Circulation: Genomic and Precision Medicine

PERSPECTIVE

Steps Toward Minimal Reporting Standards for Lipidomics Mass Spectrometry in Biomedical Research Publications

Valerie B. O'Donnell, PhD; Garret A. FitzGerald, MD; Robert C. Murphy, PhD; Gerhard Liebisch, PhD; Edward A. Dennis, PhD; Oswald Quehenberger, PhD; Shankar Subramaniam, PhD; Michael J.O. Wakelam, PhD

Lipids in blood and tissues can serve as markers of normal and pathophysiological function in humans and can even reflect functions in specific tissues and organs. Lipidomics describes the analysis of large numbers of lipids using mass spectrometry (MS). The proper implementation of these methods in a manner that ensures data quality requires care and rigorous manual checking. Issues of reproducibility and overall data quality in publications and guidelines for authors submitting research are well-developed for areas that include genetics/genomics, proteomics, and clinical trials. For example, the Human Proteome Organization has developed minimum information publication guidelines for proteomics (https://www.hupo.org/HUPO-Minimum-Information-Publication-Guidelines). However, apart from specialized lipid publications, such as the *Journal of Lipid Research*, which adopted the Lipid Metabolites and Pathways Strategy Consortium (LIPID MAPS) classification, nomenclature, and structural drawing formats in their guidelines,[1,2] there are few reporting guidelines in use for lipidomics data. This issue is particularly relevant to studies that are not focused on underpinning methodological approaches but instead cover broader issues of human health and disease. In many such articles, multiple analytical methods are applied, making it difficult to engage sufficient technical expertise to afford rigorous and comprehensive review.

We developed a short set of guidelines for lipidomics submissions that we hope will contribute to improving reproducibility and standards in published work (Table). This is a living document, expected to be expanded as the field evolves. It is not intended to serve as a definitive final set of guidelines. To support this sort of activity, the Lipidomics Standard Initiative was recently established to create guidelines for major lipidomic workflows.[3]

In the lipidomics field, different considerations apply to targeted and untargeted workflows, and it would be impossible to cover all of these in a short set of guidelines. However, we highlight some that we feel are worth special mention below.

GUIDELINES FOR MANUSCRIPT REVIEW

Given the complexity of this field, it is desirable that at least 1 individual with analytical domain expertise be included among the reviewers.

Authentic Standards

Primary standards are generally highly defined in terms of stereochemistry, since they are chemically synthesized and thus highly purified single isomers or enantiomers. It is critical to confirm that the retention time of lipids in biological samples matches that for synthetic standards, as multiple isomers of lipids may overlap during elution. A separate issue relates to identification of complex lipids using MS without fragmentation, which does not define fatty-acyl composition. Here, appropriate shorthand annotation should be applied. Guidelines have already

Key Words: biomarkers ■ genomics ■ lipids ■ lipoproteins ■ publications

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

Correspondence to: Valerie B. O'Donnell, PhD, Systems Immunity Research Institute, Cardiff University, Cardiff CF14 4XN, United Kingdom. Email o-donnellvb@cardiff.ac.uk

For Sources of Funding and Disclosures, see page 741.

Circulation: Genomic and Precision Medicine is available at www.ahajournals.org/journal/circgen

Table. Guidelines

Part A	This section is designed to help reviewers and editors
1	When MS lipidomics are included in a study, it would be optimal that at least 1 reviewer be an analytical biochemist or chemist with expertise in lipid identification and quantification by MS
2	Figures that contain TIC MS chromatograms (eg, for liquid chromatography tandem MS) or mass spectral data can be requested for the reviewers as high-resolution images during review
3	Chromatograms and MS/MS spectra for exemplar/representative lipids that have been measured and quantified using targeted methods should be provided with minimal processing and the processing pipeline used described. Representative chromatograms can be included in supplementary data sections for confirmatory viewing
Part B	This section is designed to help authors include in Materials or Data Supplement sufficient data to allow readers of the manuscript to understand the experimental details supporting the reported results. Additional details can be found in the references[1,2,4]
1	Lipidomics identification of lipids by MS (with/without fragmentation)
1.1	Experimental parameters
	Shotgun (with references to exact details) or chromatographic
	Ion trap, tandem quadrupole, orbitrap, quadrupole–time of flight (manufacturer and model)
	Resolution of mass analyzer (if high resolution, provide the resolution used, FWHM)
	Mass accuracy, state the Δppm value applied for putative identification of lipids
	Details of chromatographic separation (column, flow rate, mobile phases, gradient, injection volume, and post-column additions) and manufacturer identification
	Oven temperature and autosampler temperature
	Lipid standards (manufacturer) used in chromatographic method to identify lipid class and species
1.2	Ionization used
	Positive or negative ions
	Settings for ionization (ESI voltage and EI ionization energy)
	MALDI matrix
1.3	Tandem MS
	Collision gas and pressure with point at which it is measured
	Collision voltage
	Optimization details
1.4	Software used to assist in lipid identification and mass spectrometric identification level
	Level of lipid identification should comply with recommendations by Liebisch et al[4]
	If automation was used for checking peak quality, the software and version should be stated
	For studies involving measures of large numbers of lipid species in cohorts or other relevant biological samples to generate big data sets, data should be deposited at either LIPID MAPS (http://lipidmaps.org/resources/data/index.php) directly (or via the Metabolomics Workbench) or MetaboLights or other suitable databases. The processed data should indicate amounts (relative or quantified) of each lipid species measured in individual samples
2	Targeted quantitative analysis by MS (precursor-product ion measurements)
2.1	Reference standards (supplier)
2.2	Internal standards (supplier)
	Homolog or stable isotope labeled
	Position of stable isotopes (correct chemical nomenclature)
	Isotopic purity of stable isotopes and atom percentage
2.3	Naming of lipids to be based on the level of identification achieved, rather than the name of standard used
2.4	Precursor-product ion choice should be based on the following
	Significant product ion (typically carbon chain fragmentation)
	Avoid loss of small neutral species from molecules such as H_2O or CO_2 (or combination) as choice for an MRM transition or structure confirmation
	Reference to standard tandem mass spectrum (or provide in the Data Supplement)
2.5	Chromatographic identity
	Identical retention time as standard (reproducibility evaluation)
	Evaluation of reproducibility of retention time
	Coelution with stable isotope internal standard, if available (deuterium labeling often results in slight preelution due to isotope effect)

(*Continued*)

Table. Continued

	Monitor another MRM ion pair as additional confirmation of lipid identity where feasible
	Representative chromatogram (MRM trace recordings) of actual samples presented as a Data Supplement
	Statement of duty cycle (total time interval to complete all MRM pairs)
	Collision energy settings
2.6	Absolute minimum of 6–10 points across a peak to determine (integrate) a peak
	Peak area and peak height to be provided
	Chromatographic time window sufficiently wide to assess noise level and the presence of closely related isomers
	Minimal smoothing of data and if used, examples of the impact of smoothing on the original raw data provided with justification
	Basis for determination of LOD and LOQ
	No thresholding of data
	Use of chiral chromatography to establish, even if using enantiomerically pure standards
2.7	Validation method described (as conducted)
	Detection method (analog or pulse counting) used
	Performed in matrix of the experiments (plasma, cell media, and urine)
	Calibration curve details (at least 6 points covering the range of quantitation)
	LOQ to be S/N at least 5–10 and stated
	Percentage error of theoretical <15% (validation) or 20% for MS omics
	LOD to be S/N at least 3 and stated
	Do not assign quantitative value below LOQ unless clearly stated
	Verification of LOD by spike of reference standard (at LOD) to matrix
	Percentage error of theoretical (validation experiment)
	Periodic blank samples placed in batch series to verify little ghosting
2.8	Example of raw chromatographic data supporting analysis at LOQ or LOD (Data Supplement) that describes processing pipeline
3	Identification and structural characterization of new lipids
	A defined level of chemical characterization is required to claim a new lipid structure. It is recognized that MS is far more sensitive than NMR; thus in biological samples, full structural characterization may not be possible to obtain for a biological lipid only present in small amounts in tissue samples. This is a common situation, but it should not preclude biological studies. Thus, a putative or proposed structure can be claimed, pending confirmation using NMR, once enantiomerically pure synthetic lipids become available
3.1	Tests recommended to propose a structure include
	LC/MS(/MS), including reverse phase, straight/chiral phase LC
	MS and MSn, using a high-resolution instrument
	Confirmation of MS data using deuterated analogs generated using the same reaction/biosynthetic pathway
	GC/MS with derivatization for functional group confirmation
	Stability testing using acid, base, temperature, etc
	New structures should be deposited in relevant database, such as the LIPID MAPS database: http://lipidmaps.org/resources/data/index.php

EI indicates electron impact; ESI, electrospray; FWHM, full width half maximum; GC, gas chromatography; LIPID MAPS, Lipid Metabolites and Pathways Strategy Consortium; LOD, limit of detection; LOQ, limit of quantitation; MALDI, matrix-assisted laser desorption ionization; MRM, multiple reaction monitoring; MS, mass spectrometry; NMR, nuclear magnetic resonance; S/N, signal:noise; and TIC, total ion count.

been developed concerning the type of MS used and the valid level of identification possible.[4]

Nomenclature

One of the greatest sources of confusion in lipidomics research is the nomenclature of lipids. The wide use of disparate names for lipids and the lack of a standard naming system makes comparative analysis of lipidomics data across different studies generated by different equipment and investigators difficult. LIPID MAPS1,2 and the National Metabolomics Data Repository have tried to address this problem through the development of a Reference Set of Metabolite Names—a highly curated analytical chemistry-centric database of common names for lipids and other metabolite structures and isobaric species.[5] Reference Set of Metabolite Names has been linked to the LIPID MAPS classification system enabling data sharing and meta-analysis. Easy tools for mapping and cross-referencing other lipid names have also been made available on the Metabolomics Workbench.

Peak Quality

The fundamental analytical parameter is chromatographic resolution. Resolution may be enhanced by MS resolution of distinctive ion masses. Good peak shape is desirable and required when many analytes are recorded in a single run, but chromatographic and MS resolution is the criterion.

A pragmatic approach to assessing peak quality should be applied for targeted assays, especially since many laboratories are implementing new multiplex targeted methods that can quantify literally hundreds of lipids in a single run. A useful rule of thumb widely used in analytical chemistry would be that chromatographic peaks, generated from targeted multiple reaction monitoring experiments, which are usually gaussian in shape, should be at least 3× signal:noise for a limit of detection, rising to 5:1 or even 10:1 for defining the limit of quantitation. Care must be taken to be sure that any chromatogram used for analysis should have sufficient data points across the peaks analyzed (eg, preferably 6–10 data points across a peak) to identify its central retention time and peak shape enabling proper compound identification and quantification.

Determination of limit of detection and limit of quantitation, as well as the number of data points across a peak, should always be performed using raw data, with no smoothing applied. Representative raw data chromatograms (without any smoothing or other alterations) should be included for reviewers to inspect, and these may include expansions (blowups) of typical peaks used for signal:noise analysis and limit of detection/limit of quantitation determination. These should also be provided as published supplements or be deposited in suitable databases for readers.

Minimal smoothing may help to determine the apex of the peak and its shape relative to standards more accurately to assign the retention time, identify the compound, and quantify based on the area under the peak. Smoothing is primarily of use for cosmetically improving peaks, and since it enhances signal:noise, it should not be applied until data are tested for peak quality.[6] A quick manual inspection of all raw data peaks will avoid computational errors that can be incurred if the automated determination of peak quality is solely applied using vendors' software.

A separate use case is untargeted lipidomics, where lipids are detected based on high-resolution MS without fragmentation. Here, the identification of lipids is based only on accurate mass and retention time. Including spectral data in the acquisition process provides additional information for structural identification and validation, such as fatty acid composition of complex lipids. Here, automated methods are becoming more generally applied, and their use is likely to increase in the future. In the attempt to automate the identification, integration, and interpretation of MS data, peak analysis becomes critical. But with software-guided analysis, there is often the need to make some compromises to include all peaks of interest in the specific analysis. However, it is always important to at least spot-check peaks across the resulting chromatogram to verify the software-generated assignment of the relevant lipids of interest to the investigator, as well as the peak (and its shape) used for integration. One useful tool that may help with this is Lipid Data Analyzer (http://genome.tugraz.at/lda2/lda_download.shtml).[7] This can be downloaded free of charge or used online. The original version is optimized for the annotation of phospholipids and glycerolipids and is platform independent.

Quantitation of Lipids

Methods to carry out quantitative analysis by MS include stable isotope dilution and targeting specific lipids. This is the approach typically used for those lipids present at low abundance and for which reference standard material and isotope-labeled internal standards are available. Alternative methods for those lipids present as multiple molecular species utilize an unnatural species (homolog or isotope labeled) as an internal standard and a specific reference lipid to generate several standard curves. The mass spectrometer can be operated in a nontargeted or targeted (tandem precursor/product monitoring) mode, and each approach has advantages and disadvantages. Rather than absolute quantitation, relative quantitation can be used in a controlled experimental series that does not require standard curve generation and is accurate for fold changes. The use of the same m/z value for precursor-to-product ion (multiple reaction monitoring) experiments is discouraged since it does not allow lipids with the same m/z to be discriminated if they coelute; thus for quantitative methods, a product ion that is unique to the lipid of interest should be chosen, if possible. If the same m/z value is required, then in such cases, chromatography should provide sufficient specificity.

Structural Analysis of Lipids

We suggest a pragmatic approach to the description of new lipid mediators, reporting putative structures based on the information available. As an example, with a newly discovered lipid, part of the structure (eg, carbon chain length, number of rings/double bonds, position of oxygenation, and nature of functional groups) may be known, but details that include stereochemistry or double bond isomers may not yet be elucidated. Here, the known biological information such as enzymatic and cellular source and putative bioactivity can be placed in the public domain to encourage others to follow the work and expand it, including through more detailed structural characterization. Supporting publication of structures where the full stereochemistry may not yet be known ensures that ongoing biological studies can proceed but

has the advantage of enabling updates to be provided as additional information, such as nuclear magnetic resonance and chiral identification become available. Where reference spectra are included, the major product ions should match the synthetic standard.

Data Deposition

In relation to untargeted workflows, large datasets should be deposited at recognized repositories for future data mining and integration into systems biology, such as the Metabolomics Workbench (https://www.metabolomicsworkbench.org/repository/index.php), which has a portal through LIPID MAPS (http://lipidmaps.org/resources/data/index.php), or MetaboLights (https://www.ebi.ac.uk/metabolights/presubmit). We recommend the LIPID MAPS nomenclature and Reference Set of Metabolite Names be used as a common standard approach, either the shorthand or fully annotated nomenclature as appropriate.[1,2]

Last, we highlight that there are established international guidelines for validation of analytical procedures, from the World Health Organization, Food and Drug Administration, and European Medicines Agency. These are designed for drug or toxicology applications that require a higher level of validation than research assays in general, but they provide excellent information relating to accepted approaches for validation of quantitative methods in the field, as follows: https://www.who.int/medicines/areas/quality_safety/quality_assurance/28092018Guideline_Validation_AnalyticalMethodValidation-Appendix4_QAS16-671.pdf; https://www.fda.gov/regulatory-information/search-fda-guidance-documents/analytical-procedures-and-methods-validation-drugs-and-biologics; https://www.ema.europa.eu/en/ich-q2-r1-validation-analytical-procedures-text-methodology.

ARTICLE INFORMATION

Affiliations

Systems Immunity Research Institute, School of Medicine, Cardiff University, United Kingdom (V.B.O.). Institute for Translational Medicine and Therapeutics, Smilow Center for Translational Research, University of Pennsylvania, Philadelphia (G.A.F.). Department of Pharmacology, University of Colorado Denver, Aurora (R.C.M.). Institute of Clinical Chemistry and Laboratory Medicine, University of Regensburg, Germany (G.L.). Department of Chemistry and Biochemistry (E.A.D.), Department of Pharmacology (O.Q.), and Department of Bioengineering (S.S.), University of California, San Diego. Babraham Institute, Babraham Research Campus, Cambridge, United Kingdom (M.J.O.W.).

Sources of Funding

Funding from the Wellcome Trust for LIPID MAPS is gratefully acknowledged (203014/Z/16/Z). LIPID MAPS receives sponsorship from Avanti Polar Lipids, Cayman Chemical, and Merck. This study was supported by National Institutes of Health grants UL1 TR001878 (Dr FitzGerald), GM20501 (Dr Dennis), U54 GM069338 (Drs Dennis, Murphy, Subramaniam, and Quehenberger), U01 DK097430 Metabolomics Workbench (Dr Subramaniam), and U2C DK119886 Metabolomics Workbench (Dr Subramaniam). Dr O'Donnell is a Royal Society Wolfson Merit Awardee.

Disclosures

None.

REFERENCES

1. Fahy E, Subramaniam S, Brown HA, Glass CK, Merrill AH Jr, Murphy RC, Raetz CR, Russell DW, Seyama Y, Shaw W, et al. A comprehensive classification system for lipids. *J Lipid Res.* 2005;46:839–861. doi: 10.1194/jlr.E400004-JLR200
2. Fahy E, Subramaniam S, Murphy RC, Nishijima M, Raetz CR, Shimizu T, Spener F, van Meer G, Wakelam MJ, Dennis EA. Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res.* 2009;50 suppl:S9–14. doi: 10.1194/jlr.R800095-JLR200
3. Liebisch G, Ahrends R, Arita M, Arita M, Bowden JA, Ejsing CS, Griffiths WJ, Holčapek M, Köfeler H, Mitchell TW, et al. Lipidomics needs more standardization. *Nat Metab.* 2019;1:745–747.
4. Liebisch G, Vizcaíno JA, Köfeler H, Trötzmüller M, Griffiths WJ, Schmitz G, Spener F, Wakelam MJ. Shorthand notation for lipid structures derived from mass spectrometry. *J Lipid Res.* 2013;54:1523–1530. doi: 10.1194/jlr.M033506
5. Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK, Merrill AH Jr, Murphy RC, Raetz CR, Russell DW, et al. LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* 2007;35:D527–D532. doi: 10.1093/nar/gkl838
6. Enke CG, Nieman TA. Signal-to-noise ratio enhancement by least-squares polynomial smoothing. *Anal Chem.* 1976;48:705A–712A.
7. Hartler J, Triebl A, Ziegl A, Trötzmüller M, Rechberger GN, Zeleznik OA, Zierler KA, Torta F, Cazenave-Gassiot A, Wenk MR, et al. Deciphering lipid structures based on platform-independent decision rules. *Nat Methods.* 2017;14:1171–1174. doi: 10.1038/nmeth.4470