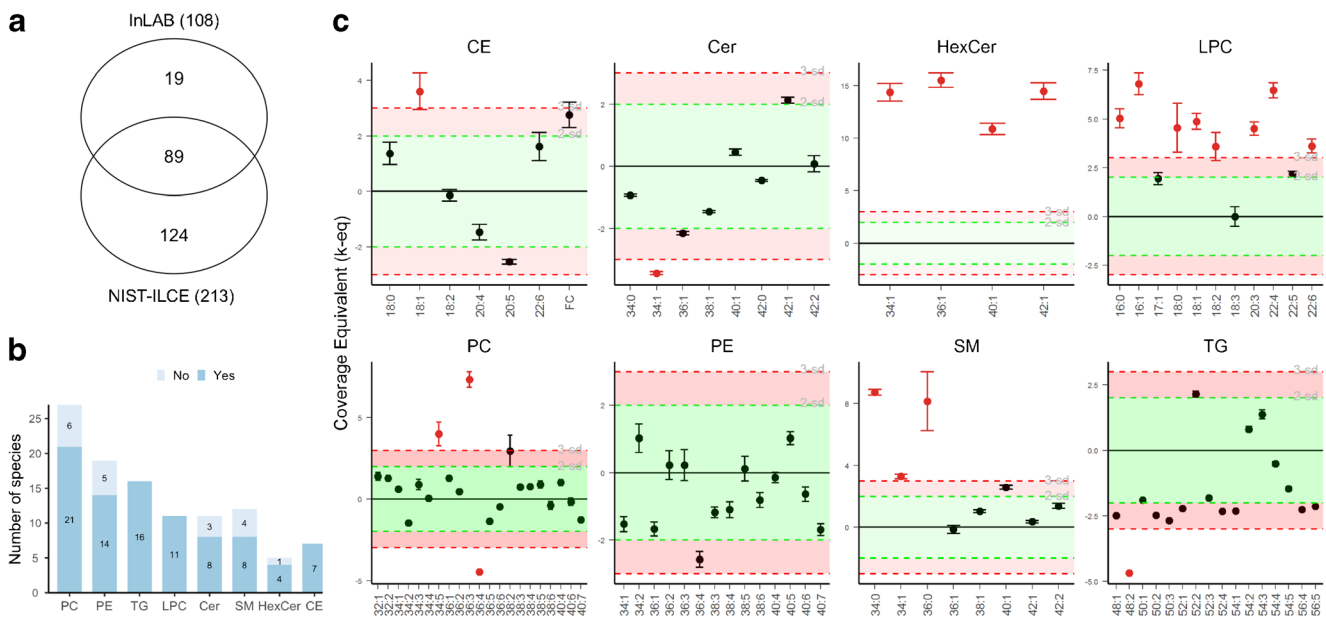


Fig. 1 Evaluation of the accuracy in the concentration of lipid species. **a** The NIST SRM-1950 plasma was analyzed in the laboratory ($n = 5$). The intersection diagram shows the number of matching species that were reported by the Interlaboratory Lipidomics Comparison Exercise (ILCE) and the species detected by our methodology (inLAB). **b** Number of species detected by our methodology classified by lipid class, whose values were present in the NIST-ILCE. **c** Bias of inLAB values to

the NIST SRM-1950 mean estimate consensus values. Results are presented as normalized coverage equivalents (k-eq) at the mean (dots) and standard deviation (bars). The concentrations of species were obtained using the IS-calibration approach. The black solid line represents the consensus mean. The green region represents the 95% (2-sd) expanded uncertainty interval for the NIST-ILCE consensus mean; the red dotted lines bound the 99% (3-sd) expanded uncertainty

buffer containing the same amounts of IS. It was expected that the deviation of the concentration of the dilutions did not exceed more than 30% of the value corresponding to the original sample (1x), which was quantified using the original extraction protocol from 50 μ L of sample. This approach allowed us to adjust the methodology within the linear dynamic range for the lipid species (ESM Table S6). Some lipid species such as Cer 42:1 showed a wide dynamic range (Fig. 2), whereas others, for example SM 34:1, demonstrated a narrower dynamic range. This exemplified the difficulties in choosing optimal analytical conditions within the acceptance dynamic range for all lipid species.

Evaluation of the within-batch and between-batch precision

In order to evaluate the within-batch and between-batch precision, we prepared three plasma sample pools denoted as (QC-low, QC-medium, and QC-high) based on the concentration values of total-TG and total-COL performed by routine clinical chemistry methods. These QC pools were processed in four consecutive batches, using the same IS-mix concentrations, but different IS-mix preparations, sample extractions, and LC runs. We also included three replicates of the LTR-calibrator in each batch to perform a NS-calibration approach and compared the precision obtained by the two approaches.

The concentration of lipid species was first obtained by the IS-calibration approach (Fig. 3). A systematic check of data quality can be made by performing a principal component

analysis (PCA) on the complete dataset. When plotting the first two principal component scores (a projection describing the maximum orthogonal variance in the data) and labelling the data points either as the QC-level and their analysis batch (Fig. 3a), we observed a significant between-batch dispersion for all the QC levels, shown by a 11.3% of explained total variance by principal component 2 (PC 2). The total dispersion in the concentration of lipid species, exemplified in the QC-medium, demonstrated a significant deviation from the average values with many individual lipid species measurements out of acceptable margins (Fig. 3b). The relative standard deviation (RSD)-within batch was below 20% for all lipid species, but 18 of 106 species showed a RSD-between batch dispersion above 20% (Fig. 3c). The lack of between-batch precision was especially noted for TG, PC, and dhCer classes (Fig. 3d). We attributed this to poor between-batch IS area stability for some classes, as exemplified for Cer and dhSM (ESM Fig. S3). In contrast, the quantification based on NS-calibration showed excellent stability between the different analytical runs, clearly noticeable when comparing the RSD (total- and between-batch) of classes (Fig. 4). Similar results were observed for QC levels low and high, which are accessible at <https://clipidomics.com/QCTool>.

Evaluation of the robustness of the methodology

The robustness of an analytical method is the resistance to change in the results produced by an analytical method when

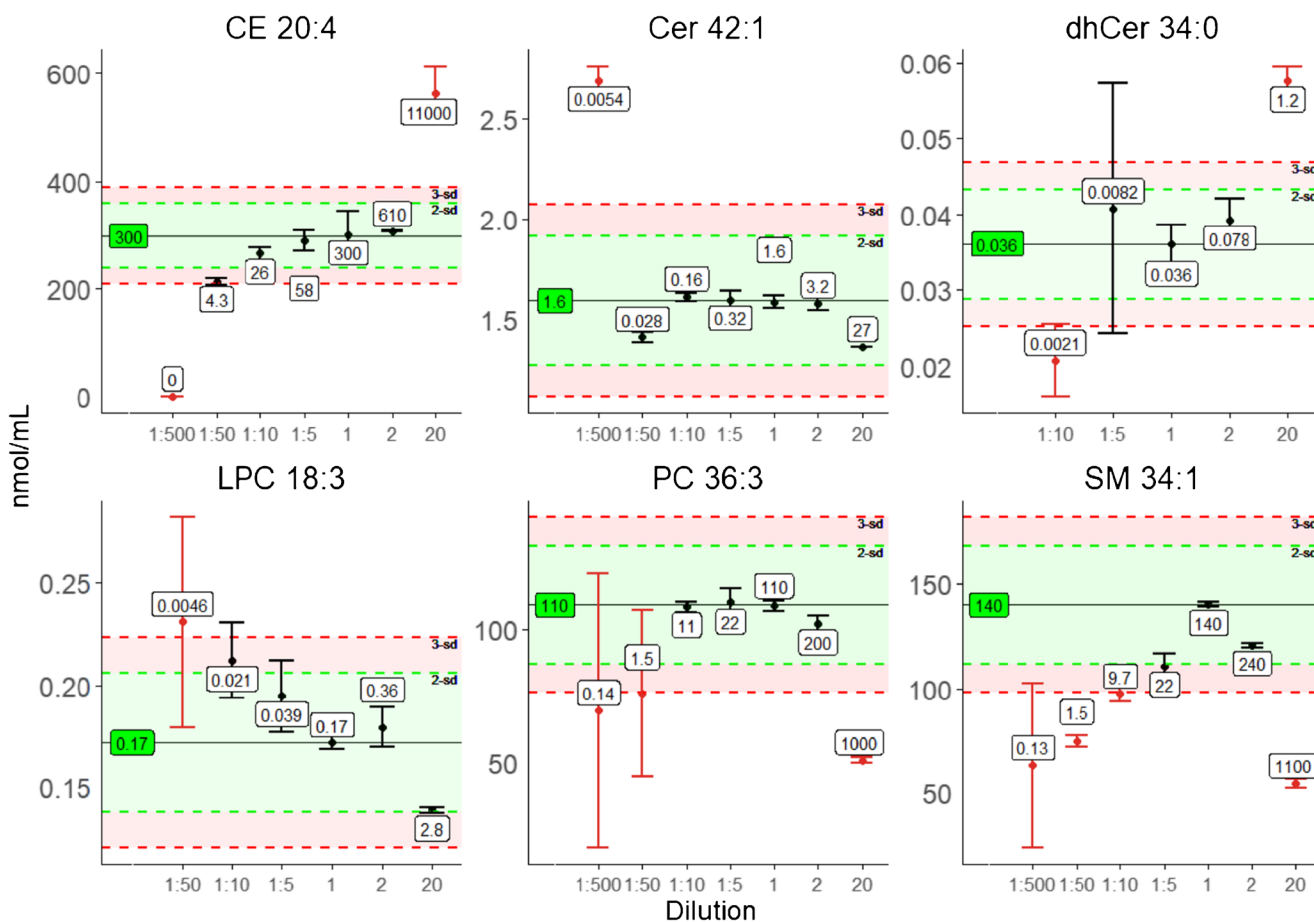


Fig. 2 Linear dynamic range of representative quantified lipid species. We plot the concentrations of representative lipid species (one for each lipid class) after a serial dilution of a concentrated plasma pool extract from 20x to 1:500x. The same amount of IS-mix was added to all dilutions and the concentration of species was obtained by the IS-calibration approach. The mean target value (black horizontal line) was evaluated (three replicates) using the standard protocol from 50 μ L of the plasma

pool. The standard deviation (sd) was set at 10% of the target mean and the upper and lower intervals (red line limits) fixed at 3-sd (30%). Points represent the mean and bar the standard deviation of three replicates ($n = 3$) done for each dilution. Values out of the acceptance intervals are colored in red. Labels correspond to the actual concentration values of lipid species obtained for every dilution

minor deviations are made from the experimental conditions described in the procedure [30]. The reliability of the quantification made by the IS-calibration approach rests on the stability of the IS area values for the different lipid species throughout the same, and between different, analysis batches. Therefore, the stability of the IS-mix solution is a critical point likely to affect the results of quantification. Based on our previous results (ESM Fig. S3), the IS-mix solution prepared with organic solvents, in our case methanol:chloroform (1:2, vol:vol), was prone to rapid evaporation, which affected the between-batch reproducibility of the process. As a consequence, different batches of IS preparation might differ slightly in the concentration of species, even when a careful preparation protocol is followed.

We decided to examine the robustness of the methodology, testing the effect of larger sample batch preparations on the real measurement conditions, provided within the analysis of 120 plasma samples. To minimize the variability, we prepared a single batch of IS-mix solution following a standard procedure of

aliquoting and storing at -80°C until use. The analysis of plasma samples was performed on three separate batches, each batch containing three replicates of QC. We also included three replicates of the LTR-calibrator SRM-1950 on each batch to perform quantification using the NS-calibration approach. The IS areas demonstrated better stability for different analytical runs within classes (ESM Fig. S4), yet the quantification made by the NS-calibration demonstrated even better reproducibility (Fig. 5). This improvement was clearly noticeable in the between day variability when comparing the distribution of the number of species and classes according to their RSD (ESM Fig. S5).

Comparison with clinical chemistry validated methods

We compared our methodology with the results provided by fully validated clinical chemistry methods for total-COL and total-TG. To this end, we added the concentration of FC and

Internal Standard Calibration (IS-calibration)

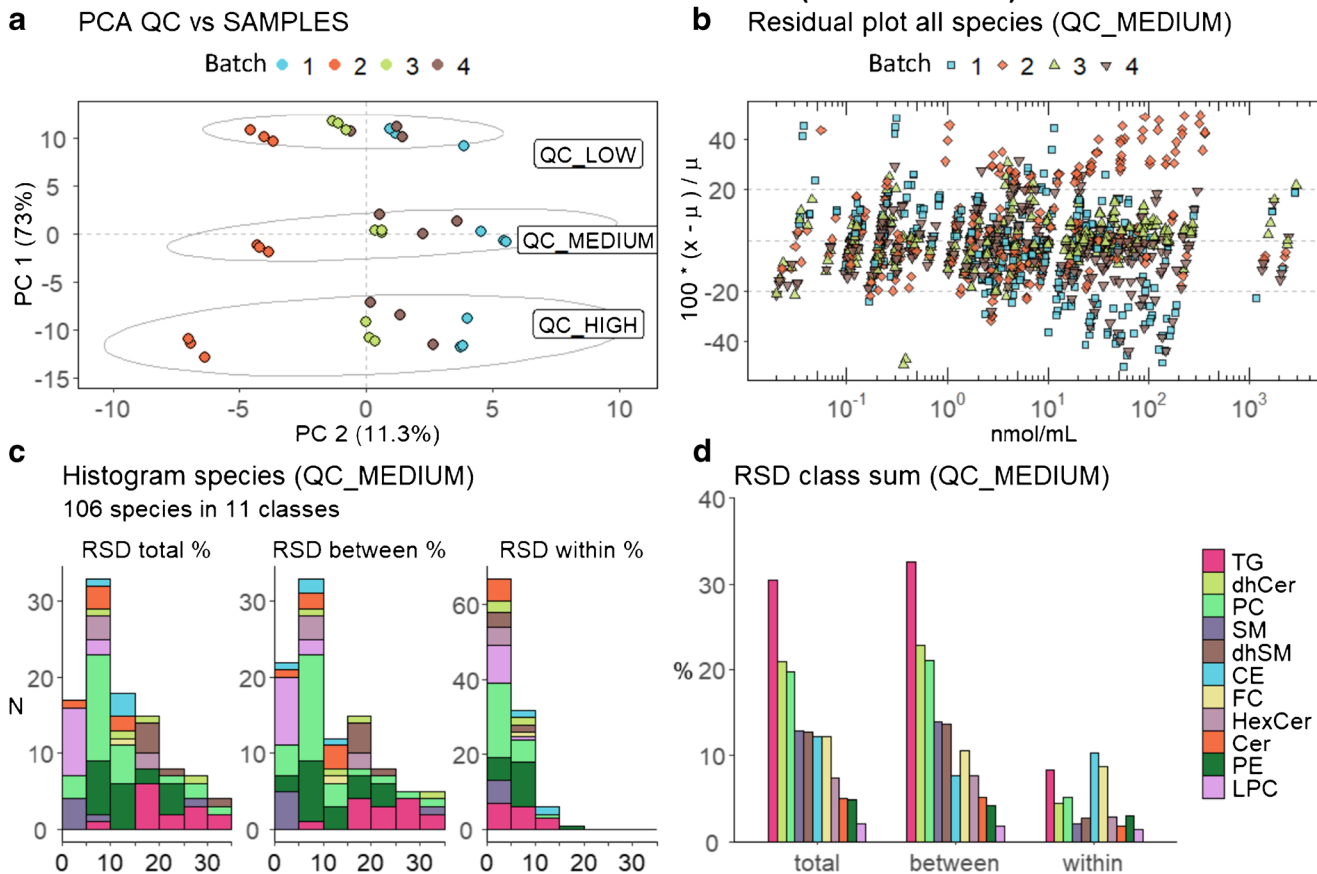
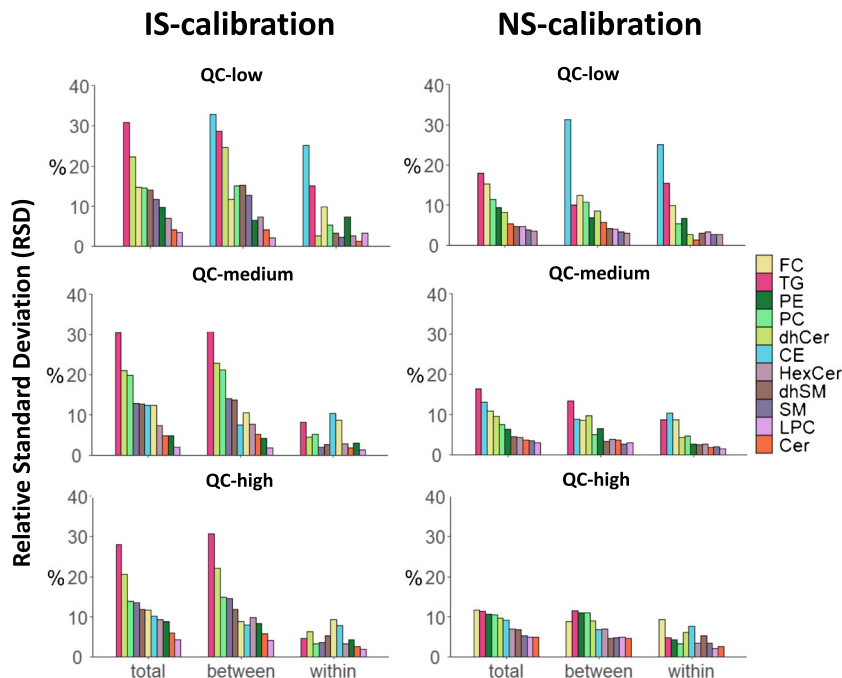


Fig. 3 Evaluation of the within-batch and between and batch precision by the IS-calibration approach. **a** Principal component analysis (PCA) score plot of quality controls (QCs) in different batches. **b** Residual plot corresponding to QC-medium. Each point represents the normalized concentration for each lipid species obtained in three replicates in four different

batches. The horizontal lines show the percentage of deviation from the average values. **c** Histogram of the distribution of lipid species measured in QC-medium according to the values of imprecision, and expressed as relative standard deviation (RSD), total, between batch, and within batch. **d** RSD total, between, and within batch for lipid classes

Fig. 4 Comparison between precision obtained by IS- and NS-calibration approaches. Relative standard deviation (RSD) of sum's lipid classes obtained for each quality control (QC)-level (low, medium, and high). Sum concentrations of class were obtained from the concentration of individual lipid species calculated by both the IS-calibration and the NS-calibration approaches



CE species performed by lipidomics. In the same way, we compared the class concentration of TG from lipidomics with the routine method. The molar concentrations of total-COL determined by MS and by clinical chemistry were of the same order of magnitude (Fig. 6a), whereas total-TG was much lower, which could have been due to limitations in the number of TG species followed by our targeted approach (Fig. 6b). Correlation with the clinical chemistry methodology was better for TG ($R = 0.82$) than for COL ($R = 0.752$).

Discussion

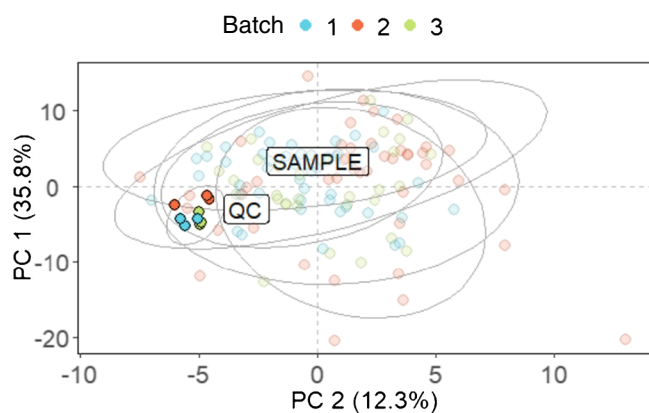
The goal of the new field of clinical lipidomics is to apply the advances in MS, allowing the quantification of hundreds (even thousands) of lipid species in clinical practice. Contrary to advances in genomics, evidence shows that there are few cases in which proteomic and metabolomic studies have been translated into a successful clinical application [31, 32]. A systematic review of 107 biomarkers provided

by *-omic* studies prior to 2006 found that despite having generated the publication of more than four thousand scientific articles, only three of these articles show some progress on their clinical utility [33]. There are two main reasons why *-omic* clinical applications (including lipidomics) develop this way. The first is that lipid abundance in clinical studies is often expressed as qualitative fold changes, and the second reason is that reported lipid concentrations widely differ between laboratories [16]. Expressing lipidomic results in molar concentrations has the great advantage of allowing the comparison of results issued on different platforms and approaches regardless of the methodology or platform used and guarantees the reliability of the conclusions derived from the study. As some authors have warned, if the lipidomic data that we deliver are not quantitative and reproducible, they lack comparability and their contribution to systems biology (also clinical chemistry) is bordering on being meaningless [12, 34].

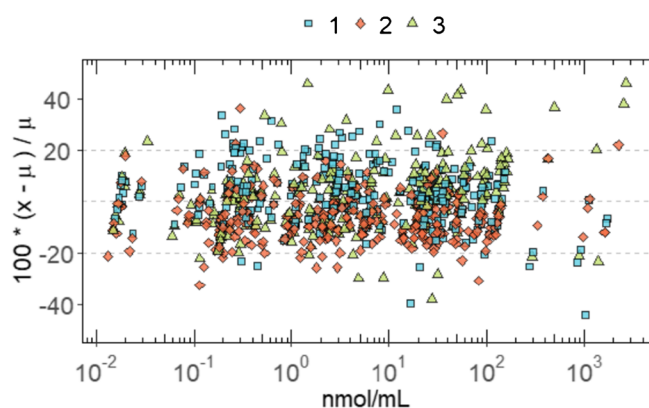
Reliable quantitation is an absolute requisite in clinical chemistry, and specific conditions need to be fulfilled along the entire lipidomics workflow in order to qualify the data as

Internal Standard Calibration (IS-calibration)

a PCA QC vs SAMPLES

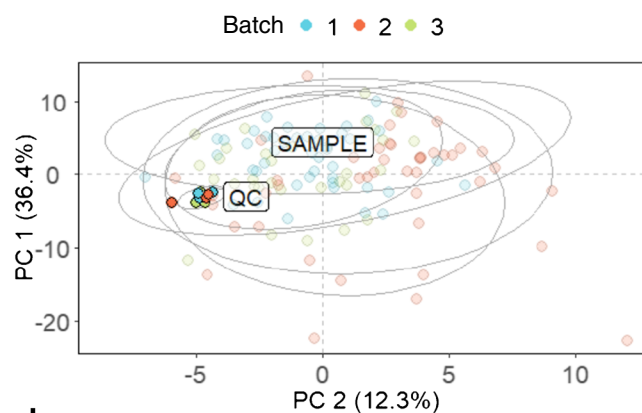


c Residual plot all species (QC_PL)



Normalized Signal Calibration (NS-calibration)

b PCA QC vs SAMPLES



d Residual plot all species (QC_PL)

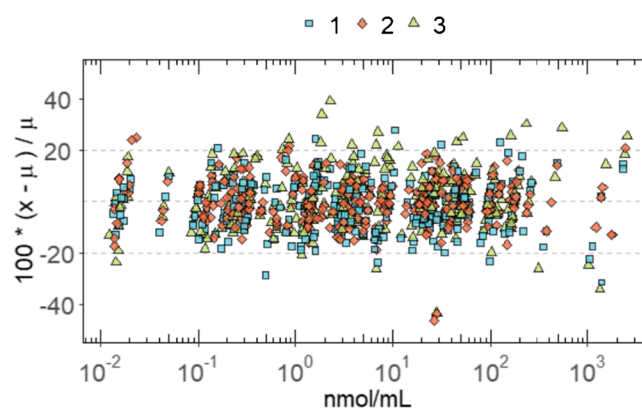
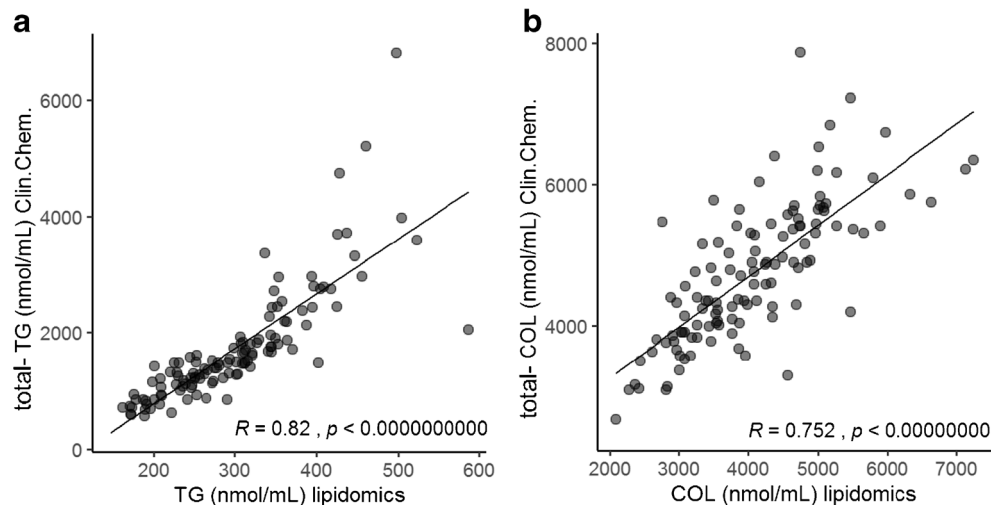


Fig. 5 Robustness of the QC concentrations throughout the analysis of 120 samples. Concentration of individual lipid species calculated by both the IS-calibration and the NS-calibration approaches for 120 samples analyzed in three batches; each batch included three replicates of QC. **a**,

b PCA score plot of QCs and samples in different batches by the two calibration approaches (IS- and NS-calibration). **c**, **d** Residual plot showing the observed individual species dispersion in the QC analyzed with the two approaches

Fig. 6 Correlation between lipidomic results with clinical chemistry values. **a** Total triglycerides (total-TG) and **b** total cholesterol (total-COL) for 120 plasma samples were analyzed using a routine clinical chemistry analyzer (Clin. Chem) and the results compared with the sum of lipid class concentrations obtained for TG and CE+FC performed by lipidomics using the NS-calibration approach (lipidomics). *R*, Pearson correlation coefficient; *p*, significance value



accurately quantitative. This implies follow-up of a series of recommendations in the pre- and analytical phase, as well as complying with a minimum of quality requirements, such as that produced by the International Lipidomics Society [17, 35]. In addition, the Food and Drug Administration and European Medicines Agency position is that lipidomic biomarker assays should address the same questions as method validation for individual molecule assays [18, 19]. Therefore, they must include an intralaboratory evaluation of the accuracy, precision, linear dynamic range, parallelism (comparison with already standardized methods), and long-term reproducibility (robustness) of the biomarker assay. In our study, we evaluated the accuracy of our methodology against the best available reference material, the NIST SRM-1950. Although the concentrations of lipid species reported for this material cannot be considered absolute, the NIST SRM-1950 is a first step in the harmonization of lipid plasma measurements on a worldwide scale and it is accessible to all laboratories.

Wolrab et al. demonstrated the importance of the IS-mix choice on the concentration of lipid species [29]. The authors performed an exhaustive validation of two methods, HILIC- and supercritical fluid chromatography-MS, using more than one IS for lipid class and comparing quantitative results for different IS-mix selections. The quantification was carried out by application of the IS-calibration. They found important differences in the concentrations in the NIST SRM-1950 for the lipid species, depending both on the reference IS to which the quantification of a class was performed, and on the IS-mix chosen. Therefore, carrying out an early comparison of accuracy with respect to this sample has the advantage of being able to make early decisions about the advisability of replacing some of the IS to gain harmonization in other laboratories, as we did in our study for those initially chosen for the quantification of Cer and HexCer classes. To double-check the accuracy of our quantification, we compared the lipid concentrations of the SRM-1950 material with the concentrations

obtained from non-matrix-matched external calibration curves, for a selection of lipid species. In general, our concentration values were in good agreement between the lipid species, and in comparison, with other exercises (see ESM Table S7).

Similarly, there is no accepted consensus on how to evaluate the linearity range of lipidomic measurements. Ideally, the linear dynamic range must be investigated using lipid standards in the presence of a real sample matrix [26], but due to the absence of an analyte lipid-free biological matrix, we devised a strategy which consisted of diluting a 20-fold concentrated extract. This approach allowed us to estimate the linear dynamic range for all lipid species over five orders of magnitude from 10^{-2} to 10^3 nmol/mL.

Once the accuracy of our methodology was verified, we evaluated the intralaboratory within- and between-batch precision using three levels of QC over 4 days. This evaluation was performed by simulating a worst-condition scenario choosing different IS-mix preparations for every batch and different sample extractions and injections into the LC, which were carried out on non-consecutive days. The stability of the IS areas for the different batches yielded a discrete between-batch reproducibility by IS-calibration method, with a total RSD above 20% for many lipid species especially TG, CE, and dhCer. In order to improve the precision between different batches of analysis and given that the main source of error identified was the imprecision of the IS-mix between days, we tried the alternative NS-calibration approach. The NS-calibration requires the inclusion of a single-point matrix-matched external calibrator of known concentrations in all the analysis batches. However, given the limitation of using SRM-1950 material directly as an external calibrator, due to its cost and limited availability, we decided to match the concentrations of a self-made LTR with the NIST SRM-1950. We subsequently used the LTR as an external calibrator adding to this reference, the QC and samples, the same amounts of IS-

mix, in every batch. This procedure is commonly used to improve the accuracy and precision for single analytes measured by LC-MS in clinical chemistry [36].

The use of a LTR, or a commercially available shared reference material, to obtain better harmonization between different laboratories and MS settings has been proposed in other studies [20, 37, 38]. These authors also suggested their use to improve the intralaboratory reproducibility in long-term longitudinal studies. In the work by Cajka et al. [38], the authors measured the between-batch reproducibility for some lipid species, but their concentration values and RSD were not reported. It has also been demonstrated that the intralaboratory accuracy and precision of lipidomic results is improved when a full method validation, which included the incorporation of a quality control strategy, was carried out [29]. Similarly, the application of the LTR in combination with the NS-calibration approach significantly improved the repeatability between days in the measurement of lipid species in the tested worst-condition scenario and was maintained even when a single IS-mix pool was used (best-condition scenario), as demonstrated in the robustness analysis performed on 120 patient samples. We attribute the better between-day to two main reasons. Firstly, the NS-calibration method does not depend on the concentration of the IS-mix; it only depends on the accuracy of the concentration of lipid species of the calibrator (in our case the LTR). As can be deduced from the application of Eq. 2, the concentration of each species of lipid added into the LTR, the QC, and samples in the batch are the same, $\text{Conc}_{IS}^{\text{CAL}} = \text{Conc}_{IS}^m$. Secondly, application of the normalization factor does not assume that matrix effects are constant between species of the same class, providing a “matrix environment,” of the calibrator, which is close to that of the real samples on every extraction. Thus, the advantages of an external calibrator and the addition of an IS are joined and the drawbacks of each one when applied separately are minimized [22].

Ensuring the quality of the quantification requires, in addition to having a reliable methodology, software tools to translate in a simple way the results of area or peak intensities of lipid species into concentrations. There are software packages which perform one or many of the steps implicated in lipidomic quantification: lipid identification (LipidXplorer) [39], signal integration (Lipid Data Analyzer) [40], signal quantification and normalization (LipidMatch Normalizer, Batch Normalizer) [41, 42], data analysis (Lipidr) [43], and comparison with the SRM-1950 (LipidQC) [21]. Some of the latest solutions provide an integration with fully developed software such as Skyline (LipidCreator) [44]. Providing a whole description of the capabilities of all these tools is beyond the scope of this paper [45], but in general none of these software packages is aimed at QC assurance. We decided to

develop an easy to use tool to carry out tasks such as monitoring the stability of the IS between samples, check the reproducibility and stability QC measures in the different batches, and compare the accuracy of values with historically stored values of QCs or the consensus reference material. It also provides a visualization of the concentration results of the individual samples and simple descriptive group statistics at the level of the class and lipid species. The results of our methodological evaluation are available on a web tool to allow those without programming skills to perform such tasks (<https://clipidomics.com/QCTool/>).

We recognize that our study has several limitations. First, despite the fact that our methodology demonstrated an improvement in robustness, the correlations demonstrated that the lipidomic concentrations for total-COL and total-TG, in comparison with clinical chemistry values, were discrete. The concentration values reported for TG by lipidomics were much lower than expected. We believe that this is due to the fact that only a partial quantification of TG species, those containing oleic acid, was quantified, which leaves many other TG species out. On the other hand, to obtain the results of total-COL by lipidomics, we decided to use atmospheric pressure chemical ionization (APCI), because it allows analysis of FC and CE species in a single LC run [46]. In our hands, this strategy seems very reliable in the analysis of FC, but it has limitations in the detection of CE species because the transitions used to trace CE species in the APCI mode overlap between them and the identification is based on their retention time. Nevertheless, the values obtained using our method were of the same order of magnitude as those reported by the clinical chemistry assay.

The second main limitation of our approach is that three independent injections in the LC-MS system are needed to analyze a single sample. While the introduction of faster MS instruments will allow the analysis of more comprehensive lists of lipid species on a single run, the limitation imposed by the different linear dynamic range of the lipid species does not preclude the need for careful adjustment in the dilution and the injection volume, to maintain good accuracy and precision in the results. Finally, it is important to recognize that the use of the SRM-1950 as an absolute reference for the accuracy for our results is debatable, and since their concentration values are based on a consensus and not on actual absolute values obtained for lipid species, they are not absolute molar concentrations. The concentrations provided correspond to a *level 3* type of quantification values [9] whose trueness rests in their similitude with those reported by other laboratories. In this sense, we could call them *harmonized molar concentrations*. However, the application of the NS-calibration method in combination with the proposed software tool will allow an immediate transformation of the results into absolute concentration values, to the extent that there are in the future true

absolute concentration values for the lipid species present in the NIST SRM-1950.

In conclusion, we demonstrate that the use of a matrix-matched single-point external calibration with internal standardization improves the intralaboratory robustness of our lipidomic workflow. The use of the NS-calibration approach in combination with the developed application will provide valuable information on the analytical variability, within- and between-day, batches or platforms, recognizing problematic lipid species and classes whose quantification is dubious. The use of this strategy in combination with the NIST SRM-1950 reference will help in the early detection of inaccuracy problems, thus enabling the continuous improvement and standardization of quantitative plasma lipidomics. Ongoing initiatives will improve the accuracy and coverage of lipid species reported by the NIST SRM-1950 and provide better reference materials, to harmonize lipidomic results. We hope that the workflow and tools described in this paper can help to improve the reliability of lipidomics.

Abbreviations *ACN*, Acetonitrile; *APCI*, Atmospheric pressure chemical ionization; *CE*, Cholesteryl esters; *Cer*, Ceramide; *CF*, Calibration factor; *COL*, Cholesterol; *dhCer*, Dihydroceramide; *dhSM*, Dihydrosphingomyelin; *EMA*, European Medicines Agency; *FDA*, Food and Drug Administration; *FC*, Free cholesterol; *HexCer*, Hexosylceramide; *HILIC*, Hydrophobic interaction liquid chromatography; *HRMS*, High-resolution mass spectrometry; *ILCE*, Interlaboratory Lipidomics Comparison Exercise; *IPA*, Isopropanol; *IS*, Internal standard; *LC*, Liquid chromatography; *LCL*, Lower confidence limit; *LLOQ*, Lower limit of quantification; *LPC*, Lysophosphatidylcholine; *MeOH*, Methanol; *MRM*, Multiple reaction monitoring; *MS*, Mass spectrometry; *NIST SRM-1950*, National Institute of Standards and Technology Standard Reference Material-1950; *NS*, Normalized signal; *PC*, Phosphatidylcholine; *PCA*, Principal component analysis; *PE*, Phosphatidylethanolamine; *QC*, Quality control; *RP*, Reverse phase; *RSD*, Relative standard deviation; *SM*, Sphingomyelin; *TG*, Triglycerides; *UCL*, Upper confidence limit; *ULOQ*, Upper limit of quantification

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00216-021-03364-x>.

Acknowledgements We thank the Quantification and Molecular Characterization Unit (UCA-CCM IRYCIS) for their technical help, and International Science Editing for English style revision.

Author contribution O. Pastor conceived and designed the study and the software tools. O. Pastor and B. Babiy performed analytical measurements and software tools. O. Pastor wrote the paper with inputs from B. Babiy and R. Busto.

Funding This work was supported by grants from the Instituto de Salud Carlos III, Spain (PI18/01152), the Ministerio de Ciencia e Innovación, Spain (RTI2018-098113-B-I00, Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020), and CIBER de Fisiopatología de la Obesidad y Nutrición, CIBEROBN, (OBN-18PI032018); all the grants co-financed by the European Development Regional Fund (ERDF). CIBER de Fisiopatología de la Obesidad y Nutrición is an initiative of ISCIII. B.B. is supported by pre-doctoral contracts (PEJD-2017-PRE/

BMD-4142 and PEJD-2019-PRE/BMD-15962) of the Comunidad Autónoma de Madrid (CAM). R.B. is a researcher from FIBio-HRC supported by Consejería de Sanidad (CAM).

Data availability All the raw data of this manuscript is available at the web page <https://clipidomics.com/QCTool>.

Declarations

Statement of ethical approval The study received approval from the clinical research ethical board of Hospital Ramón y Cajal (EC 276-14).

Conflict of interest The authors declare no competing interests.

References

1. Quehenberger O, Dennis EA. The human plasma lipidome. *N Engl J Med*. 2011;365(19):1812–23.
2. Lipidomics Standards Initiative Consortium. Lipidomics needs more standardization. *Nat Metab* 2019;1(8):745–747.
3. Vvedenskaya O, Wang Y, Miranda Ackerman J, Knittelfelder O, Shevchenko A. Analytical challenges in human plasma lipidomics: a winding path towards the truth. *Trends Analyt Chem*. 2019;120(115277).
4. Holčapek M, Liebisch G, Ekroos K. Lipidomic analysis. *Anal Chem*. 2018;90(7):4249–57.
5. Bowden JA, Ulmer CZ, Jones CM, Koelmel JP, Yost RA. NIST lipidomics workflow questionnaire: an assessment of community-wide methodologies and perspectives. *Metabolomics*. 2018;14(5): 53.
6. Rampler E, Abiead YE, Schoeny H, Rusz M, Hildebrand F, Fitz V, et al. Recurrent topics in mass spectrometry-based metabolomics and lipidomics-standardization, coverage, and throughput. *Anal Chem*. 2021;93(1):519–45.
7. European Medicines Agency (EMA). Guideline on bioanalytical method validation. <https://www.ema.europa.eu/en/bioanalytical-method-validation>. Access date: Feb 17th 2021.
8. Food and Drug Administration (FDA). Bioanalytical method validation guidance for industry. <https://www.fda.gov/files/drugs/published/Bioanalytical-Method-Validation-Guidance-for-Industry.pdf>. Access date: Feb 17th 2021.
9. Lipidomics Standards Initiative (LSI). <https://lipidomics-standards-initiative.org/guidelines/lipid-species-quantification>. Access date: Feb 26th 2020.
10. Koivusalo M, Haimi P, Heikinheimo L, Kostianen R, Somerharju P. Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response. *J Lipid Res*. 2001;42(4):663–72.
11. Höring M, Ejsing CS, Hermansson M, Liebisch G. Quantification of cholesterol and cholesteryl ester by direct flow injection high-resolution Fourier transform mass spectrometry utilizing species-specific response factors. *Anal Chem*. 2019;91(5):3459–66.
12. Eisenhofer G, Fassnacht M. Steroid profiling for adrenocortical disorders: a pathway for omics-based diagnostics. *Clin Chem*. 2017;63(12):1787–9.
13. Lange M, Fedorova M. Evaluation of lipid quantification accuracy using HILIC and RPLC MS on the example of NIST® SRM® 1950 metabolites in human plasma. *Anal Bioanal Chem*. 2020;412(15):3573–84.
14. Schuhmann K, Moon H, Thomas H, Ackerman JM, Groessl M, Wagner N, et al. Quantitative fragmentation model for bottom-up shotgun lipidomics. *Anal Chem*. 2019;91(18):12085–93.

15. Thompson JW, Adams KJ, Adamski J, Asad Y, Borts D, Bowden JA, et al. International ring trial of a high resolution targeted metabolomics and lipidomics platform for serum and plasma analysis. *Anal Chem.* 2019;91(22):14407–16.
16. Bowden JA, Heckert A, Ulmer CZ, Jones CM, Koelmel JP, Abdullah L, et al. Harmonizing lipidomics: NIST interlaboratory comparison exercise for lipidomics using SRM 1950-metabolites in frozen human plasma. *J Lipid Res.* 2017;58(12):2275–88.
17. Burla B, Arita M, Arita M, Bendt AK, Cazenave-Gassiot A, Dennis EA, et al. MS-based lipidomics of human blood plasma: a community-initiated position paper to develop accepted guidelines. *J Lipid Res.* 2018;59(10):2001–17.
18. Timmerman P, Goodman J, Golob M, Haslberger T, Knutsson M, Nelson R, et al. European Bioanalysis Forum feedback on draft ICH M10 guideline on bioanalytical method validation during the Step 2b public consultation period. *Bioanalysis.* 2020;12(6):1–11.
19. Lowes S, Ackermann BL. AAPS and US FDA Crystal City VI workshop on bioanalytical method validation for biomarkers. *Bioanalysis.* 2016;8(3):163–7.
20. Triebel A, Burla B, Selvalatchmanan J, Oh J, Tan SH, Chan MY, et al. Shared reference materials harmonize lipidomics across MS-based detection platforms and laboratories. *J Lipid Res.* 2020;61(1):105–15.
21. Ulmer CZ, Ragland JM, Koelmel JP, Heckert A, Jones CM, Garrett TJ, et al. LipidQC: method validation tool for visual comparison to SRM 1950 using NIST Interlaboratory comparison exercise lipid consensus mean estimate values. *Anal Chem.* 2017;89(24):13069–73.
22. Cuadros-Rodríguez L, Bagur-González MG, Sánchez-Viñas M, González-Casado A, Gómez-Sáez AM. Principles of analytical calibration/quantification for the separation sciences. *J Chromatogr A.* 2007;1158(1–2):33–46.
23. Serna J, Garcia-Seisdedos D, Alcazar A, Lasuncion MA, Busto R, Pastor O. Quantitative lipidomic analysis of plasma and plasma lipoproteins using MALDI-TOF mass spectrometry. *Chem Phys Lipids.* 2015;189:7–18.
24. Folch J, Lees M, Sloane Stanley GH. A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem.* 1957;226(1):497–509.
25. Adams KJ, Pratt B, Bose N, Dubois LG, St John-Williams L, Perrott KM, et al. Skyline for small molecules: a unifying software package for quantitative metabolomics. *J Proteome Res.* 2020;19(4):1447–58.
26. Wang M, Wang C, Han X. 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.* 2017;36(6):693–714.
27. Poots AJ, Woodcock T. Statistical process control for data without inherent order. *BMC Med Inform Decis Mak.* 2012;12:86.
28. Koch J, Lackner K, Wohlfarter Y, Sailer S, Zschocke J, Werner ER, et al. Unequivocal mapping of molecular ether lipid species by LC-MS/MS in plasmalogen-deficient mice. *Anal Chem.* 2020;92(16):11268–76.
29. Wolrab D, Chocholoušková M, Jirásko R, Peterka O, Holčapek M. Validation of lipidomic analysis of human plasma and serum by supercritical fluid chromatography-mass spectrometry and hydrophilic interaction liquid chromatography-mass spectrometry. *Anal Bioanal Chem.* 2020;412(10):2375–88.
30. González O, Blanco ME, Iriarte G, Bartolomé L, Maguregui MI, Alonso RM. Bioanalytical chromatographic method validation according to current regulations, with a special focus on the non-well defined parameters limit of quantification, robustness and matrix effect. *J Chromatogr A.* 2014;1353:10–27.
31. Taylor DR, Ghataore L, Couchman L, Vincent RP, Whitelaw B, Lewis D, et al. A 13-steroid serum panel based on LC-MS/MS: use in detection of adrenocortical carcinoma. *Clin Chem.* 2017;63(12):1836–46.
32. Bancos I, Taylor AE, Chortis V, Sitch AJ, Jenkinson C, Davidge-Pitts CJ, et al. Urine steroid metabolomics for the differential diagnosis of adrenal incidentalomas in the EURINE-ACT study: a prospective test validation study. *Lancet Diabetes Endocrinol.* 2020;8(9):773–81.
33. Parker LA, Chilet-Rosell E, Hernández-Aguado I, Pastor-Valero M, Gea S, Lumberras B. Diagnostic biomarkers: are we moving from discovery to clinical application? *Clin Chem.* 2018;64(11):1657–67.
34. Simons K. How can omic science be improved? *Proteomics.* 2018;18(5–6):e1800039.
35. International Lipidomics Society (ILS). <https://lipidomicsociety.org/2020/09/30/elm-2020-and-lipidomics-forum-session-recordings/>. Accessed on Feb 26th 2020.
36. Honour JW. Development and validation of a quantitative assay based on tandem mass spectrometry. *Ann Clin Biochem.* 2011;48(Pt 2):97–111.
37. Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, Lewis MR, et al. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics.* 2018;14(6):72.
38. Cajka T, Smilowitz JT, Fiehn O. Validating quantitative untargeted lipidomics across nine liquid chromatography-high-resolution mass spectrometry platforms. *Anal Chem.* 2017;89(22):12360–8.
39. Herzog R, Schwudke D, Shevchenko A. LipidXplorer: software for quantitative shotgun lipidomics compatible with multiple mass spectrometry platforms. *Curr Protoc Bioinformatics.* 2013;43:14.2.1–2.30.
40. Hartler J, Trötz Müller M, Chittraju C, Spener F, Köfeler HC, Thallinger GG. Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics.* 2011;27(4):572–7.
41. Wang SY, Kuo CH, Tseng YJ. Batch Normalizer: a fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Anal Chem.* 2013;85(2):1037–46.
42. Koelmel JP, Cochran JA, Ulmer CZ, Levy AJ, Patterson RE, Olsen BC, et al. Software tool for internal standard based normalization of lipids, and effect of data-processing strategies on resulting values. *BMC Bioinformatics.* 2019;20(1):217.
43. Mohamed A, Molendijk J, Hill MM. Lipidr: a software tool for data mining and analysis of lipidomics datasets. *J Proteome Res.* 2020;19(7):2890–7.
44. Peng B, Kocpczynski D, Pratt BS, Ejsing CS, Burla B, Hermansson M, et al. LipidCreator workbench to probe the lipidomic landscape. *Nat Commun.* 2020;11(1):2057.
45. Schwudke D, Shevchenko A, Hoffmann N, Ahrends R. Lipidomics informatics for life-science. *J Biotechnol.* 2017;261:131–6.
46. Gardner MS, McWilliams LG, Jones JL, Kuklennyik Z, Pirkle JL, Barr JR. Simultaneous quantification of free cholesterol, cholesteryl esters, and triglycerides without ester hydrolysis by UHPLC separation and in-source collision induced dissociation coupled MS/MS. *J Am Soc Mass Spectrom.* 2017;28(11):2319–29.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.