A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP15465.

Categorizing Concentration Confidence: A Framework for Reporting Concentration Measures from Mass Spectrometry-Based Assays

Lauren M. Petrick,[1,2] David Achaintre,[1,2] Amith Maroli,[1,2] Julio Landero,[1,2] Priyanthi S. Dessanayake,[1,2] Susan L. Teitelbaum,[1,2] Mary S. Wolff,[1,2] Manish Arora,[1,2] Robert O. Wright,[1,2] and Syam S. Andra[1,2]

[1]Department of Environmental Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA
[2]The Institute for Exposomics Research, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Background: Innovation in mass spectrometry-based methods to both quantify and perform discovery has blurred the lines between targeted and untargeted assays of biospecimens. Continuous data—concentrations or intensity values generated from both methods—can be used in statistical analysis to determine associations with health outcomes, but concentration values are needed to compare measurements from one study to another to inform policy making decisions and to develop clinically relevant thresholds. As a single solution for discovery and quantitation, new hybrid-type assays derive concentration values for chemicals or metabolites but with varying degrees of uncertainty that may be greater than traditional quantitative assays. There is no current single standard to guide reporting bioassay concentrations or their uncertainty in concentration values from hybrid assays. Even when measures are robust, obtained with high scientific rigor, and provide valuable data toward risk assessment, unknown uncertainty can lead to bias in interpretation of reported data or omission of reported data that does not meet the strict criteria for absolute quantitation.

Objective: The objective of this commentary is to articulate a scheme that enables investigators across bioanalytical fields to easily report analyte measurement assurance on the same scale from quantitative, untargeted, or hybrid assays that include a range of concentration confidences.

Discussion: We propose a simple scheme to report concentrations for targeted and untargeted analytes. Level 1 is a confirmed concentration following established tolerances in a fully quantitative assay while level 5 is a tentative intensity from a typical untargeted assay. This framework enables easy communication of uncertainty in concentration measurements to aid cross-validation, meta-analysis, and extrapolation across studies. It will facilitate interpretation while supporting analytical advancement and allow clear and concise measurement reporting across a broad range of confidences. https://doi.org/10.1289/EHP15465

Introduction

Uncertainty in measures affects the direction and size of estimates. Uncertainty can lead to epidemiological reports with conflicting results, even in well-designed studies designed to test causal hypotheses in humans. In addition, uncertainty and then extrapolation among studies can affect the accuracy of dose–response relationships in toxicology or drug efficacy. While reporting uncertainty in quantitative analysis uses established standards, developing new analytical methods requires new tools to report the uncertainty associated with those methods. For example, the Biomonitoring, Environmnetal Epidemiology, and Short-lived Chemicals (BEES-C) instrument evaluates the quality of biomonitoring data on short-lived chemicals and helps users better understand the limitations.[1] In addition, annotation confidence criteria help users understand the specificity of measurements in untargeted analysis for biological and exposure interpretation.[2] However, while new analytical advances combine untargeted and quantitative analysis using high-resolution mass spectrometry (HRMS) and low resolution mass spectrometry (QqQ) techniques, there is currently no tool to report the uncertainty of measures from these hybrid workflows easily.

Researchers combine the strengths and weaknesses of untargeted and targeted analyses to develop hybrid strategies to quantify exogenous chemicals and endogenous metabolites, while simultaneously measuring additional metabolites and chemicals with less precision and accuracy (Figure 1).[3–6] A traditional discovery workflow for an untargeted assay includes analysis of samples using HRMS. Data are collected on hundreds to thousands of analytes, and relative peak intensities across two or more groups (e.g., cases and controls, exposed and not exposed) are compared using advanced statistics. In contrast, quantitative assays with targeted analysis focus on the precise measurement of one to tens of *a priori* determined chemicals to obtain absolute concentrations using strict criteria to minimize uncertainty. These assays provide higher throughput and lower costs than an untargeted assay that employs HRMS techniques. However, there is increased interest in performing quantitation and discovery at the same time (i.e., a hybrid approach) because of limited sample availability, the cost and time to perform multiple assays, and the advancement of HRMS instrumentation and quantitative software.[7] Newly developed workflows include simultaneous quantitation and discovery (SQUAD) with HRMS technology: It combines an *a priori* set of analytes with authentic chemical standards and calibration curves in a targeted workflow and then remines the data for new analytes in an untargeted workflow.[8] There is a label-free hybrid approach combining HRMS tandem mass spectrometry (MS/MS) data collection across the metabolome coupled with sequential dilution and quantitation using a QqQ.[9] There is a pseudo-targeted method that uses HRMS to generate a targeted list of multiple reaction monitoring (MRM) ion pairs, which are then measured in samples on a QqQ to increase sensitivity.[10,11] These advancements in quantitation using untargeted methodologies accelerates science with valuable data, but they also bring ambiguity in error associated with the reported concentrations that are not easy to convey clearly or concisely.

Here, we propose a scheme to report confidence in concentration measures that do not meet the accuracy and uncertainty criteria

Address correspondence to Lauren M. Petrick, 1 Gustave Levy Place, Box 1057, New York, NY 10029 USA. Email: lauren.petrick@mssm.edu

M.A. is the founder and officer of Linus Biotechnology, Inc., a start-up company of the Mount Sinai Health System. His conflicts of interest are managed by the Mount Sinai Health System in accordance with institutional policies. The company had no role in the work reported in this manuscript. All other authors declare they have no conflicts of interest related to this work to disclose.

Conclusions and opinions are those of the individual authors and do not necessarily reflect the policies or views of EHP Publishing or the National Institute of Environmental Health Sciences.

Received 29 May 2024; Revised 27 January 2025; Accepted 20 March 2025; Published 12 May 2025.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

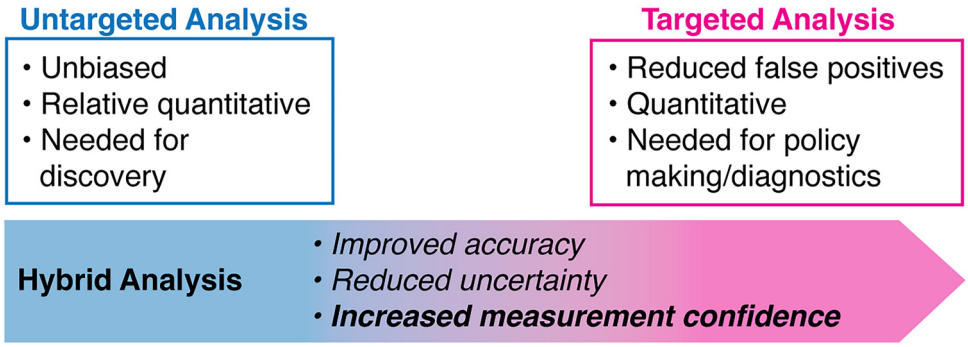

Figure 1. Hybrid analysis combines the advantages of untargeted and targeted analysis for risk assessment.

expected from an absolute quantitative assay but are still robust measures valid to help address policy and risk. Our scheme enables investigators to compare across assays more easily by binning reporting measures. It is not intended to replace established protocols to report absolute quantitation or untargeted analysis. It can complement these existing protocols and fill gaps in reporting for a range of assays not classically described by either technique in a single inclusive reporting scheme. For example, it is common for quantitative and untargeted assays to use quality control (QC) samples to calculate interbatch and intrabatch variability.[6,12] Therefore, using these metrics as criteria for a congruent scale across the range of assays is straightforward. Through this framework, we hope to facilitate transparency, improve data sharing, and initiate discussions with epidemiologists and risk assessment researchers for more robust and versatile use of these chemical measurements.

Concentration Confidence Scheme

In Table 1, we propose a concentration confidence scheme to report confidence in analyte concentrations generated from nontraditional targeted and untargeted assays (e.g., hybrid assays) that is also applicable to traditional targeted and untargeted workflows. To determine the confidence level for reporting a given analyte, we suggest that a user begins at level 1 and work down to level 5 until the analyte meets all method criteria and metric thresholds in that level. This should be repeated for each analyte in the study because the measurement confidence varies across the different analytes, especially in hybrid assays. Reporting the concentration confidence as level 1, 2, 3, 4, or 5 for each analyte should be included in the manuscript Methods section or provided in the supplementary tables along with additional method validation metrics such as limit of detection (LOD), detection frequencies, coefficient of variation (CV), and etc. Further, we encourage reporting the analyte concentration confidence into shared data repositories, such as the Human Health Exposure Analysis Resource (HHEAR) Data Center (https://hheardatacenter.mssm.edu/) and Metabolomics Workbench (https://www.metabolomicsworkbench.org), to enable researchers to include metadata reporting with their individual analyte measures. This framework also accommodates variability in calibration practices, such as the use of surrogate vs. matrix-matched standards,[13] and addresses quantification challenges across both xenobiotics and endogenous chemicals.[14] It enhances harmonizability for targeted, untargeted, and hybrid analyses by providing guidance on selecting confidence levels applicable for all workflows.

Level 1, confirmed. The criteria for level 1 is the highest confidence reporting and is based on extensive literature regarding absolute quantitation.[15] To achieve confidence level 1, every batch uses isotope dilution with a labeled version of the analytes of interest (i.e., internal standard calibration curve) and a matrix-matched (recommended) or non–matrix-matched calibration curve (performing within 5% deviation of matrix-matched slope calibration curve of at least six nonzero levels). Calibration curves must have $r^2 > 0.95$ and cover >95% of the range of study samples (100% is ideal). While an analyte-specific sample preparation protocol is suggested, it is not required as long as thresholds for reporting concentration accuracy of reference standards are met [$\pm 20\%$ relative standard deviation (RSD)], and there is repeatability (intrabatch variability $\leq 15\%$) and reproducibility (interbatch variability $\leq 15\%$) of QC results. If any of the method criteria or metric thresholds are not met, the analyte does not meet the requirements to report a level 1 and a less confident level must be reported. Confirmed quantification applies when authentic reference standards are analyzed concurrently with samples, matching exact mass, isotope pattern, retention time, and MS/MS spectrum.[16]

Level 2, strong. Level 2 confidence includes method criteria such as isotope dilution using a labeled analog from the chemical class of the analyte of interest (i.e., surrogate calibration curve) and/or a non–matrix-matched calibration curve with six or more levels. While an analyte-specific sample preparation protocol is suggested, it is not required as long as the metric thresholds for reporting concentration accuracy ($\pm 30\%$ RSD) and intrabatch variability ($\leq 15\%$) of the reference standards are met for every batch, and interbatch variability is $\leq 30\%$. This level includes analytes for which no externally validated reference standards are available. Therefore, even when metric thresholds meet those for level 1 [i.e., accuracy is $\pm 20\%$ RSD, uncertainty ($\leq 15\%$)], this is the highest confidence level possible when analyte-specific labeled standards are not used.

Level 3, good. Analysis is performed using the method criteria for level 1 or level 2 (analyte-specific standards or surrogate, matrix-matched curve or non–matrix matched with six or more levels) but does not achieve the metric threshold criteria for uncertainty for level 2. Accuracy should be $\pm 40\%$ RSD, and intrabatch and interbatch variability should be $\leq 30\%$ and $\leq 50\%$, respectively. If any of these metric thresholds are not met, a less confident level must be reported.

Level 4, fair. Analysis is performed using an external calibration curve without isotope dilution and a limited number of calibration levels ($n = 1$–5). Reference standards are not required, although uncertainty of the analyte should be reported even if there is no threshold. There is no threshold for intrabatch or interbatch variability at this level. Therefore, this is the most confident level that can be reported for an analyte that uses only one to five calibration curve levels (analyte-specific standards or surrogate, matrix matched curve or non–matrix matched) in the method. This is also the confidence level to report when at least one of the metric thresholds is not met for level 3 (i.e., accuracy $> \pm 40\%$

Table 1. Concentration confidence scheme. Proposed reportable quantification confidence levels for analytes based on thresholds of accuracy and uncertainty obtained from quantitative, hybrid, and untargeted assays.

Quantification confidence level	Method criteria			Threshold metrics			Quantification confidence indicator
	Native standards	Labeled internal standards	Calibration curve	Accuracy[a]	Intrabatch variability[b]	Interbatch variability[b]	
1	Analyte-specific	Analyte-specific	Multipoint (≥6 levels), internal	± Up to 20% RSD	≤15%	≤15%	Confirmed
2	Analyte-specific	Analyte-specific, class-specific, or similar physicochemical properties	Multipoint (≥6 levels), internal or surrogate	± Up to 30% RSD	≤15%	≤30%	Strong
3	Analyte-specific	Analyte-specific, class-specific, or similar physicochemical properties	Multipoint (≥6 levels), internal or surrogate	± Up to 40% RSD	≤30%	≤50%	Good
4	Nonspecific	None	Single or multipoint with 1–5 levels, internal or surrogate	± More than 40% RSD, or none	No threshold	No threshold	Fair
5	None	Not used for quantitation	None	None	Not applicable[c]	Not applicable[c]	Tentative

Note: RSD, relative standard deviation.
[a]Accuracy for level 1 is calculated from standard reference materials (SRM), which can be National Institute of Standards and Technology (NIST) samples, proficiency testing materials, or well-characterized pools used by multiple labs. For levels 2–4, reference material can also be spiked in the sample matrix. Quality controls (QCs) are injected into every analytical run of a project. Analyst confidence in the SRM results can influence the level selection.
[b]Interbatch and intrabatch variability (i.e., uncertainty) is calculated from any quality QC sample that is measured repeatedly throughout the analysis.
[c]Intrabatch variability <30% is typical but not required. After normalization and batch correction, interbatch variability is assessed and reported for selected metabolites or chemicals.

RSD and intrabatch and interbatch variability >30% and >50%, respectively).

Level 5, tentative. Criteria for the lowest confidence level are reserved for analytic measures that do not meet any of the requirements for quantitation. Intrabatch and interbatch variability should be reported. Robust data analysis allows intensity measures (not concentrations) to be used as continuous measures in statistical analysis. Concentration units are not reported with these measures. This is the most confident level that can be reported for an analyte in the absence of any calibration curve or the absence of isotope labeled standards (analyte-specific or surrogate) used in the method.

We recognize that an analyte may be classified at a lower level on the scheme due to analytical restrictions but may be practically equivalent to a higher level. For example, labeled standards may not be commercially available, or an externally validated reference sample for assessing accuracy may not be available to reach level 1. It is common practice in complex lipidomics experiments that seek to quantify hundreds of lipid species in the same run[17] to use nonphysiologically occurring lipid species[18] or to use an external standard curve (constructed from a common lipid) to quantify all individual lipid species in a particular class.[19] Additionally, nonconventional matrices such as hair, teeth, dried blood spots, or nails may not have proper QC materials, so surrogates must be used.[20] While these criteria may require reporting a "lower" level of quantitation confidence according to the current scheme and expert experience, concentrations measured in these biological samples are still valuable and valid.

Selected Examples

Here, we provide examples of typical and novel analysis workflows used in biomedical and environmental analysis and the respective concentration confidence levels that will be reported using the proposed level scheme (Table 1).

Case 1. Urinary metabolites of phthalates[21] and phenols[22] were measured according to CDC-published methods.[23,24] This involved sample-specific preparation using enzyme deconjugation and sample cleanup after solid-phase extraction. Labeled standards for each analyte of interest were utilized for isotope dilution with a seven-level calibration curve. Analytes were quantified using a QqQ mass spectrometer with MRM. Batch-wise RSDs of QCs during analysis of the study specimens were <20% for target analytes present in reference or fortified materials above the LOD. Intrabatch precision for QC analytes above the LOD had a CV of <10%, while interbatch precision remained below 15%. Recovery rates for batch QCs ranged from 80% to 120%. Quality assurance (QA) and QC protocols include those described by the Mount Sinai HHEAR Targeted Lab,[25,26] and these assays meet the concentration confidence Level 1 criteria for all analytes.

Case 2. A dilute-and-shoot assay was developed to measure six carotenoids by QqQ mass spectrometry (lutein, β-cryptoxanthin, zeaxanthin, α-carotene, β-carotene, and lycopene).[27] To reduce labor time and prevent potential metabolite degradation, plasma samples were deproteinized with methanol and centrifuged, followed by injection. Seven level calibration curves for each analyte were generated with a single internal standard, α-tocopheryl acetate. Among the six carotenoids, lutein and zeaxanthin coeluted and were measured as a mixture. Furthermore, lycopene stereoisomers were summed to obtain the final concentrations. Linearity for all carotenoids was >0.99. Intrabatch and interbatch precision values were evaluated with National Institute of Standards and Technology (NIST) samples (ranging from 6.2% for β-carotene to 12.4% for β-cryptoxanthin for interbatch CV). A matrix effect from 105% to 115% was observed. Accuracy at low concentrations of NIST samples was <20% for all metabolites, except for α-carotene, which

was 30%. Accuracy at high concentrations was 35.9% for lycopene and 108% for β-cryptoxanthin. α-Carotene and β-carotene meet the criteria for confidence reporting level 2, lycopene meets the criteria for confidence reporting level 3, whereas β-cryptoxanthin and a mix of lutein and zeaxanthin meet the criteria for level 4.

Case 3. Using an untargeted metabolomics workflow with HRMS, a hybrid assay was developed to quantify per- and polyfluoroalkyl substances (PFAS) including perfluorooctanesulfonic acid (PFOS) and perfluorooctanoic acid (PFOA) in plasma (example based on Goodrich et al.[28]). A six-level matrix-matched calibration curve using isotope dilution and labeled PFOS and PFOA was used. All endogenous metabolites were measured in the absence of labeled standards or calibration curves. Plasma metabolites, PFOS, and PFOA in the study samples and NIST SRM1957 and 1958 were extracted using a simple solvent protein crash. PFOS and PFOA were measured using the MS1 ion within a *m/z* of 5 ppm from the standards. Accuracy and uncertainty for PFOS and PFOA were within 10% and <15%, respectively. All other endogenous metabolites quality control was unreported. PFOS and PFOA meet the criteria for confidence reporting level 1, while all other metabolites and chemicals measured in the assay meet the criteria for confidence reporting level 5.

Case 4. A hybrid assay named globally optimized targeted mass spectrometry (GOT-MS) was developed to expand and optimize detection capabilities and improve quantification, notably amino acids.[29] A global search of precursors and products ions by liquid chromatography-triple quadrupole (LC-QqQ) mass spectrometry was followed by selecting MRM transitions for identification and quantitative analysis. Five samples were spiked with labeled U-^{13}C-^{15}N-amino acids and diluted to generate five-level calibration curves for amino acids without using native standards. Isoleucine and Leucine coeluted and were measured as a mixture. Intrabatch variability was $<3.0 \pm 2.2\%$ without amino acid labeled standard normalization and $<4.1 \pm 3.4\%$ with amino acid labeled normalization. Interbatch variability of amino acids decreased from $8.3 \pm 3.4\%$ to $4.8 \pm 4.0\%$ after normalization with labeled amino acids. The linearity (R^2) of amino acids was $0.82 \pm 0.26\%$ for GOT-MS. Amino acids meet the criteria for confidence reporting level 4, whereas all other metabolites and chemicals measured meet the criteria for confidence reporting level 5.

Discussion

The availability of analytic methods that can increase the biomarker information obtained from nonrenewable biological samples using HRMS is an important advancement in environmental health research. However, the authors' experience is that there is skepticism about the comparability of biomarker concentrations obtained through HRMS and the traditional targeted sample analyses, and this has been a roadblock to accepting study results that report HRMS-derived concentrations. In particular, establishing exposure–outcome associations requires results from multiple studies in varied populations where exposure measurements are considered comparable. The proposed concentration confidence levels describe the uncertainty around concentration measures from untargeted, hybrid, and quantitative analysis on the same scale. Therefore, it can provide information for the evaluation of whether the results of a study that uses HRMS-derived concentrations can be used in combination with other studies to establish exposure–outcome associations. For example, an analyte reported from different studies with measurement confidence level 1 can easily be combined for meta-analysis and cross-study evaluation. Within the HHEAR program, measurements that reach levels 1 and 2 are considered harmonizable based on the published or publicly available proficiency testing materials' reference values and tolerance ranges that indicate an accuracy of ≥70% recoveries and a precision of ≤30% RSD as an acceptable benchmark (CHEAR/HHEAR Targeted Lab Analysis and Quality Assurance Working Group, personal communication). While it is likely that additional levels can also be considered, when and how studies from multiple confidence levels can be aggregated requires further evaluation, which will influence the development and modification of future schemes. For example, an analyte reported from different studies with measurement confidence level 1 can be easily combined for meta-analysis and cross-study evaluation.

In addition to reporting concentration confidence for each analyte according to Table 1, we propose a way to report assay performance using accuracy criteria as well as intrabatch and interbatch variability of the analytes. Therefore, this scheme can provide a reporting tool for cross-validation of untargeted and hybrid assays in the absence of proficiency testing and round-robin validation schemes, which are common to quantitative analysis. By using these confidence levels as targets when optimizing new assays, this framework can influence study design and analytical method development.

There are particular challenges to quantitation of some exogenous and endogenous chemicals in complex matrices like plasma and urine. As a result, greater confidence level in the concentration confidence scheme is achieved through approaches that account for preanalytical variability and matrix effects. For example, a matrix-matched calibration curve is used as the same biological matrix as the samples to account for potential degradation during sample preparation and analysis.[30] This approach ensures that calibration standards and samples are affected similarly by matrix components.[31] To manage ionization variability and matrix effects, especially at low concentrations typical of exogenous chemicals, usage of isotopically labeled internal standards is highly effective[32] because these standards have identical chemical and physical properties to the target analyte and experience the same ionization effects, allowing for accurate quantification even in complex matrices.[33] When analytical standards are limited, which is particularly relevant for emerging exogenous chemicals, surrogate analyte or surrogate matrix approaches can be utilized.[34] The surrogate analyte method uses a stable isotope–labeled standard as a surrogate for calibration in the actual biological matrix, while the surrogate matrix method employs an alternative matrix devoid of the target analyte.[14] Both approaches require demonstration of parallelism between calibration standards and the target analyte in the biological matrix to ensure accurate quantification.[34] Where analyte-free matrices are not available, typical of endogenous chemicals, techniques such as standard addition or background subtraction can be employed.[14] By combining these strategies as needed for selected exogenous and endogenous chemicals, analysts can develop robust methods that achieve high confidence in concentration measures across different types of analytes, contributing to a comprehensive framework for reporting concentration measures from mass spectrometry-based assays.

While valuable as providing a framework to compare across methods, we recognize some limitations to this scheme, which can lead to subjectivity in reporting. In particular, there may be higher or lower confidence in a given sample measurement due to measurements below a limit of detection (LOD), the use of blank correction or recovery correction, calibration curve extrapolation outside calibration levels, or other analytical conditions that could apply during data acquisition and curation. The scheme does not address the range of annotation in untargeted data, which can affect analyte specificity.[2] Further, the choice of a surrogate for quantitation can also impact the reliability of the measures, especially for complex biological samples.[35] As a result, regulatory clinical or environmental risk assays, in particular,

will likely require significant adaptation to the proposed scheme to meet standards or fit-for-purpose.[36] However, as an initial framework, we capture most bioanalytical scenarios while maintaining a simple reporting process.

The variability in analytical capacity across institutions and limited access to high-quality standards are indeed significant barriers to achieving level 1 confidence in the proposed framework similar to challenges discussed by Clarke et al.[37] For those laboratories with reliability concerns for quantification confidence, we propose adopting a tiered approach similar to that suggested by Sarmad et al.[38] This "fit-for-purpose" strategy can accommodate various research settings and analytical capabilities, allowing laboratories to implement the highest level of quality assurance feasible within their constraints.[38] For laboratories with limited resources, we recommend prioritizing the principle of utility in analyte selection, focusing on practical value rather than quantity.[39] Additionally, emphasizing methodological innovation for developing high-throughput equipment could contribute to increased automation and accessibility.[39] To address the challenge of limited access to high-quality standards, we suggest focusing on the development of reference methods and materials across laboratories.[39] This can help ensure reliability and reproducibility across different laboratories, even those with resource constraints. By incorporating these considerations and providing flexible solutions, we believe the framework can be used across a wide range of research settings, ultimately improving the quality of reporting and comparability of mass spectrometry-based assays in various fields.[2,39]

Our confidence scheme focused on molecular mass spectrometry methods for organic chemicals. However, there are similarly stringent QC thresholds in atomic mass spectrometry for quantifying trace elements. Atomic mass spectrometry is based on argon plasma, which dramatically reduces matrix components compared to molecular mass spectrometer applications.[40] In addition, the main factors that determine the relative sensitivity of different elements are fixed by the atomic first ionization potential, atomic mass, and isotopic abundance.[40] Therefore, in some applications, external calibration methods or using single standards can generate quantitative measures that may or may not reach the thresholds considered for absolute quantification. Therefore, the development of a similar scheme may be helpful to the trace elements community for discussion and use.

Technology advancements in HRMS now enable increasingly reliable measurements. These methods approach reliability provided by quantitative assays that were once limited to QqQ MS. As novel bioanalytic workflows continue to strive toward broader analyte coverage, lower limits of detection, and better precision, established criteria for reporting only targeted or untargeted measures—at opposite ends of the spectrum—are no longer sufficient. Therefore, we propose a scheme for reporting concentration confidence measures from MS-based bioassays. The scheme maximizes the utility of the available measurement techniques while still communicating the associated error. We anticipate that this scheme will help clarify the muddled middle—to facilitate clear and concise reporting of measures on the same scale across a broad range of confidences, including targeted and semiquantitative measures. The current scheme is simple and provides flexibility for application in emerging technologies to report measurements or as target metrics for method development. Through individual use and dissemination of the concept at symposia, conferences, and consortia with the metabolomics, exposomics, and biomonitoring communities, we aspire that the community will adopt and refine this framework. As analytical methods evolve and adapt, we encourage discussion across the bioanalytical communities on concentration confidence schemes and best reporting practices.

Acknowledgments

We are grateful for support from the National Institutes of Health grants P30ES023515 and U01CA066572 and HHEAR U2CES030859, U2CES026561, and U2CES026555.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. LaKind JS, Sobus JR, Goodman M, Barr DB, Fürst P, Albertini RJ, et al. 2014. A proposal for assessing study quality: biomonitoring, environmental epidemiology, and short-lived chemicals (BEES-C) instrument. Environ Int 73:195–207, PMID: 25137624, https://doi.org/10.1016/j.envint.2014.07.011.
2. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. 2014. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ Sci Technol 48(4):2097–2098, PMID: 24476540, https://doi.org/10.1021/es5002105.
3. Chen Y, Zhou Z, Yang W, Bi N, Xu J, He J, et al. 2017. Development of a data-independent targeted metabolomics method for relative quantification using liquid chromatography coupled with tandem mass spectrometry. Anal Chem 89(13):6954–6962, PMID: 28574715, https://doi.org/10.1021/acs.analchem.6b04727.
4. Zhang L, Zheng W, Li X, Wang S, Xiao M, Xiao R, et al. 2021. A merged method for targeted analysis of amino acids and derivatives using parallel reaction monitoring combined with untargeted profiling by HILIC-Q-Orbitrap HRMS. J Pharm Biomed Anal 203:114208, PMID: 34148019, https://doi.org/10.1016/j.jpba.2021.114208.
5. Goodrich JA, Walker D, Lin X, Wang H, Lim T, McConnell R, et al. 2022. Exposure to perfluoroalkyl substances and risk of hepatocellular carcinoma in a multiethnic cohort. JHEP Rep 4(10):100550, PMID: 36111068, https://doi.org/10.1016/j.jhepr.2022.100550.
6. Petrick LM, Wolff MS, Barupal D, Teitelbaum SL. 2022. Comparison of untargeted and targeted perfluoroalkyl acids measured in adolescent girls. Chemosphere 290:133303, PMID: 34921852, https://doi.org/10.1016/j.chemosphere.2021.133303.
7. Chen L, Zhong F, Zhu J. 2020. Bridging targeted and untargeted mass spectrometry-based metabolomics via hybrid approaches. Metabolites 10(9):348, PMID: 32867165, https://doi.org/10.3390/metabo10090348.
8. Amer B, Deshpande RR, Bird SS. 2023. Simultaneous quantitation and discovery (SQUAD) analysis: combining the best of targeted and untargeted mass spectrometry-based metabolomics. Metabolites 13(5):648, PMID: 37233689, https://doi.org/10.3390/metabo13050648.
9. Zhou Z, Chen Y, Gao Y, Bi N, Yue X, He J, et al. 2020. Development of a high-coverage metabolome relative quantitative method for large-scale sample analysis. Anal Chim Acta 1109:44–52, PMID: 32252904, https://doi.org/10.1016/j.aca.2020.02.049.
10. Xu J, Li J, Zhang R, He J, Chen Y, Bi N, et al. 2019. Development of a metabolic pathway-based pseudo-targeted metabolomics method using liquid chromatography coupled with mass spectrometry. Talanta 192:160–168, PMID: 30348373, https://doi.org/10.1016/j.talanta.2018.09.021.
11. Zheng F, Zhao X, Zeng Z, Wang L, Lv W, Wang Q, et al. 2020. Development of a plasma pseudotargeted metabolomics method based on ultra-high-performance liquid chromatography-mass spectrometry. Nat Protoc 15(8):2519–2537, PMID: 32581297, https://doi.org/10.1038/s41596-020-0341-5.
12. Broeckling CD, Beger RD, Cheng LL, Cumeras R, Cuthbertson DJ, Dasari S, et al. 2023. Current practices in LC-MS untargeted metabolomics: a scoping review on the use of pooled quality control samples. Anal Chem 95(51):18645–18654, PMID: 38055671, https://doi.org/10.1021/acs.analchem.3c02924.
13. Cheng WL, Markus C, Lim CY, Tan RZ, Sethi SK, Loh TP, IFCC Working Group on Method Evaluation Protocols. 2023. Calibration practices in clinical mass spectrometry: review and recommendations. Ann Lab Med 43(1):5–18, PMID: 36045052, https://doi.org/10.3343/alm.2023.43.1.5.
14. Thakare R, Chhonker YS, Gautam N, Alamoudi JA, Alnouti Y. 2016. Quantitative analysis of endogenous compounds. J Pharm Biomed Anal 128:426–437, PMID: 27344632, https://doi.org/10.1016/j.jpba.2016.06.017.
15. US FDA (US Food and Drug Administration). 2018. *Bioanalytical Method Validation Guidance for Industry*. Silver Spring, MD: US FDA. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/bioanalytical-method-validation-guidance-industry [accessed 15 January 2025].
16. Charbonnet JA, McDonough CA, Xiao F, Schwichtenberg T, Cao D, Kaserzon S, et al. 2022. Communicating confidence of per- and polyfluoroalkyl substance identification via high-resolution mass spectrometry. Environ Sci Technol Lett 9(6):473–481, PMID: 35719859, https://doi.org/10.1021/acs.estlett.2c00206.
17. Wenk MR. 2005. The emerging field of lipidomics. Nat Rev Drug Discov 4(7):594–610, PMID: 16052242, https://doi.org/10.1038/nrd1776.

18. Jenkins B, West JA, Koulman A. 2015. A review of odd-chain fatty acid metabolism and the role of pentadecanoic acid (C15:0) and heptadecanoic acid (C17:0) in health and disease. Molecules 20(2):2425–2444, PMID: 25647578, https://doi.org/10.3390/molecules20022425.
19. Cajka T, Smilowitz JT, Fiehn O. 2017. Validating quantitative untargeted lipidomics across nine liquid chromatography-high-resolution mass spectrometry platforms. Anal Chem 89(22):12360–12368, PMID: 29064229, https://doi.org/10.1021/acs.analchem.7b03404.
20. Tehrani MW, Yang KX, Parsons PJ. 2020. Development and characterization of reference materials for trace element analysis of keratinized matrices. Anal Bioanal Chem 412(8):1847–1861, PMID: 32020317, https://doi.org/10.1007/s00216-020-02432-y.
21. Babadi RS, Riederer AM, Sampson PD, Sathyanarayana S, Kavanagh TJ, Krenz JE, et al. 2022. Longitudinal measures of phthalate exposure and asthma exacerbation in a rural agricultural cohort of Latino children in Yakima Valley, Washington. Int J Hyg Environ Health 243:113954, PMID: 35588565, https://doi.org/10.1016/j.ijheh.2022.113954.
22. Yoon LS, Binder AM, Pereira A, Calafat AM, Shepherd J, Corvalán C, et al. 2022. Variability in urinary phthalates, phenols, and parabens across childhood and relation to adolescent breast composition in Chilean girls. Environ Int 170:107586, PMID: 36302292, https://doi.org/10.1016/j.envint.2022.107586.
23. Silva MJ, Slakman AR, Reidy JA, Preau JL, Herbert AR, Samandar E, et al. 2004. Analysis of human urine for fifteen phthalate metabolites using automated solid-phase extraction. J Chromatogr B Analyt Technol Biomed Life Sci 805(1):161–167, PMID: 15113553, https://doi.org/10.1016/j.jchromb.2004.02.038.
24. Ye X, Kuklenyik Z, Needham LL, Calafat AM. 2005. Automated on-line column-switching HPLC-MS/MS method with peak focusing for the determination of nine environmental phenols in urine. Anal Chem 77(16):5407–5413, PMID: 16097788, https://doi.org/10.1021/ac050390d.
25. Mazzella MJ, Barr DB, Kannan K, Amarasiriwardena C, Andra SS, Gennings C. 2021. Evaluating inter-study variability in phthalate and trace element analyses within the Children's Health Exposure Analysis Resource (CHEAR) using multivariate control charts. J Expo Sci Environ Epidemiol 31(2):318–327, PMID: 33603093, https://doi.org/10.1038/s41370-021-00293-w.
26. Kannan K, Stathis A, Mazzella MJ, Andra SS, Barr DB, Hecht SS, et al. 2021. Quality assurance and harmonization for targeted biomonitoring measurements of environmental organic chemicals across the children's health exposure analysis resource laboratory network. Int J Hyg Environ Health 234:113741, PMID: 33773388, https://doi.org/10.1016/j.ijheh.2021.113741.
27. Bukowski MR, Voeller K, Jahns L. 2018. Simple and sensitive dilute-and-shoot analysis of carotenoids in human plasma. J Chromatogr B Analyt Technol Biomed Life Sci 1095:32–38, PMID: 30041087, https://doi.org/10.1016/j.jchromb.2018.07.020.
28. Goodrich JA, Walker DI, He J, Lin X, Baumert BO, Hu X, et al. 2023. Metabolic signatures of youth exposure to mixtures of per- and polyfluoroalkyl substances: a multi-cohort study. Environ Health Perspect 131(2):027005, PMID: 36821578, https://doi.org/10.1289/EHP11372.
29. Gu H, Zhang P, Zhu J, Raftery D. 2015. Globally optimized targeted mass spectrometry: reliable metabolomics analysis with broad coverage. Anal Chem 87(24):12355–12362, PMID: 26579731, https://doi.org/10.1021/acs.analchem.5b03812.
30. Houghton R, Horro Pita C, Ward I, Macarthur R. 2009. Generic approach to validation of small-molecule LC–MS/MS biomarker assays. Bioanalysis 1(8):1365–1374, PMID: 21083084, https://doi.org/10.4155/bio.09.139.
31. Kwon H, Lehotay SJ, Geis-Asteggiante L. 2012. Variability of matrix effects in liquid and gas chromatography–mass spectrometry analysis of pesticide residues after QuEChERS sample preparation of different food crops. J Chromatogr A 1270:235–245, PMID: 23182936, https://doi.org/10.1016/j.chroma.2012.10.059.
32. Panuwet P, Hunter RE, D'Souza PE, Chen X, Radford SA, Cohen JR, et al. 2016. Biological matrix effects in quantitative tandem mass spectrometry-based analytical methods: advancing biomonitoring. Crit Rev Anal Chem 46(2):93–105, PMID: 25562585, https://doi.org/10.1080/10408347.2014.980775.
33. Heudi O, Winter S, Picard F. 2020. Matrix matching in quantitative bioanalysis by LC-MS/MS a dream or a reality? Bioanalysis 12(2):67–69, PMID: 31854194, https://doi.org/10.4155/bio-2019-0273.
34. Jones BR, Schultz GA, Eckstein JA, Ackermann BL. 2012. Surrogate matrix and surrogate analyte approaches for definitive quantitation of endogenous biomolecules. Bioanalysis 4(19):2343–2356, PMID: 23088461, https://doi.org/10.4155/bio.12.200.
35. Liigand P, Liigand J, Cuyckens F, Vreeken RJ, Kruve A. 2018. Ionisation efficiencies can be predicted in complicated biological matrices: a proof of concept. Anal Chim Acta 1032:68–74, PMID: 30143223, https://doi.org/10.1016/j.aca.2018.05.072.
36. Pu S, McCord JP, Bangma J, Sobus JR. 2024. Establishing performance metrics for quantitative non-targeted analysis: a demonstration using per- and polyfluoroalkyl substances. Analytical and Bioanalytical Chemistry 416(5):1249–1267, PMID: 38289355, https://doi.org/10.1007/s00216-023-05117-4.
37. Clarke W, Rhea JM, Molinaro R. 2013. Challenges in implementing clinical liquid chromatography-tandem mass spectrometry methods–seeing the light at the end of the tunnel. J Mass Spectrom 48(7):755–767, PMID: 23832931, https://doi.org/10.1002/jms.3214.
38. Sarmad S, Viant MR, Dunn WB, Goodacre R, Wilson ID, Chappell KE, et al. 2023. A proposed framework to evaluate the quality and reliability of targeted metabolomics assays from the UK Consortium on Metabolic Phenotyping (MAP/UK). Nat Protoc 18(4):1017–1027, PMID: 36828894, https://doi.org/10.1038/s41596-022-00801-8.
39. Yu S, Zou Y, Ma X, Wang D, Luo W, Tang Y, et al. 2024. Evolution of LC-MS/MS in clinical laboratories. Clin Chim Acta 555:117797, PMID: 38280490, https://doi.org/10.1016/j.cca.2024.117797.
40. Clases D, Gonzalez de Vega R. 2022. Facets of ICP-MS and their potential in the medical sciences—part 1: fundamentals, stand-alone and hyphenated techniques. Anal Bioanal Chem 414(25):7337–7361, PMID: 36028724, https://doi.org/10.1007/s00216-022-04259-1.