



Best practices in NMR metabolomics: Current state

Robert Powers^{a,p,*}, Erik R. Andersson^b, Amanda L. Bayless^c, Robert B. Brua^d, Mario C. Chang^e, Leo L. Cheng^f, Chaevien S. Clendinen^g, Darcy Cochran^{a,p}, Valérie Copié^h, John R. Cort^g, Alexandra A. Crook^{a,p}, Hamid R. Eghbalniaⁱ, Anthony Giacalone^e, Goncalo J. Gouveia^j, Jeffrey C. Hochⁱ, Micah J. Jeppesen^{a,p}, Amith S. Maroli^{a,p,2}, Matthew E. Merritt^e, Wimal Pathmasiri^k, Heidi E. Roth^{a,p}, Anna Rushin^e, Isin T. Sakallioglu^{a,p}, Saurav Sarma^{l,1}, Tracey B. Schock^c, Lloyd W. Sumner^l, Panteleimon Takis^{m,q}, Mario Uchimiyaⁿ, David S. Wishart^o, Metabolomics Association of North America (MANA), NMR Special Interest Group, Metabolomics Quality Assurance & Quality Control Consortium (mQACC), NMR Technology Group

^a Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0304, USA

^b Department of Biological Sciences, University of Illinois at Chicago, 845 W. Taylor Street, Chicago, IL 60607, USA

^c National Institute of Standards and Technology (NIST), Chemical Sciences Division, Charleston, SC 29412, USA

^d Environment and Climate Change Canada, National Hydrology Research Centre, 11 Innovation Blvd, Saskatoon, Saskatchewan S7N 3H5, Canada

^e Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL, 32611, USA

^f Department of Pathology and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

^g Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352, USA

^h Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59715, USA

ⁱ Department of Molecular Biology and Biophysics, UConn Health, Farmington, CT 06030-3305, USA

^j Institute for Bioscience and Biotechnology Research, National Institute of Standards and Technology, University of Maryland, Godelsky Drive, Rockville, MD, 20850, USA

^k Nutrition Research Institute, Department of Nutrition, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

^l University of Missouri, Department of Biochemistry, MU Metabolomics Center, Bond Life Sciences Center, Interdisciplinary Plant Group, Columbia, MO 65211, USA

^m Section of Bioanalytical Chemistry, Division of Systems Medicine, Department of Metabolism, Digestion and Reproduction, Imperial College London, London SW7 2AZ, UK

ⁿ Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA

^o Departments of Biological Sciences and Computing Science, University of Alberta, Edmonton, Alberta, T6G 2E8, Canada

^p Nebraska Center for Integrated Biomolecular Communication, University of Nebraska-Lincoln, Lincoln, NE 68588-0304, USA

^q National Phenome Centre, Department of Metabolism, Digestion and Reproduction, Imperial College London, London W12 0NN, UK

ARTICLE INFO

Keywords:

NMR
Metabolomics
Best practices
Literature survey

ABSTRACT

A literature survey was conducted to identify current practices used by NMR metabolomics investigators when conducting and reporting their metabolomics studies. A total of 463 papers from 2020 to 80 papers from 2010 were selected from PubMed and were manually analyzed by a team of investigators to assess the extent and completeness of the experimental procedures and protocols reported. A significant number of the papers did not report on essential experimental details, incompletely stated which statistical methods were used, improperly applied supervised multivariate statistical analyses, or lacked validation of statistical models. A large diversity of protocols and software were identified, which suggests a lack of consensus and a relatively limited use of commonly agreed upon standards for conducting and reporting NMR metabolomics studies. The overall intent of the survey is to inform and encourage the NMR metabolomics community to develop and adopt best practices for the field.

* Corresponding author. University of Nebraska-Lincoln, Department of Chemistry, Nebraska Center for Integrated Biomolecular Communication, 722 Hamilton Hall, Lincoln, NE, 68588-0304, USA.

E-mail address: rpowers3@unl.edu (R. Powers).

¹ Current affiliation: Bayer Crop Science, 700 Chesterfield PKWY W., Chesterfield MO 63017.

² Present address: Icahn School of Medicine at Mount Sinai, New York, New York-10029, USA.

<https://doi.org/10.1016/j.trac.2023.117478>

Received 14 July 2023; Received in revised form 20 October 2023; Accepted 4 December 2023

Available online 12 December 2023

0165-9936/© 2023 Elsevier B.V. All rights reserved.

1. Introduction

NMR metabolomics has beneficially impacted a diversity of fields, from issues related to the environment [1] and nutrition [2] to drug discovery [3–5] and disease evaluation [6–9]. As a result, NMR metabolomics, as a field, continues to experience expansive growth [10, 11], particularly for medical applications due to its potential for biomarker discovery [6–9,12] and *in vivo* examinations [13–17]. Overall, metabolomics has benefited from the technical advantages of NMR, which includes its simplicity in measurements, often without the need for sample pre-processing, ready metabolite quantification, and redundant means to validate metabolite assignments [16,18–23]. These advantages present a low barrier of entry into the field that has contributed to the growth of NMR metabolomics, but, unfortunately, has also led to an abundance of tenuous metabolomics research studies populating the scientific literature. Routine concerns include misunderstandings about the proper usage of statistical tools [24–28], missing details fundamental to the rigor and reproducibility of the study [29,30], and the absence of quality controls necessary to validate the conclusions [21, 31]. Research studies without a clearly defined scientific hypothesis that just catalog and measure metabolites raise additional concerns. Numerous poorly designed or executed studies that fail to generate reproducible or meaningful data have the potential to severely undermine the long-term perspective and value of metabolomics [32]. Simply put, the field risks a loss of confidence by the broader scientific community.

The lack of well-established best practices that have been widely adopted, deployed, and validated by the broader scientific community is clearly a contributing factor to the proliferation of scientifically questionable metabolomics reports. This is despite the efforts of several community-led efforts by COSMOS: Coordination of Standards in Metabolomics, MSI: Metabolomics Standards Initiative, mQACC: Metabolomics Quality Assurance & Quality Control Consortium, and MANA: Metabolomics Association of North America to establish standard practices and minimum reporting criteria [33–39]. Furthermore, a properly executed metabolomics study is highly dependent on an assortment of unique skills that includes expertise in analytical chemistry, separation techniques, and statistics, among others. The false perception that metabolomics is easy often leads to poorly conducted or designed studies due to missing expertise, an absence of pertinent experiences, or methods lacking a clear explanation. Again, insufficient guidance from community-certified experimental protocols may lead to these problematic outcomes. To address this critical issue, the NMR metabolomics community must move toward establishing and widely adopting a set of best practices along with standards specifically relevant to an NMR metabolomics study. Of course, to achieve this laudable goal, it is first necessary to identify the common problems or limitations routinely encountered in reported NMR metabolomics studies. This, in turn, requires assessing and characterizing the different experimental protocols, statistical methods, and software used by the metabolomics community because each investigator or research group follows a different if not unique data processing pipeline. Herein, we report the results of a detailed survey of the scientific literature to document exactly how NMR has been routinely used in metabolomics research. The overarching intent of this analysis is to outline both the successes and existing problems with current NMR metabolomics studies to inform best practices. The outcome of the literature survey lays a foundation for how an NMR metabolomics study should be designed, executed, and reported in the future.

2. Discussion

2.1. The set of 2010 and 2020 NMR-based metabolomics papers

The PubMed database was first queried to identify papers published in 2020 that described an NMR-based metabolomics study. An initial

query identified 844 potentially relevant papers. A manual analysis to remove papers focused on software, NMR protocol reviews and natural product identification reduced the total number of papers that were surveyed to 487. A query check confirmed that 61 of the removed papers were reviews. A further evaluation of the assigned papers by the project participants removed an additional 24 papers for various reasons, but mainly because the focus of the study was not metabolomics. In total, 463 manuscripts from 2020 were used in the survey. Using an identical approach, a total of 80 manuscripts from 2010 were used in the survey. Please see **Appendix A** for a detailed description of the methods used to complete the literature survey. The survey results from the 2010 and 2020 papers were then compared to determine if the practices used to report NMR metabolomics data and findings remained consistent or improved with time.

2.2. Differences in reported protocols between 2010 and 2020 papers

Excluding magnitude differences due to the larger number of papers in 2020, a comparison of the survey data between the two periods indicated an overall high similarity in the results. Most responses were remarkably similar between the two sets of papers. This overall lack in improvement in reporting standards between 2010 and 2020 suggests a strong reliance on citing prior studies as a source of established procedures, which themselves may not contain a complete protocol. This practice has resulted in the perpetuation of poor reporting practices. Breaking this cycle requires the community to adhere to new recommendations of best practices promoted by societies and consortia while acknowledging that the current scientific literature is riddled with problematic inaccurate, confusing, or incomplete content.

There were a few minor areas that indicated a modest improvement from 2010 to 2020. The average number of identified and quantified metabolites increased by over 50 % from 2010 to 2020. Similarly, there was a small increase from 0 to ~11 % in the number of studies that reported submitting metabolomics data to a repository. Another promising trend was the general increase of ~20–~30 % in the application and/or the reporting of standard statistical validation parameters such as *p*-values [40], R^2/Q^2 [41], and false discovery rate (FDR) corrections [42–44].

2.3. Study design results

A total of nine Yes-No survey questions were grouped together to broadly characterize how investigators reported their overall study design. A series of bar graphs are depicted in **Fig. 1** that summarizes these survey responses.

One of the most surprising and concerning outcomes of the survey was the observation that less than 50 % of all papers provided a clearly stated scientific hypothesis. One potential explanation for this result is the fact that metabolomics lends itself to metabolite discovery and the relatively simple goal of just cataloging detectable metabolites in a given biological sample. While the rationale behind such a study may be to inform future investigations, it may not be particularly valuable since any follow-up study would also likely discover the metabolites that are detectable as a first step of the project. Further, without the availability of the complete data set and procedures, the potential utility of any metabolomics project limited to cataloging the observed metabolites is dubious. In this regard, not a single paper in 2010 reported depositing their metabolomics data to an appropriate repository. To be fair, few journals required in 2010 or currently require the deposition of metabolomics data in data repositories, as the Metabolomics Workbench (established 2016) [45] and MetaboLights [46] are relatively new enterprises. With the emergence of metabolomics specific repositories, approximately 11 % of the 2020 papers indicated the data was deposited in such a repository, which represents a modest improvement. The use and reuse of metabolomics data is an important issue and an expanding need as evident by a recent request for information by the National

Cancer Institute (NOT-CA-23-00). Thus, the deposition of NMR-based metabolomics data into publicly accessible repositories must become more routine, seamless, and mandated by journals and funding agencies to enhance data deposition compliance.

Not unexpectedly, the vast majority, nearly 100 %, of both the 2010 and 2020 papers employed one-dimensional (1D) ^1H or ^{13}C NMR experiments. Troubling, only 45 % (2010) or 36 % (2020) of the papers surveyed relied on 2D NMR experiments for metabolite identification and validation. This correlates with the nearly non-existent (~6–7%) NMR metabolomics projects that relied on isotope labeling. Since 2D NMR experiments provide a valuable and complementary approach to validating metabolite assignments, the lack of a routine usage of 2D NMR raises concerns about the accuracy and reproducibility of metabolite assignments for compounds in matrices that are not well-established within the NMR metabolomics literature. Instead, it would be prudent to commonly employ a series of 2D NMR experiments such as the 2D ^1H - ^{13}C HSQC, HMBC, HSQC-TOCSY, and 2D ^1H - ^1H TOCSY for spectra collected on pooled samples to validate metabolite annotations [47]. Specifically, 2D NMR provides multiple, correlated chemical shifts that may uniquely identify a metabolite. COLMAR and similar software tools provide an efficient and reliable approach to leverage these 2D NMR spectral signals to accurately annotate complex metabolomics mixtures [48–50]. This level of confidence cannot be achieved by relying solely on 1D NMR.

On the other hand, the semi-automatic assignments of well-known and abundant metabolites commonly detected in 1D ^1H NMR spectra is likely to be reliable when analyzing a specific type of biofluid using specific small molecule spectral libraries [51–53]. It is important to note that the false positive rate increases proportionally with the size of the reference library, necessitating the use of a targeted library for accurate metabolite annotation extracted from 1D ^1H NMR spectral patterns [53]. Nevertheless, NMR can partially overcome this challenge given the fact that most metabolites produce numerous signals in a 1D ^1H NMR spectrum. In this regard, matching multiple redundant signals between the experimental and reference spectrum greatly improves confidence in the assignment. This is a unique and valuable advantage of NMR compared to other analytical techniques that produce only a single signal per compound, especially for the analysis of complex mixtures.

Only ~46–50 % of the surveyed papers reported an observed consistency between the key identified metabolites with other replicate studies. An observed consistency with these previous studies raises the likelihood that the metabolites identified have been correctly assigned. The reason why a comparison with prior literature is not a common occurrence in the metabolomics field is perplexing. Routine literature retrospection in the discussion section of any NMR metabolomics publications provides confidence in the analyzed data.

A failure to demonstrate consistency with previous studies raises broad concerns regarding the robustness and reproducibility of metabolomics data. Part of the issue may be a conservative outlook in which variations in sample collections, handling, and preparations may deem a comparison between studies inappropriate. It is also perfectly legitimate that a given study is unique and lacks prior replicate studies to compare

against. Given the exponential growth of metabolomics studies, especially the large number of replicate clinical studies designed to identify biomarkers, the complete absence of any related prior studies is highly unlikely. Even in cases where there are legitimate concerns regarding important differences in experimental design decisions, a comparison of outcomes and procedures may still be informative. Metabolomics papers that incorrectly present a study “as the first of its kind” and fail to place their results in context with the scientific literature may hinder progress in the field. Opportunities to establish results that have been successfully replicated or to identify potential contradictions have been missed. Importantly, if disagreements are not revealed at the time of publication, the opportunity to understand or identify the source of the problem and to potentially be able to fix the inconsistencies may be lost. More troubling is the possibility that incorrect scientific conclusions may be propagated through the community leading to detrimental outcomes that may involve patients.

2.4. Statistics and quality control

The application of univariate [54] and multivariate [41] statistics is a common and often necessary aspect of a metabolomics study. This is consistent with the results of the survey (Fig. 1) where ~70–75 % of the 2010 and 2020 papers reported using univariate statistics, ~75 % of the studies used unsupervised multivariate statistics, ~52–73 % of the papers reported using supervised multivariate statistics, and ~61–68 % of the studies used both univariate and multivariate statistics. Since most metabolomics investigators are not necessarily trained biostatisticians, the proper application of statistics to a metabolomics study is a common concern, especially if a biostatistician has not been part of the study design [26–28,40,41,55]. This concern is underscored by the results of our survey. Summarized in Fig. 2A are the results of the five Yes-No survey questions related to proper statistical usage. The poor performance in terms of properly reporting or using statistics is readily apparent. The good news is that there is a notable improvement between papers published in 2010 compared to 2020. For example, only ~60 % of the 2010 papers reported a p -value, a fundamental assessment of the statistical significance of identified metabolites. Without this statistical measure of confidence, how is it even possible to determine reliability or the utility of the results from a metabolomics study? The number reached nearly 80 % for 2020 papers. Nevertheless, the number of papers reporting an FDR or multiple hypothesis correction method such as Benjamini-Hochberg [56] or Bonferroni [43,44] was unacceptably low, corresponding to only ~14 % for 2010 and ~34 % for 2020. An FDR correction of any omics data set is essential since errors accumulate as illustrated in eqn. (1):

$$p = 1 - (1 - \alpha)^m \quad (1)$$

where m is the number of hypotheses or metabolites and α is the statistical significance level or desired p -value, typically ≤ 0.05 . The lack of reported FDR corrected p -values likely means a significant number of the putatively identified metabolites are false positives. This, in turn, will negatively impact any biological significance or outcomes of the study.

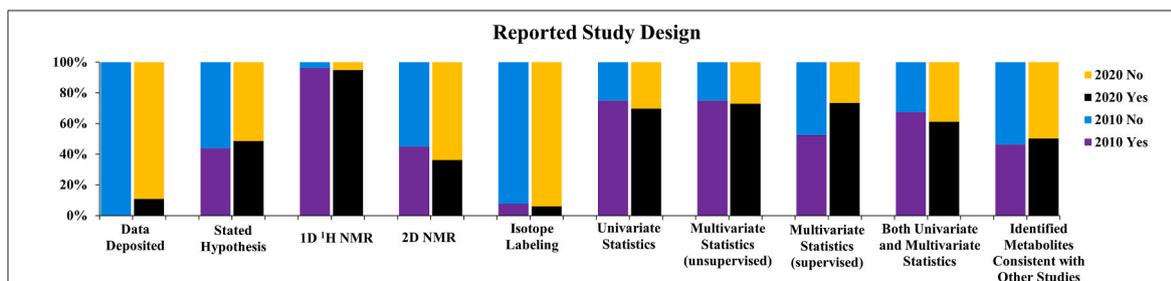


Fig. 1. Bar graphs of Yes-No survey questions related to study design of NMR-based metabolomic studies.

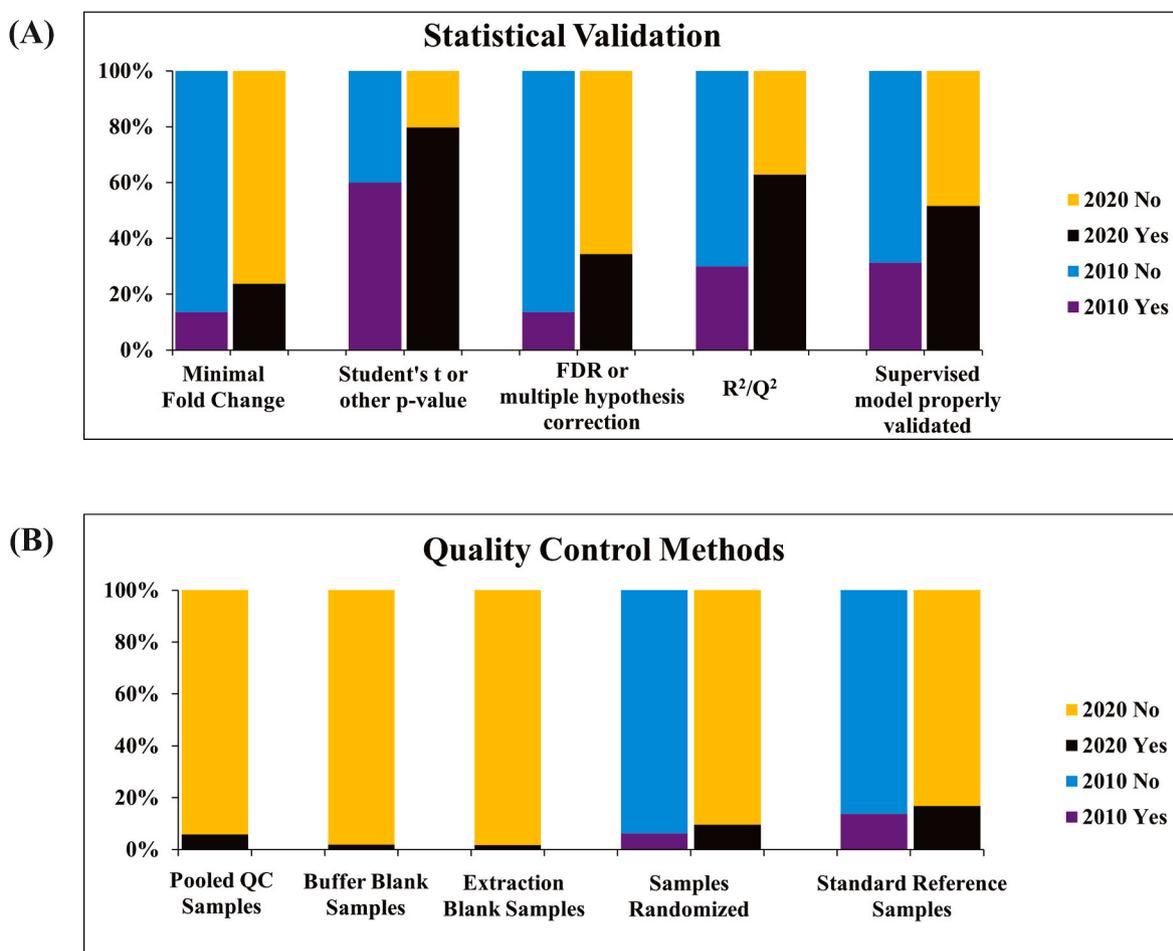


Fig. 2. Bar graphs of Yes-No survey questions related to (A) statistical validation and (B) quality control methods of NMR-based metabolomic studies.

An FDR corrected p -value [57,58] along with a quantitative measure such as fold change is recommended as a requirement for reporting all metabolites identified as differing significantly between two or more groups in a published metabolomics manuscript.

Effect size is another important statistical parameter to report, but rarely is, and, accordingly, was not specifically tabulated in our survey. Fold change, which was tabulated in our survey, is only one of various metrics that can be used to assess effect size. The measurement of an effect size is mandatory for the validation of a biomarker (e.g., a disease model predictive accuracy) and requires an independent set of samples to accomplish this assessment, which was not the main subject of the papers comprising our literature survey. Therefore, it has been difficult to make any assessment on the usage of effect size from the current survey.

The problems with statistics are not limited to univariate analysis. The survey results also indicated that a low number of papers listed standard quality factors, R^2 and Q^2 , associated with multivariate models, such as principal component analysis (PCA), partial least squares or projections to latent structures (PLS), orthogonal projections to latent structures (OPLS), and projections to latent structures-discriminant analysis (PLSDA) that are routinely used in analysis of metabolomics data [41]. Only ~30 % of 2010 papers reported R^2 and Q^2 values, but fortunately, the number more than doubled to ~63 % in 2020. Still, in the same way that a p -value is needed to assess the statistical significance of a metabolite change, R^2 and Q^2 values are important to evaluate the quality of a multivariate statistical model. Specifically, R^2 measures the degree of fit to the model and Q^2 provides a measure of consistency between the predicted and original data. Typically, R^2 is greater than Q^2 , $Q^2 \geq 0.4$, and Q^2 is within 20 % of R^2 for an acceptable model [41].

Importantly, R^2 and Q^2 do not provide a measure of model validation, especially for supervised methods including PLS, OPLS, and PLSDA. Instead, an additional model validation approach is needed, which usually consists of p -values calculated by CV-ANOVA [59] and/or a permutation test ($n = 1000$) [60]. Disappointingly, our literature survey indicated that only ~31 % of 2010 papers reported an acceptable validation method for a supervised multivariate statistical model. The numbers did modestly improve to ~43 % in 2020. Nevertheless, the low routine validation of statistical models raises serious concerns regarding the scientific validity of the entirety of these studies. Simply, if the statistical models are revealed to be invalid by being overfit, then any resulting scientific insight or biological inference is equally suspect. In essence, the entire outcome of the study is in doubt. Therefore, we suggest statistical model validations be reported for metabolomics studies that employ multivariate statistical methods, which may include reporting R^2 , Q^2 and p -values calculated by CV-ANOVA [54] and a permutation test, or other validation results (CER, AUROC, random forest analysis).

Unfortunately, the survey results summarizing the NMR community's application of quality control (QC) methods directly pertinent to the reported study were worse than the reported statistical data (Fig. 2B). In fact, the response rate was so low, near zero for most categories, that a few of the questions were eliminated from the 2010 survey. The response rates ranged from a low of ~2 % to a high of nearly 17 % for the 2020 papers. Given the intrinsically high reproducibility of NMR experiments, the high instrument stability, and the relatively low variance (CV 5–10 %) of quantitative NMR (qNMR) [22], QC has not been as strong a concern among NMR-based metabolomics investigators as it is for LC/GC-MS metabolomics data [61]. In fact, a few of the survey

questions, such as pooled QC samples, are standard protocols in MS-based metabolomics studies, but have not been universally adopted by the NMR community. Nevertheless, randomization of samples to minimize potential sample-order bias is an important practice for all metabolomics studies. To note, pooled QC samples offer an evaluation of the experimental variance that may occur during sample preparation and metabolite extraction and remain important for an NMR-based study. For example, combining pooled samples with 2D NMR may prove to be an efficient approach to validate metabolite assignments. Other important QC samples include process blanks, buffer blanks, standard reference samples for system suitability, standards for chemical shift referencing and even standard reference materials for quantitation and assessment of technical variance. Of course, most NMR facility managers conduct routine quality control checks of NMR instrument performance that are rarely reported in the literature. We suggest that manuscripts reporting on the use of NMR for metabolomics include pertinent information regarding the QA/QC procedures implemented for assessing spectrometer performance, including stability of 90° pulses, water suppression efficiency, and signal-to-noise values on standard samples, sample temperatures, or ERETIC signal calibrations, among other routine QA/QC procedures. The survey questions regarding QA/QC did not address whether these routine protocols were employed, but, instead, were directed to QA/QC protocols relevant to the reported study to determine variability that may emerge during sample preparation. Thus, we also recommend that other QC protocols specific to the published study such as the use of pooled, standard, and blank samples, and the randomization of sample preparation and data collection be reported.

2.5. Reporting of experimental parameters and data processing parameters

The inclusion of standard experimental parameters should be considered relatively routine and common practice. In this regard, a survey of scientific literature would be expected to reach nearly 100 %.

Any outcome significantly below this mark would be troubling, which is generally what was observed (Fig. 3A). The worst outcome from both the 2010 and 2020 survey was reporting sample pH, which only reached ~56 % of papers. Now, the relatively low reporting rate may be a partial artifact of studies that use only organic solvents where pH has no meaning. This may also partially explain the observation that only ~81 % of 2010 papers and ~74 % of 2020 papers reported a buffer. Nevertheless, it doesn't explain why only ~65 % (2010) or ~75 % (2020) of the survey papers reported a sample temperature. The number of samples used in a study were commonly reported but decreased from ~96 % in 2010 to ~83 % in 2020. More importantly, the number of samples was often not summarized in the methods section. Thus, readers had to deduce sample numbers by counting symbols in plots or careful examination of tables. A summary of the experimental design in the results section of a manuscript that includes the number of groups, the number of replicates (biological and analytical) per group, the number and type of NMR experiments would be a simple remedy to this problem.

Unfortunately, the reporting of processing parameters was significantly worse than the listing of experimental parameters for both the 2010 and 2020 papers (Fig. 3B). It is well-known that the methods and parameters of NMR data processing, such as baseline correction method (~68–75 %), automated or manual phasing (~58–98 %), choice of window function (~34–55 %), application of zero-filling (~21–33 %), and the removal of solvent/buffer peaks (~58–63 %), will directly impact the composition of the data matrix and the resulting statistical analysis. The 2020 papers reported a lower percentage for most of these important parameters.

As discussed above, the low reporting rate of important statistical parameters occurred again in the context of processing parameters. The fact that the data matrix was normalized and scaled is of critical importance [62,63]. This also requires knowledge of the type of normalization and data scaling technique that has been used. The absence or improper usage of normalization and scaling techniques would completely negate the value of any multivariate statistical model and any subsequent biological interpretation. Despite the importance of

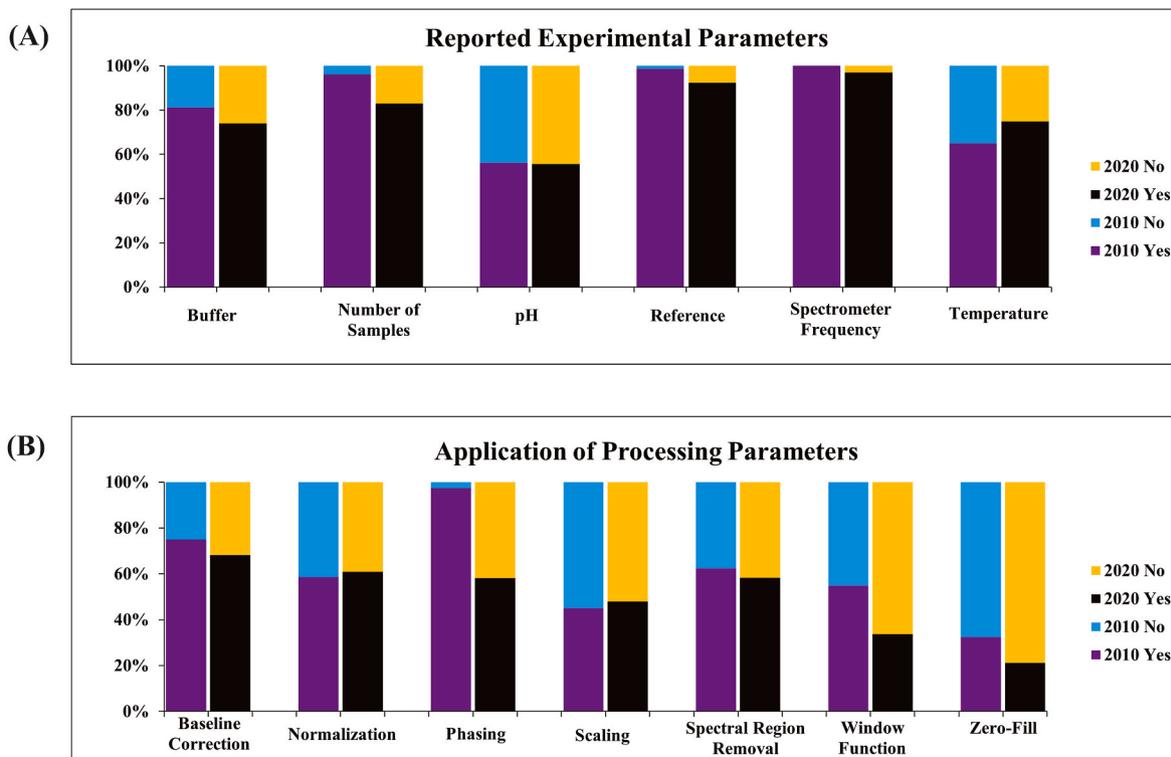


Fig. 3. Bar graphs of Yes-No survey questions related to (A) experimental parameters and (B) processing parameters of NMR-based metabolomic studies.

reporting such data manipulation, only ~60 % of both the 2010 and 2020 papers reported a normalization method which dropped to ~45 % of papers that reported data scaling.

A potential explanation for the low reporting of experimental, processing, and statistical parameters is the reliance on citing previously published papers that are expected to provide a detailed explanation of the metabolomics protocols used by the study. Frustratingly, sometimes this approach leads to the backwards referencing through multiple generations of papers before the original manuscript with the complete list of protocols are located. While the procedures themselves may eventually be found in the literature, the simple specification of the protocol is inadequate. As noted in a *Nature* survey from 2016 [64], more than 60 % of respondents were unable to repeat other scientists' experiments. The American Society for Cell Biologists similarly found that more than 70 % of its members were unable to replicate a published experimental result using standard well-established protocols because of incomplete details reported in the original protocols (<https://www.ascb.org/science-policy-public-outreach/advocacy-policy/ascb-examines-difficulty-reproducing-research-data/>). The individual parameters for each of those procedures are critical to ensure reproducibility of results and findings, which are often study and/or sample dependent and unlikely to be identical to the cited literature reference. A reference to a previous manuscript could be adequate if the citing paper contained sufficient data to indicate that all parameters from the previous study are identical to the current study. Lack of specific clear-cut statements indicating identical conditions, or mentioning specific exceptions, indicates that the simple description of the protocol or literature citation is insufficient. Regardless of the reason, the absence of key experimental and processing parameters (Fig. 3) may lead to legitimate concerns regarding the validity of the presented study and the scientific value of any biological interpretation. In our view, a well-defined and established procedure is one where the specific biological, physiological, experimental, and study conditions are fully described in the published manuscript. At a minimum, detailed study-specific experimental and

org/science-policy-public-outreach/advocacy-policy/ascb-examines-difficulty-reproducing-research-data/). The individual parameters for each of those procedures are critical to ensure reproducibility of results and findings, which are often study and/or sample dependent and unlikely to be identical to the cited literature reference. A reference to a previous manuscript could be adequate if the citing paper contained sufficient data to indicate that all parameters from the previous study are identical to the current study. Lack of specific clear-cut statements indicating identical conditions, or mentioning specific exceptions, indicates that the simple description of the protocol or literature citation is insufficient. Regardless of the reason, the absence of key experimental and processing parameters (Fig. 3) may lead to legitimate concerns regarding the validity of the presented study and the scientific value of any biological interpretation. In our view, a well-defined and established procedure is one where the specific biological, physiological, experimental, and study conditions are fully described in the published manuscript. At a minimum, detailed study-specific experimental and

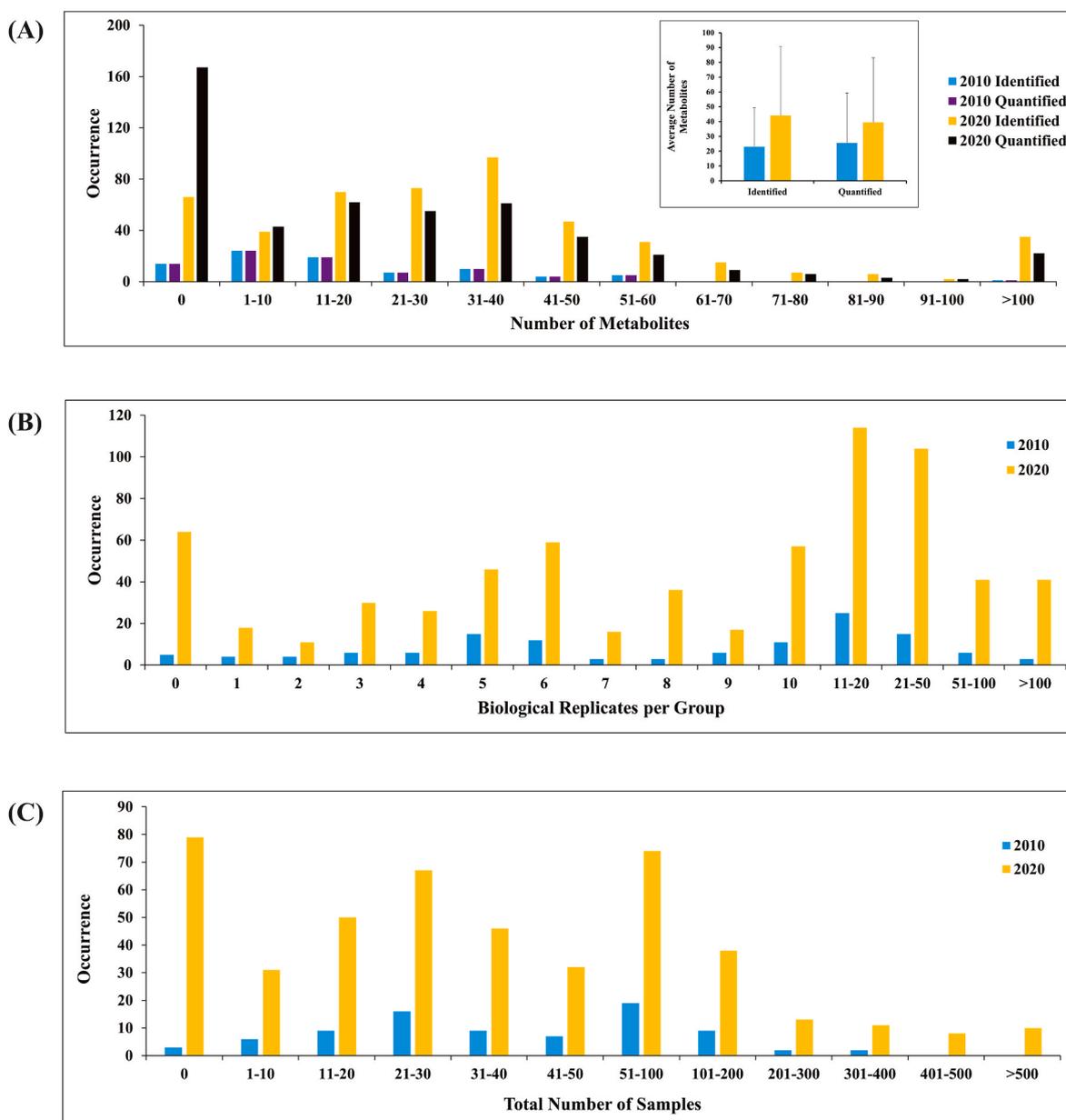


Fig. 4. Bar graphs summarizing the number of 2010 and 2020 papers reporting (A) the number of identified and quantified metabolites (binned), (B) the number of biological replicates per group, and (C) the total number of samples. The insert in (A) plots the average and standard deviation of the number of identified and quantified metabolites in 2010 and 2020.

processing parameters be described in supplemental material and included when depositing data in public data repositories.

2.6. Distribution of experimental parameters and study outcomes

Another clear outcome of the literature survey is the diversity of experimental protocols and the range of study designs used by the NMR metabolomics community. For example, the number of metabolites identified and quantified ranged from 0 to >100, with a peak at 0 and 31–40 metabolites in 2020 (Fig. 4A). Encouragingly, the average number of identified/quantified metabolites increased from 23/26 in 2010 to 44/39 in 2020. The standard deviations in all cases were essentially equal to the average values, which again highlights the high diversity of NMR metabolomics studies. To further clarify the study outcomes, the number of metabolites identified and quantified were also tabulated for several different sample types: animal models, biofluids, cell cultures, food/beverages, humans, plants, and tissues. There was only a modest difference in metabolite counts per sample type. For example, in 2020, biofluids and human samples had the largest average number of identified metabolites ranging from approximately 45 to 56, while animal models and plants had the smallest average number of metabolites ranging from 12 to 29. The remaining groups had an average that ranged between 32 and 37. Consistent with the overall trends, the number of metabolites identified and quantified were lower for 2010 publications, but the relative rankings of sample types changed. In 2010, animal models and food/beverages had the lowest average number of metabolites, ranging from approximately 7 to 13. Cell cultures, plants and tissues had the largest average number of metabolites, which ranged from 18 to 77. The remaining sample types had an average number ranging from 18 to 21. Given both group type and yearly variations, it is difficult to assign any importance to differences in sample types beyond limited knowledge and/or complexity of some metabolomes that were sampled by NMR. Nevertheless, it is concerning that an equally likely outcome of NMR-based metabolomics studies is to report 0 identified/quantified metabolites. This is easily fixable and largely a result of the authors' simple failure to report a summary of the number of metabolites identified and quantified.

The number of biological replicates per group and the total number of samples provide another factor to gauge the overall quality of a metabolomics study (Fig. 4B and C). In general, the more samples that are analyzed, the better for establishing statistical significance and the reliability of any observations. The peak in the number of biological replicates per group being centered around bins of 11–20 and 21–50 is a positive trend, as is the number of studies that analyzed 50–200 total samples. Again, a large range of values was observed, further establishing the diversity of NMR metabolomics studies that are undertaken. A common occurrence was zero, in which the study doesn't clearly state the number of samples used. Equally troubling is the large number of studies using a few (<6) and even a single biological replicate per group or only 1 to 10 total samples. The biological significance of these studies is highly suspect. Of course, there are practical considerations that constrain the number of available samples, such as in rare medical conditions, but a minimal number of samples still needs to be acquired to obtain the necessary statistical power. In highly controlled metabolomics studies using cell cultures, animal models, food, and beverage samples, etc., 10 to 20 biological replicates per group are likely sufficient for statistical significance. Conversely, human clinical metabolomics studies may require 60 or more replicates per group to obtain sufficient statistical power (α 0.8). Of course, these are only minimal recommendations for the number of biological replicates, as more replicates are always better.

It should also be mentioned that the number of analytical replicates reported across all 2010 and 2020 papers was one. There were only a few rare exceptions of a study reporting more than one analytical replicate. While the reproducibility of the NMR measurement is superior, possible processing deviations can be detected when including

analytical replicates for one or more samples. The practice of using analytical replicates ensures consistency within the metabolomics analytical workflow and offers confidence that the data is of high quality.

Other experimental parameters exhibited a similarly large range of values from the literature survey (*data not shown*). For example, reported pHs ranged from 2.5 to 7.50, with peaks at pH 6, 7, and 7.4. Despite the use of an apparently large range of pHs, about 80 (2010) to 90 (2020) percent of the reported pHs were within the 6.5 to 7.5 range. Although many studies used a pH consistent with typical biological conditions (*i. e.*, pH 7.2) for their metabolomics analyses, some studies demonstrated an occasional need to use different pHs based upon the specific nature of the samples. For example, analysis of samples at lower pH is important in food preservation studies. Similarly, reported temperatures ranged from 198 K to 448 K, with peaks at 298 K, 300 K, and 310 K. Despite the apparently large range of temperatures used in diverse studies, about 79 (2010) and 82 (2020) percent of the reported temperatures were either 298 K or 300 K, essentially ambient temperature. However, it is also clear that other temperatures were needed for specific studies to accommodate, for example, the need to cool samples for solid state NMR. For comparisons across studies and between different groups, it would be ideal to establish sample-type specific pH and temperature values for conducting metabolomics studies. Several studies have previously reported the evaluation of sample preparation for metabolomics studies [2,65,66].

Every field between 400 and 900 MHz has been used, where 600 MHz was the most popular choice, followed by 500 MHz. All modern NMR spectrometers can perform sophisticated multidimensional experiments that are pertinent to metabolomics, in addition to simple 1D NMR experiments. The advent of ultra-high field instruments, with the state of the art now being 1.2 GHz (^1H frequency), affords unprecedented sensitivity and resolution, and simplification of the spectra for strongly coupled spin systems [67]. While their staggering cost precludes widespread application, ambiguities resolved using high field instruments will enhance the utility of lower field instruments. Recent advances in inexpensive benchtop instruments [68], which can execute the sophisticated pulse sequences available on high field instruments, will enable wider application and penetration of NMR-based metabolomics in more diverse settings, for example field studies and point-of-care. Thus, the choice of a specific spectrometer frequency for a given metabolomics study is likely dictated by instrument availability, project needs, and access to high-throughput capabilities including automatic sample loading systems or cryoprobe systems. Thus, standardizing NMR field strength for metabolomics studies is not practical, which, in turn, requires more flexible pipelines and software for NMR data processing and analyses.

Finally, the choice of 1D NMR experiments, 2D NMR experiments, and the software and databases used to process and analyze the metabolomics datasets exhibited the largest variance between research groups. Consistent with prior responses, "No Response" was the most common or among the most common answers to the type of NMR experiments and software used. This is a concerning outcome considering how fundamentally important it is to know specific details regarding the type of NMR experiments conducted and how the spectral data were processed and analyzed. Again, a partial explanation for the omission of this critical information may be the reliance on citing previous work to provide experimental details, instead including this information in supplementary material as a minimal requirement is a reasonable expectation. The popular choices of 1D NMR experiments were 1D NOESY, CPMG, and a generic 1D ^1H NMR pulse sequence with or without the explicit inclusion of a presaturation pulse [69–74]. While widely used, each of these 1D NMR experiments has well-known and unique limitations with associated problems that may impact the veracity of the reported biological insights. Accordingly, a community-led recommendation to adopt specific NMR pulse sequences and experimental parameters would again be beneficial to establish standards and

enable the comparison of datasets across studies and groups. Conversely, popular choices for 2D NMR experiments are routinely employed for metabolite identification and validation and do not present any unique concerns. These experiments include expected selections such as HSQC, TOCSY, COSY, and HMBC [47,75]. A notable exception is the 2D J-resolved experiment, which could be viewed as an alternative 1D ^1H NMR experiment [76].

By far, the largest diversity in experimental protocols was the choice of software and database used to process and analyze the NMR metabolomics data sets. Once again, the number one response was “No Response.” While this assertion has become redundant, it is extremely challenging to evaluate the value of a metabolomics study and infer the reliability of its outcomes without a fundamental understanding of the data analysis protocols. The common reliance on prior published work to describe experimental protocols is a likely partial explanation for the absence of software being listed in the 2010 and 2020 papers. The top five software out of the 20 commonly reported programs were, in order, Chenomx, TopSpin, SIMCA, and R [77]. The missing member from this list, which was third overall, was the general category of “Other.” The other category contained over 90 different software programs. To further emphasize this point, of the 543 papers surveyed from 2010 to 2020, over 110 unique software packages were identified as being used to process or analyze the NMR metabolomics data sets. Nearly one out of every five papers used a unique NMR spectra processing approach. Again, community-led standardization of data processing pipelines and software may greatly benefit the field of metabolomics while minimizing problematic or tenuous outcomes. NMRbox (<https://nmrbox.nmrhub.org/>) [78], MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca/>) [79], and MVAPACK (<https://mvapack.unl.edu/>) [80], among others, are current on-going efforts to resolve this issue.

3. Conclusion

Herein we reported on an assessment of NMR-based metabolomics studies published in the scientific literature during the years 2010 and 2020. The outcome of our literature survey identified several areas of concern, but it also provided an important framework to establish best practices. One troubling observation was the high occurrence of papers that lacked important experimental details. A reliance on citing prior published papers as the source of these experimental details remains problematic. Instead, at a minimum, a proposed practice of including all experimental details and protocols in supplementary files is suggested. This information is critical to evaluating the proper execution of an NMR-based metabolomics study and ensuring the veracity of the results and biological insights, in addition to providing a means to reproduce the study results. The most severe omission was the incomplete reporting of statistics, which was made worse by routine improper usage of both univariate and multivariate statistical methods. Thus, robust guidelines need to be adopted by the NMR metabolomics community to ensure proper application, reporting, and validation of statistical methods critical to metabolomics analysis.

In addition to the lack of reporting key information regarding the design and execution of a study, it was extremely unlikely that the resulting data would be deposited in publicly available metabolomics repositories. Metabolomics data that is not publicly accessible eliminates the possibility to reexamine and/or reprocess the datasets to determine the reliability and reproducibility of the original results or even to reanalyze the research as software and data analytic tools progress. A lack of test data sets also hinders progress in software and method development and undermines the broad impact of metabolomics findings. Further perplexing was the observation that a majority of the papers did not include a stated scientific hypothesis. The importance of metabolomics goes well beyond simply cataloging the metabolites detected in a biological sample. Instead, the value of

metabolomics comes from its complementarity with other experimental techniques that provide a distinct and holistic view of the system under investigation. Thus, metabolomics is often uniquely positioned for hypothesis generation.

The final impression garnered from the survey data was the large diversity of protocols employed by NMR communities. The absence of a generally accepted and adhered-to standards for data acquisition and processing pipeline almost guarantees that the results acquired across multiple investigations will not be reproducible or interoperable. Without any semblance of consistency, the accuracy of the results is also highly questionable. The diversity of protocols reported in the literature presents a valuable framework for devising standards and best practices for NMR metabolomics protocols that address: (i) sample handling and processing, (ii) NMR data collection, (iii) data preprocessing protocols and software choice, (iv) statistical analysis, validation, and software choice, (v) metabolite identification, quantification, validation, software and database choices, (vi) network analysis, and (vii) combining NMR with other analytical platforms. In fact, a body of suggested standards related to various segments of the metabolomics pipeline does exist [31,33–37,81], including recommendations about data acquisition and processing [27,28,62,63,82–84], but compliance with reporting standards is lacking. For example, a study by Spicer et al. (2017) [38] pointed out that many of the minimal reporting standards established by the MSI: Metabolomics Standards Initiative [39] in 2007 were not adhered to in metabolomics datasets deposited in MetabolomeXchange (<http://www.metabolomexchange.org/>) or other data repositories. Unfortunately, our recent survey of the NMR metabolomics scientific literature confirms this ongoing challenge facing the metabolomics community, and the lack of an approved and agreed upon set of best practices is still a major drawback. Resolving this issue will require a concerted and joint effort by metabolomics researchers, journal publishers, metabolomics data repositories, societies, and funding agencies to embrace and enforce an adherence to community standards.

As a final note, given the assortment of protocols revealed by the survey, it is difficult to imagine a single agreed-upon approach to NMR-based metabolomics among so many different studies involving humans, cells, animals, food, and others. Instead, we expect that by focusing on a single type of biological system, for example, on human serum or urine, only then will an improved set of best practices appear for the specific system. For instance, metabolomics studies performed with cells or animal models usually require fewer biological replicates than human studies because of the controlled environment and lower sources of variance. Similarly, different sample matrices define sample handling and metabolome extraction protocols which determine the number and chemical class of metabolites that can be detected. Describing the reporting habits among the many investigators comprising this survey was important to better understand whether the methods currently employed are sufficiently detailed and adequate for achieving such a consensus reporting standard. Conversely, it is difficult to conclude something about the standardization of a metabolomics methodology if the data are not presented in a coherent and homogeneous manner. Establishing an analytical standard framework for NMR-based metabolomics necessitates a thorough assessment of the experimental, processing, and statistical parameters currently employed by the scientific community.

NIST disclaimer

Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

Declaration of competing interest

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

Data availability

All data is included as supplemental material.

Acknowledgments

This work was supported in part by funding from the National Institutes of Health (R01 AI148160, NIAID to R.P.), and the Nebraska Center for Integrated Biomolecular Communication (NIH NIGMS P20 GM113126 to R. P.). L. L. C. is supported by NIH AG070257. J.C.H. and H.R.E. are supported by NIH R24GM150793. P. T. is supported by the NIHR Imperial Biomedical Research Centre (BRC). V.C received support from NIH R01- DK117473 and NSF- MCB-1714556. This material is based upon work supported by the Defense Health Agency and U.S. Strategic Command under Contract No. FA4600-12-D-9000. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Health Agency, U.S. Strategic Command, or 55th Contracting Squadron.

Appendix B. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.trac.2023.117478>.

Appendix A. Methods

A.1 Assembling a collection of NMR-based metabolomics papers

The scientific literature deposited in the PubMed [85–87] database (<https://pubmed.ncbi.nlm.nih.gov/>) was searched for peer-reviewed NMR-based metabolomics manuscripts published in 2010 and 2020. The database search employed a combination of an automatic and manual filtering process. The following PubMed query was used to generate an initial list of manuscripts:

```
((((NMR [Title/Abstract]) AND (metabolomics [Title/Abstract]))) AND ("2020/01/01" [Date - Publication]: "2020/12/31" [Date - Publication]))) NOT (Review).
```

Any paper focused on software and/or NMR protocols, and any reviews that were not detected with the first filter were manually removed. Also, any paper describing the identification of a natural product that used mass spectrometry (MS) metabolomics to detect the compounds followed by NMR to determine the structure were removed. The following PubMed query was then used to cross check the list of publications that were identified with the first PubMed query to confirm that the articles selected via the review criteria were review articles:

```
((((NMR [Title/Abstract]) AND (metabolomics [Title/Abstract]))) AND ("2020/01/01" [Date - Publication]: "2020/12/31" [Date - Publication])))
```

The process was repeated with the year 2020 changed to 2010.

A.2 Creating and conducting the survey of NMR-based metabolomics papers

A list of 72 survey questions was developed by the project participants and are available in the supplemental information (Table B1). The survey questions were designed to thoroughly catalog current NMR metabolomics best practices as described in the scientific literature. The number of survey questions were reduced to 59 for 2010 papers to eliminate uninformative questions. Omitted questions were associated

with automation, number of analytical replicates that were either omitted or nearly always one, number of controls and cases that was redundant with number of groups, number of zero fills, final file size, and baseline correction function that were difficult to properly assess, and quality control and machine learning protocols not employed by NMR investigators. Most questions were designed with a simple response of “No”, “Yes”, or “Not Indicated.” A few questions had a defined pull-down list of likely answers, such as spectrometer frequencies. The remaining questions were open-ended such as the type of buffer, software, or biological sample used in the study. Overall, the survey questions were comprised of the following general categories: (i) study design, type of samples, number of groups and biological replicates, (ii) experimental details like field strength, temperature, pH, and solvent, (iii) processing details like zero fills, baseline correction, normalization, and scaling, (iv) quality controls, software used, metabolites identified, and metabolites quantified, and (v) statistical details like univariate techniques, multivariate techniques, and validation methods.

An Excel file of the list of authors, titles, and PMIDs from the PubMed query results was combined with the survey questions. A separate file was created for the 2010 and the 2020 papers. Each Excel file was shared with the survey participants. The 2010 papers were randomly distributed between 14 research groups and the 2020 papers were randomly distributed between 17 research groups to manually complete the survey questions. Survey participants read each assigned paper and any available supplemental information to manually answer all survey questions. A master 2010 (Table B2) and 2020 (Table B3) Excel file was created by concatenating the individual literature survey results compiled by each research group. The master Excel files were then used to tabulate a summary of the 2010 (Table B4) and 2020 (Table B5) results for each individual question using standard Excel COUNT and SUM functions such as COUNTA, COUNTBLANK, and COUNTIF. For Yes-No questions, the number of “Yes”, “No”, “Not Indicated”, and blank cells were separately counted. For questions with a defined list of answers, the number of each defined answer, each generic “other”, and each blank cell was separately counted. Open ended questions were manually cataloged and counted. The two Excel files that summarize the results of the literature survey are available as supplemental information.

References

- [1] M.R. Viant, *Mol. Biosyst.* 4 (2008) 980.
- [2] L. Brennan, *Prog. Nucl. Magn. Reson. Spectrosc.* 83 (2014) 42.
- [3] R. Powers, *Magn. Reson. Chem.* 47 (2009) S2.
- [4] R. Powers, *J. Med. Chem.* 57 (2014) 5860.
- [5] D.S. Wishart, *Drugs R 9* (2008) 307.
- [6] I.F. Duarte, S.O. Diaz, A.M. Gil, *J. Pharm. Biomed. Anal.* 93 (2014) 17.
- [7] A.-H.M. Emwas, R.M. Salek, J.L. Griffin, J. Merzaban, *Metabolomics* 9 (2013) 1048.
- [8] T. Gebregiorgis, R. Powers, *Comb. Chem. High Throughput Screening* 15 (2012) 595.
- [9] W.J. Griffiths, T. Koal, Y. Wang, M. Kohl, D.P. Enot, H.-P. Deigner, *Angew. Chem., Int. Ed.* 49 (2010) 5426.
- [10] D.S. Wishart, L.L. Cheng, V. Copié, A.S. Edison, H.R. Eghbalian, J.C. Hoch, G. J. Gouveia, W. Pathmasiri, R. Powers, T.B. Schock, L.W. Sumner, M. Uchimiyama, *Metabolites* 12 (2022) 678.
- [11] J.L. Markley, R. Bruschweiler, A.S. Edison, H.R. Eghbalian, R. Powers, D. Raftery, *D.S. Wishart, Curr. Opin. Biotechnol.* 43 (2017) 34.
- [12] T.A. Schult, M.J. Lauer, Y. Berker, M.R. Cardoso, L.A. Vandergrift, P. Habel, J. Nowak, M. Taupitz, M. Aryee, M.A. Mino-Kenudson, D.C. Christiani, L.L. Cheng, *Proc. Natl. Acad. Sci. U.S.A.* 118 (2021).
- [13] P.F. Daly, J.S. Cohen, *Cancer Res.* 49 (1989) 770.
- [14] M. Bastawrous, A. Jenne, M.T. Anaraki, A.J. Simpson, *Metabolites* 8 (2018), 35/1.
- [15] M.T. Anaraki, D. Lane, M. Bastawrous, A. Jenne, A.J. Simpson, *Methods Mol. Biol.* 2037 (2019) 395.
- [16] G.A. Nagana Gowda, D. Raftery, *Adv. Exp. Med. Biol.* 1280 (2021) 19.
- [17] C.L. Wu, K.W. Jordan, E.M. Ratai, J. Sheng, C.B. Adkins, E.M. Defeo, B.G. Jenkins, L. Ying, W.S. McDougal, L.L. Cheng, *Sci. Transl. Med.* 2 (2010) 16ra8.
- [18] J.K. Nicholson, J.C. Lindon, E. Holmes, *Xenobiotica* 29 (1999) 1181.
- [19] T.W.M. Fan, *Prog. Nucl. Magn. Reson. Spectrosc.* 28 (1996) 161.
- [20] D.S. Wishart, *Trends Anal. Chem.* 27 (2008) 228.
- [21] A.S. Maroli, R. Powers, *NMR Biomed.* e4594.
- [22] A.A. Crook, R. Powers, *Molecules* 25 (2020) 5128.

- [23] P.G. Takis, V. Ghini, L. Tenori, P. Turano, C. Luchinat, *Trends Anal. Chem.* 120 (2019) 115300.
- [24] R.G. Brereton, *J. Chemom.* 28 (2014) 749.
- [25] D.L. Vaux, *Nature* 492 (2012) 180.
- [26] S. Mutter, C. Worden, K. Paxton, V.-P. Mäkinen, *Metabolomics* 16 (2019) 5.
- [27] K. Kjeldahl, R. Bro, *J. Chemometr.* 24 (2010) 558.
- [28] R.G. Brereton, *TrAC, Trends Anal. Chem.* 25 (2006) 1103.
- [29] H.R. Eghbalnia, P.R. Romero, W.M. Westler, K. Baskaran, E.L. Ulrich, J.L. Markley, *Curr. Opin. Biotechnol.* 43 (2017) 56.
- [30] S.M. Mische, N.C. Fisher, S.M. Meyn, K. Sol-Church, R.L. Hegstad-Davies, F. Weis-Garcia, M. Adams, J.M. Ashton, K.M. Delventhal, J.A. Dragon, L. Holmes, P. Jagtap, K.E. Kubow, C.E. Mason, M. Palmbalad, B.C. Searle, C.W. Turck, K. L. Knudtson, *J. Biomol. Tech.* 31 (2020) 11.
- [31] J.A. Kirwan, H. Gika, R.D. Beger, D. Bearden, W.B. Dunn, R. Goodacre, G. Theodoridis, M. Witting, L.R. Yu, I.D. Wilson, *Metabolomics* 18 (2022) 70.
- [32] H.E. Roth, R. Powers, *Cancers* 14 (2022) 3992.
- [33] R. Goodacre, D. Broadhurst, A.K. Smilde, B.S. Kristal, J.D. Baker, R. Beger, C. Bessant, S. Connor, G. Capuani, A. Craig, T. Ebbels, D.B. Kell, C. Manetti, J. Newton, G. Paternostro, R. Somorjai, M. Sjöström, J. Trygg, F. Wulfer, *Metabolomics* 3 (2007) 231.
- [34] J.C. Lindon, J.K. Nicholson, E. Holmes, H.C. Keun, A. Craig, J.T. Pearce, S.J. Bruce, N. Hardy, S.A. Sansone, H. Antti, P. Jonsson, C. Daykin, M. Navarange, R.D. Beger, E.R. Verheij, A. Amberg, D. Baunggaard, G.H. Cantor, L. Lehman-McKeeman, M. Earll, S. Wold, E. Johansson, J.N. Haselden, K. Kramer, C. Thomas, J. Lindberg, I. Schuppe-Koistinen, I.D. Wilson, M.D. Reilly, D.G. Robertson, H. Senn, A. Krotzky, S. Kochhar, J. Powell, F. van der Ouderaa, R. Plumb, H. Schaefer, M. Spraul, *Nat. Biotechnol.* 23 (2005) 833.
- [35] S.-A. Sansone, T. Fan, R. Goodacre, J.L. Griffin, N.W. Hardy, R. Kaddurah-Daouk, B.S. Kristal, J. Lindon, P. Mendes, N. Morrison, B. Nikolau, D. Robertson, L. W. Sumner, C. Taylor, M. van der Werf, B. van Ommen, O. Fiehn, M.S.I. B. Members, *Nat. Biotechnol.* 25 (2007) 846.
- [36] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A.N. Lane, J.C. Lindon, P. Marriott, A.W. Nicholls, M.D. Reilly, J. J. Thaden, M.R. Viant, *Metabolomics* 3 (2007) 211.
- [37] D.V. Rubtsov, H. Jenkins, C. Ludwig, J. Easton, M.R. Viant, U. Gunther, J.L. Griffin, N. Hardy, *Metabolomics* 3 (2007) 223.
- [38] R.A. Spicer, R. Salek, C. Steinbeck, *Sci. Data* 4 (2017), 170138.
- [39] O. Fiehn, D. Robertson, J. Griffin, M. van der Werf, B. Nikolau, N. Morrison, L. W. Sumner, R. Goodacre, N.W. Hardy, C. Taylor, J. Postel, B. Kristal, R. Kaddurah-Daouk, P. Mendes, B. van Ommen, J.C. Lindon, S.-A. Sansone, *Metabolomics* 3 (2007) 175.
- [40] S. Goodman, *Semin. Hematol.* 45 (2008) 135.
- [41] B. Worley, R. Powers, *Curr. Metabolomics* 1 (2013) 92.
- [42] Y. Benjamini, Y. Hochberg, *J. R. Stat. Soc. Ser. B Methodol.* 57 (1995) 289.
- [43] J. Neyman, E.S. Pearson, *Biometrika* 20A (1928) 175.
- [44] R.A. Armstrong, *Ophthalmic Physiol. Opt.* 34 (2014) 502.
- [45] J.L. Markley, R. Bruschweiler, A.S. Edison, H.R. Eghbalnia, R. Powers, D. Raftery, D.S. Wishart, *Curr. Opin. Biotechnol.* 43 (2017) 34.
- [46] K. Haug, R.M. Salek, P. Conesa, J. Hastings, P. de Matos, M. Rijnbeek, T. Mahendrakar, M. Williams, S. Neumann, P. Rocca-Serra, E. Maguire, A. González-Beltrán, S.A. Sansone, J.L. Griffin, C. Steinbeck, *Nucleic Acids Res.* 41 (2013) D781.
- [47] F. Bhinderwala, T. Vu, T.G. Smith, J. Kosacki, D.D. Marshall, Y. Xu, M. Morton, R. Powers, *Anal. Chem.* 94 (2022), 16308.
- [48] R.M. Borges, G. de Assis Ferreira, M.M. Campos, A.M. Teixeira, F. das Neves Costa, F.O. das Chagas, M.B. Colonna, *Chem* (2023) 1.
- [49] S.L. Robinette, F. Zhang, L. Bruschweiler-Li, R. Bruschweiler, *Anal. Chem.* (Washington, DC, U. S.) 80 (2008) 3606.
- [50] C. Wang, B. Zhang, I. Timari, A. Somogyi, D.-W. Li, H.E. Adcox, J.S. Gunn, L. Bruschweiler-Li, R. Bruschweiler, *Anal. Chem.* (Washington, DC, U. S.) 91 (2019), 15686.
- [51] S. Ravanbakhsh, P. Liu, T.C. Bjordahl, R. Mandal, J.R. Grant, M. Wilson, R. Eisner, I. Sinelnikov, X. Hu, C. Luchinat, R. Greiner, D.S. Wishart, *PLoS One* 10 (2015), e0124219.
- [52] A. Foroutan, C. Fitzsimmons, R. Mandal, M.V. Berjanskii, D.S. Wishart, *Metabolites* 10 (2020).
- [53] T. Vu, Y. Xu, Y. Qiu, R. Powers, *Biostatistics* 24 (2023) 140.
- [54] E. Saccenti, H.C.J. Hoefsloot, A.K. Smilde, J.A. Westerhuis, M.M.W.B. Hendriks, *Metabolomics* 10 (2014) 361.
- [55] B. Percival, M. Gibson, J. Leenders, P.B. Wilson, M. Grootveld, in: P.B. Wilson, M. Grootveld (Eds.), *Univariate and Multivariate Statistical Approaches to the Analysis and Interpretation of NMR-Based Metabolomics Datasets of Increasing Complexity*, The Royal Society of Chemistry, 2020, 0.
- [56] Y. Benjamini, Y. Hochberg, *J. Roy. Stat. Soc. B* 57 (1995) 289.
- [57] Y. Benjamini, D. Yekutieli, *Ann. Stat.* 29 (2001) 1165.
- [58] J.D. Storey, *J. Roy. Stat. Soc. B Stat. Methodol.* 64 (2002) 479.
- [59] L. Eriksson, J. Trygg, S. Wold, *J. Chemom.* 22 (2008) 594.
- [60] F. Lindgren, B. Hansen, W. Karcher, M. Sjöstrom, L. Eriksson, *J. Chemom.* 10 (1996) 521.
- [61] O. Begou, H.G. Gika, G.A. Theodoridis, I.D. Wilson, *Methods Mol. Biol.* 1738 (2018) 15.
- [62] A. Craig, O. Cloarec, E. Holmes, J.K. Nicholson, J.C. Lindon, *Anal. Chem.* 78 (2006) 2262.
- [63] T. Vu, E. Rieckeberg, Y. Qiu, R. Powers, *Metabolomics* 14 (2018) 108.
- [64] Anonymous, *Nature* 533 (2016) 437.
- [65] A.-H. Emwas, C. Luchinat, P. Turano, L. Tenori, R. Roy, R.M. Salek, D. Ryan, J. S. Merzaban, R. Kaddurah-Daouk, A.C. Zeri, G.A. Nagana Gowda, D. Raftery, Y. Wang, L. Brennan, D.S. Wishart, *Metabolomics* 11 (2015) 872.
- [66] O.A. Snytnikova, A.A. Khlichkina, R.Z. Sagdeev, Y.P. Tsentlovich, *Metabolomics* 15 (2019) 1.
- [67] L. Banci, L. Barbieri, V. Calderone, F. Cantini, L. Cerofolini, S. Ciofi-Baffoni, I. C. Felli, M. Fragai, M. Lelli, C. Luchinat, E. Luchinat, G. Parigi, M. Piccioli, R. Pierattelli, E. Ravera, A. Rosato, L. Tenori, P. Turano, *e-Print Arch., Phys.* (2019) 1, arXiv.org.
- [68] P. Alonso-Moreno, I. Rodriguez, J.L. Izquierdo-Garcia, *Metabolites* 13 (2023) 614.
- [69] O. Beckonert, H.C. Keun, T.M.D. Ebbels, J. Bundy, E. Holmes, J.C. Lindon, J. K. Nicholson, *Nat. Protoc.* 2 (2007) 2692.
- [70] J.K. Nicholson, P.J.D. Foxall, M. Spraul, R.D. Farrant, J.C. Lindon, *Anal. Chem.* 67 (1995) 793.
- [71] T.L. Hwang, A.J. Shaka, *J. Magn. Reson., Ser. A* 112 (1995) 275.
- [72] A.C. Dona, B. Jiménez, H. Schäfer, E. Humpfer, M. Spraul, M.R. Lewis, J.T. M. Pearce, E. Holmes, J.C. Lindon, J.K. Nicholson, *Anal. Chem.* 86 (2014) 9887.
- [73] A.J. Simpson, S.A. Brown, *J. Magn. Reson.* 175 (2005) 340.
- [74] A. Le Guennec, F. Tayyari, A.S. Edison, *Anal. Chem.* 89 (2017) 8582.
- [75] J.J.J. van der Hoof, N. Rankin, in: G.A. Webb (Ed.), *Metabolite Identification in Complex Mixtures Using Nuclear Magnetic Resonance Spectroscopy*, Springer International Publishing, Cham, 2018, p. 1309.
- [76] C. Ludwig, M.R. Viant, *Phytochem. Anal.* 21 (2010) 22.
- [77] R.B. Dessau, C.B. Pipper, *Ugeskr Laeger* 170 (2008) 328.
- [78] M.W. Maciejewski, A.D. Schuyler, M.R. Gryk, I.I. Moraru, P.R. Romero, E.L. Ulrich, H.R. Eghbalnia, M. Livny, F. Delaglio, J.C. Hoch, *Biophys. J.* 112 (2017) 1529.
- [79] Z. Pang, J. Chong, G. Zhou, D. Anderson de Lima Morais, L. Chang, M. Barrette, C. Gauthier, P.-E. Jacques, S. Li, J. Xia, *Nucleic Acids Res.* 49 (2021) W388.
- [80] B. Worley, R. Powers, *ACS Chem. Biol.* 9 (2014) 1138.
- [81] S. Alseekh, A. Aharoni, Y. Brotman, K. Contrepolis, J. D'Auria, J. Ewald, J.C. Ewald, P.D. Fraser, P. Giavalisco, R.D. Hall, M. Heinemann, H. Link, J. Luo, S. Neumann, J. Nielsen, L. Perez de Souza, K. Saito, U. Sauer, F.C. Schroeder, S. Schuster, G. Siuzdak, A. Skirycz, L.W. Sumner, M.P. Snyder, H. Tang, T. Tohge, Y. Wang, W. Wen, S. Wu, G. Xu, N. Zamboni, A.R. Fernie, *Nat. Methods* 18 (2021) 747.
- [82] T. Vu, P. Siemek, F. Bhinderwala, Y. Xu, R. Powers, *J. Proteome Res.* 18 (2019) 3282.
- [83] F. Bhinderwala, R. Powers, F. Bhinderwala, R. Powers, *Methods Mol. Biol.* 2037 (2019) 265.
- [84] F. Bhinderwala, S. Lei, J. Woods, J. Rose, D.D. Marshall, E. Rieckeberg, A. De Lima Leite, M. Morton, E.D. Dodds, R. Franco, R. Powers, *Methods Mol. Biol.* 1996 (2019) 217.
- [85] Z. Lu, *Database* (2011) 1.
- [86] N. Fiorini, D.J. Lipman, Z. Lu, *Elife* 6 (2017).
- [87] M.E. Falagas, E.I. Pitsouni, G.A. Malietzis, G. Pappas, *Faseb. J.* 22 (2008) 338.