DISCUSSION ARTICLE

A C G publications *J. Chem. Metrol*. 17:1 (2023) 114-127

journal of chemical metrology

Accuracy of quantitative NMR analysis: a case study of lignin

Shrikant Shivaji Pawade*, Lauri Toom, Koit Herodes and Ivo Leito

University of Tartu, Institute of Chemistry, Ravila 14a, 50411 Tartu, Estonia

(Received April 23, 2023; Revised May 16, 2023; Accepted May 20, 2023)

Abstract: Lignin analysis using quantitative NMR (qNMR) has received a lot of attention recently and has been the topic of a large number of papers. The large majority of them report high-quality research. However, when trying to understand what is the accuracy that can be achieved with such analysis and which of the approaches are more accurate, it turns out that it is difficult to impossible to compare the accuracy of the methods. The main reasons are that different authors use (1) different types of lignin, (2) different measurands and (3) different ways of presenting precision and trueness data. Precision is mostly presented as standard deviation between replicate measurement results but it is in most cases not specified whether the precision relates to repeatability, intermediate precision or some other precision type. Bias is typically termed as "error" and usually expressed as difference from a reference value obtained from an artificial model system or difference from results of independent measurements. Again, insufficient detail is often given. Accuracy in terms of "measurement uncertainty" is hardly ever presented. Some uncertainty sources, most notably variability between subsamples of the same bulk of lignin, are only seldom addressed. We present an analysis of the situation on the basis of 21 papers and give some recommendations for future workers in the field. We hope that this work will be useful for researchers using qNMR for the analysis of lignin, as well as using qNMR for the analysis of natural products more generally.

Keywords: Lignin; precision; bias; accuracy; NMR, qNMR.

1. Introduction

Lignin is the second (after cellulose) most abundant natural organic material on Earth [1]. Huge amounts of lignin are produced (approximately 50 million tons annually) as a byproduct of the cellulose industry [2]. Given its molecular structure – sophisticated phenylpropane-derived polymeric network – lignin could be valorized into a range of useful chemicals. Although numerous groups work in this field and quite some progress has been achieved [3], currently still, 98-99% of the produced lignin is simply burned for energy, instead of conversion into value-added chemicals [2].

In the context of lignin valorization, quantitative evaluation of its composition is of high importance. Numerous studies are either fully devoted to quantitative analysis of lignin [4–7] or involve it as an important part. There are a number of functional groups and structural fragments that are typically quantified when quantitative lignin analysis is carried out: methoxyl: aryl ratio [4], syringyl: guaiacyl ratio [4], hydroxyl groups [5], the content of β-O-4, β-5, β-β units, etc. [8]. There are also a number of different measurands [9] – ratios of contents of different fragments, number of certain fragments per aromatic ring, moles of certain fragments per 100 g of lignin, etc. [10]. These two dimensions of diversity

* Corresponding author E-Mail: shrikant.shivaji.pawade@ut.ee

The article was published by ACG Publications
http://www.acgpubs.org/journal/journal-of-chemical-metrology January-June 2023 EISSN:1307-6133
DOI: http://doi.org/10.25135/rnp.88.2304.2753
Available online: May 30, 2023

– different structural fragments and different measurands – result in a large number of combinations, complicating the comparison of results between different reports. The diversity is made even bigger by the number of different technologies that are used for producing lignin: Kraft lignin, lignosulfonates, soda lignin, and organosolv lignin [11,12].

In the course of our recent work, we have become interested in the quantitative analysis of lignin, in terms of abundance of different structural units/linkages/functional groups. Our interest was what techniques are available and what accuracy can be achieved.

It seems clear from the analysis of the literature that quantitative NMR (qNMR) is the most important analytical tool for both qualitative and quantitative structural analysis of lignin. Out of the different NMR approaches, the ^{13}C and ^{31}P (after phosphitylation [7]) NMR, as well as different two-dimensional approaches (e.g., HSQC), are the most useful [5]. qNMR is the use of NMR spectroscopy to determine the concentration or purity of one or more chemical species or structural fragments in samples. The basic principle of qNMR relies on the area of an NMR signal being directly proportional to the concentration of the particular analyte in the full concentration range [13,14]. qNMR is well applicable to the quantitative analysis of natural products due to the intrinsically quantitative nature of NMR, the fact that it doesn't have any concerns with the sample impurities when the impurity peaks aren't interfering with sample peaks or when the impurities possess no peaks, it also does not require standard reference materials and the whole process does not destroy to the sample [15].

As for the possible accuracy/reliability of results, it turns out that the situation is as diverse as with measurands. Different authors present accuracy-related data – precision and trueness/bias – in very different ways. In many cases, only precision is presented, often termed "reproducibility" or simply "RSD." Other authors evaluate the "errors" of their results by comparing them to different types of reference values of model compounds or model mixtures. The high diversity of ways of presenting results and their accuracy complicates comparisons of results and accuracies from different authors.

Thus, in this paper, we intend to present an analysis of the situation with the accuracy (see below for explanations) of quantitative NMR analysis of lignin on the basis of available literature, draw some conclusions and give some recommendations.

2. Literature Survey

The literature search was carried out targeting reports that focused specifically on qNMR analysis of lignin (as opposed to focusing on some technological problem and simply using qNMR as a tool). When choosing the reports to be included, we were selective rather than exhaustive. Thus, out of few hundreds of possible papers, we targeted those reports that specifically addressed the accuracy of qNMR analysis of lignin, ending up with a selection of 21 papers. We categorized the reports into two different types:

(Q) Reports addressing quantification of lignin components by qNMR, e.g., presenting new quantitative analysis methods.
(A) Reports addressing specifically the accuracy of qNMR quantification.

The term accuracy refers to the agreement of a measured quantity value with the true quantity value of a measurand (i.e. the quantity that is intended to be measured) [9]. Although mostly interpreted as qualitative concept, it can also be viewed that accuracy can be numerically expressed by measurement uncertainty [16]. Measurement uncertainty defines an interval around the measured value where with high (and predefined) probability the true value of the measurand can be found [17]. Measurement result is influenced by systematic and random effects that cause deviations of the measured value from the true value and thereby the existence of measurement uncertainty. Accuracy can be regarded as relating to both random and systematic effects (random and systematic errors) influencing the analysis result, whereby the random effects are characterized by precision (expressed as standard deviation) and systematic effects by trueness (expressed as bias) [16,17]. Thus, we have attempted to divide the different types of information given by authors about accuracy into two categories:

- Precision, which quantifies random effects, often expressed a relative standard deviation (RSD) [9,18,19], and
- Bias (often termed as “error”), which quantifies systematic effects [9,18,19].

Determining precision relies on replicate measurements with real-life samples [20]. This can almost always be done (although it can be work-intensive for measurements where long signal-collecting times are needed). In contrast, in order to estimate bias, reference samples are necessary that (1) have reliable (i.e., low uncertainty) reference values and (2) are sufficiently similar to real-life samples. Obtaining suitable reference samples is not trivial. For that reason, bias estimates are more difficult to make. Moreover, in spite of seeming simplicity, both of these concepts are nuanced and for fully understanding what is meant by a certain precision or bias estimate, more information is needed. As discussed below, this information is often missing.

The results of the literature survey are presented in Table 1.

3. Discussion

3.1. Uncertainty Sources

A number of uncertainty sources have been highlighted by the authors of the papers. There are obvious ones, such as the accuracy of integration [5], including overlapping spectral features of different structural fragments, irregularities in baseline [10], noise in the spectrum [10], etc. There are uncertainty sources that are specific to HSQC: deviations in coupling constant, resonance offset effects, effects of ^{1}H T_1 relaxation, effects of proton and X-nuclei T_2 relaxation, and effects of proton homonuclear coupling [30]. However, there are some uncertainty sources, which, although mentioned in some reports, deserve more attention. They are:

(1) Sampling and subsampling of lignin for measurement. Lignin is a solid and reasonably homogeneous, but still, different subsamples taken from the same bulk for the analysis might have somewhat different compositions, depending on mixing and on the amount taken. This variability represents an uncertainty source. In most of the works, replicate measurements were made, but in many of the papers, it is not clear whether every replicate measurement was carried from an independent subsample or from the same subsample. This uncertainty source is only seldom addressed, and it appears that the precision data in most reports are given without counting sample preparation. There are some exceptions [10,25]. The data from [25] enable back-calculating the RSD of sample preparation, and it is 5% (^{13}C NMR) and 8% (^{31}P NMR). These RSD values are larger (!) than the RSD related to NMR measurement (3% and 5%, respectively), indicating the high importance of sample preparation as uncertainty source. The higher RSD in the case of ^{31}P NMR is probably due to the need for derivatization procedure, differently from ^{13}C NMR.

(2) Lignin usually contains some amounts (a few percent) of carbohydrates (cellulose, hemicellulose) [23]. The signals from these carbohydrates overlap with signals from aliphatic moieties of lignin and thereby lead to increased estimates of some the content of aliphatic moieties.

(3) Technical lignin often contains degraded side chains or broken aromatic rings. If these degradation products are not included into the modeling or internal standardizing, they act as overlooked impurities [36] and can lead to significant errors [10,27] and should be considered as an uncertainty source.

Table 1. Chronological overview of selected reports addressing the accuracy of quantitative NMR analysis of lignin.[a]

Sample type	Quantified fragment or measurand[a]	Method, amount of material used	1st author, Year, Type[a] and Ref	Precision[b]	Bias[c]	Comments, contributions of specific uncertainty sources[d]
Milled wood lignin (MWL), acetylated milled wood lignin (MWL)	Methoxy, acetate group, aromatic group, carbonyl group, different peak ratios	^{13}C NMR small samples 70 mg, large samples 700 mg	Landucci 1985 A [21]	Precision: ±3%…±10%, RSD: ±1.3%…±8.7%	–	Only precision was characterized, no attempt was made to evaluate bias. The ^{13}C NMR of acetylated milled wood lignin can be measured with a precision of ±3% for large samples (700 mg) which was measured for 6 hours runs whereas when the samples were run for 1 hour then the precision was ±10%. When the sample was small (70 mg) run 3 times then the precision for 8000 pulses was ±10% for 22 hours run.
Softwood and hardwood lignins	Ratios methoxyl : aryl and syringyl : guaiacyl (S:G)	^{13}C NMR	Obst 1986 Q [4]	–	"Error" in the range of ±0.01…±0.07	The uncertainty (termed as "error") of the methoxyl : aryl ratio was in the range of ±0.01 to ±0.07 (the ratios ranged from 0.95 to 1.65. The uncertainty of S:G ratio is 6 times higher, i.e. ranges from ±0.06 to ±0.42 (S:G ratios ranged from 0.29 to 1.85). Neither the meaning nor the source of the "error" estimate is explained.
Tree pine kraft lignin samples, differently prepared and milled wood lignin (MWL)	Number of various functional groups (different types of OH, CH, etc.) per aromatic ring	^{13}C NMR, DEPT, 500-600 mg	Gellerstedt 1987 Q [22]	–	±5%	"Error" limit was estimated of the order of ±5%. No description is given how the "error" estimate was obtained.
Milled wood lignin	Number of carbons (quaternary, tertiary, methoxy, etc) per aromatic ring	^{1}H and ^{13}C NMR, DEPT 35 mg AcMWL 300-400 mg	Chen 1988 Q [23]	–	±5%	"Error" estimated as ±5%, which is claimed to be due to less than ideal proportionality between the signal intensity and the number of the respective ^{13}C nuclei. No details are given about how the error was estimated.

Variety of lignins: kraft, Indulin mixed softwood, sucrolin; bagasse, steam explosion; aspen Alcell-organosolv mixed hardwoods, steam explosion yellow poplar; ball milled enzyme cottonwood	Number of total and phenolic OH groups per 100 C_9 units	^{1}H NMR, ^{13}C NMR, ^{31}P NMR	Faix 1994 A [24]	^{13}C NMR: 5%	Total OH by ^{1}H NMR: -15%…13% ^{31}P NMR: -22%...3% Total phenolic OH by ^{1}H NMR: -30%...14% ^{31}P NMR: -23%...10%	The precision was termed as “relative reproducibility”, but no details are given about how the replicate measurements were carried out. The bias (number of replicates unknown) is found against a “wet chemical method”. The total OH group concentration across all lignin’s was 120/C_{900}, with an average standard deviation of 8 (excluding one ^{31}P NMR measurement).
Two hardwood (*Populus tremuloides*) and one softwood (mixture of softwoods) lignin	A variety of hydroxyls: CH_2OH, secondary-OH, phenolic-OH, carboxylic, syringyl (S), guaiacyl (G), p-hydroxyphenyl (H), β-O-4 hydroxyls, expressed as mol per mol of C_9 units	Phosphitylation and ^{31}P NMR 30-60 mg	Argyropoulos 1994 Q [7]	RSD: 1.5%…2.1%	–	Definitive paper on the Phosphitylation ^{31}P NMR approach for determining different hydroxyls. The "standard error” was interpreted by the current authors as standard deviation of the mean. Number of replicates: 30. No information is given about how the replicate measurements were carried out.
Residual lignin after kraft pulping	Aliphatic hydroxyl content (mmol/g), phenolic hydroxyl content (mmol/g), number of C^{γ} per β-O-4 aromatic unit, etc.	^{13}C NMR, ^{31}P NMR 300-400 mg	Froass 1998 Q [25]	Repeatability RSD: with the same subsample: 3% (^{13}C) and 5% (^{31}P); with different subsamples: 6% (^{13}C) and 10% (^{31}P)	–	Repeatability RSD for the same subsample of lignin (i.e. sample preparation was carried out just once) was 3% and 5%. When different subsamples of the same lignin (each with separate sample preparation) were analyzed, the repeatability RSD was 6% and 10%.

Softwood kraft residual lignin,	Hydroxy functional groups (mmol/g)	^{31}P NMR (30 mg lignin or ~30 µmol model compounds)	Zawadzki 2001 Q [26]	RSD: 0.5%…2%	–	Two different internal standards were used (cyclohexanol and *N*-hydroxy-5-norbornene-2,3-dicarboximide). No systematic difference was found between RSD values. The authors termed the standard deviations as "reproducibility" but there is insufficient detail to judge whether it is reproducibility or repeatability.
Indulin kraft lignin	Amounts of structural units related to spectral areas (ArC-1, ArC-2, etc) mmol/g	^{13}C NMR, DEPT ~200 mg/mL	Xia 2001 Q [27]	RSD 0.8%...8.4%	–	Extensive precision data (expressed as RSD) are presented for different areas of spectral integrals. The RSD depends on the spectral region. Information on the conditions of those replicate measurements is limited. Dangers of using aromatic methoxy signals as internal standards in the case of technical lignins are highlighted.
Spruce lignin and mixture of model compounds	Ratios of different structural units: β-O-4, β-5, β-β, dibenzodioxocin, OCH_3	^{1}H-^{13}C Q-HSQC NMR with apocynol as IS 100 mg	Heikkinen 2003 Q [28]	RSD 0.8%...3%	0%...14%	"Reproducibility" RSD ranged from 0.8% to 3% and was evaluated by repeating measurements 4 times. Although termed "reproducibility" the evaluated precision is more likely repeatability. The possible systematic effect was estimated via comparison with values calculated from the composition of a model compound mixture.

Nonacetylated and acetylated spruce MWL	Number of structural units (β-O-4, β-5, β-β, etc) per 100 aromatic rings	^{1}H-^{13}C HMQC NMR, quantitative ^{13}C NMR 200-600 mg, with Shigemi microtube 50-70 mg	Capanema 2004 Q [29]	Accuracy of integration 2%	Accuracy of integration 2%	The authors state that they give a "comprehensive and rather reliable picture of the structure of lignin, compatible with the whole set of other methods in lignin chemistry". There is indeed a useful comparison table of quantitative results obtained with different model approaches. The accuracy of integration in the case of well resolved peaks and clusters such as hydroxyl, methoxy, aromatic as well as oxygenated aliphatic carbon was assessed to ±2% at 2000 scans. It is not clear if this refers to only random effects or both random and systematic effects. The reliability of these estimates is considerably improved by using several independent approaches to quantify various lignin moieties. However, overall uncertainties are not presented.
Eucalyptus grandis milled wood lignin (MWL)	Number of structural units (β-O-4, β-5, β-β, etc) per aromatic ring; ratio of H, G and S units per aromatic ring	2D ^{1}H-^{13}C HSQC, HMQC, ^{1}H-^{1}H TOCSY NMR, quantitative ^{13}C NMR Shigemi microtube 60-70 mg	Capanema 2005 Q [6]	Integration error 3%	Integration error 3%	Integration "error" is estimated as 3%. It is not clear if this refers to only random effects or both random and systematic effects. The article highlights other uncertainty sources, such as uncertainty due to incomplete resolution of signals or uncertainty due to assumptions in model equations that slightly deviate from reality, but no quantitative estimates are given.
Spruce milled wood lignin,	Clusters of signals, which were different lignin methine (CH) groups, number per C_9 units	^{13}C NMR, DEPT90 and HSQC NMR 100mg	Zhang 2007 Q [30]	–	"Error": 0% (^{13}C NMR), -26 (DEPT90) and -35% (HSQC)	All "errors" are integration errors. The errors in quantitative HSQC NMR are claimed to be mainly due to five reasons: deviations in coupling constant, resonance offset effects, effects of ^{1}H T_1 relaxation, effects of proton and X-nuclei T_2 relaxation, and effects of proton homonuclear coupling. 1,3,5-trimethoxybenzene was used as IS.

Aspen MWL, Alcell, Indulin and SEAL lignins	OMe, total OH, aliphatic (total) OH, phenolic OH, S, G, S/G ratio, ArH, and oxygenated aliphatic moieties, total β-O-4, COOR, CO and EtO-groups, β-β, β-5, H-units (in wood-originated lignins), Alk-O-Alk, degree of demethylation. Expressed in ppm per 100 Ar	^{13}C NMR 190–210 mg 350–400 mg	Balakshin 2015 A [10]	RSD: highly accurate quantification <3%, with moderate accuracy 3%...10%, semi-quantitative >10%	–	The accuracy is lower in quantification (aliphatic OH (primary and secondary ones), β-O-4 units and other oxygenated aliphatic moieties) due to the degradation in moieties during processing and lesser abundance in technical lignins in comparison to MWLs. Replicate measurements (no information, whether short- or long-term) were carried out including sample preparation for NMR, acquisition of NMR and NMR processing. The deviation in NMR analysis shows RSD ranged for AWML in 0.2 to 32.1%, for Alcell lignin 0.9 to 15.0%, for Indulin it was 0.6 to 23.6% whereas for the SEAL lignin it was 0.3 to 61.0%. The paper gives a number of useful recommendations and a good overview of uncertainty sources. The importance of good baseline and appropriate S/N ratio is stressed.
A number of lignins: aspen milled wood (MWL) lignin, pine MWL, aspen and birch dioxane lignins (ADL and BDL); pine and aspen kraft lignin (PKL), pine soda lignin (PSL), Indulin lignin and Alcell lignin	Content of hydroxyl groups (aliphatic, phenolic, total). ^{13}C NMR: mol%; ^{31}P NMR: mmol/g	^{31}P NMR, ^{13}C NMR 150–200 mg	Balakshin 2015 A [5]	RSD 2%...3% (for both ^{13}C and ^{31}P NMR)	Typical differences between ^{31}P NMR and ^{13}C NMR: 10%…15%	Bias has been estimated from comparisons between ^{31}P and ^{13}C NMR results. The ^{31}P NMR method shows different precision (no information, whether short- or long-term) for different types of hydroxyl groups. The authors also present between-lab reproducibility data for measurements of “similar” samples and find RSD in the range of 26% to 65% depending on structural feature. A number of uncertainty sources (termed as such in the paper) are analyzed and the influence of some of them quantitatively estimated: sample preparation, extraction or isolation of lignin (soluble lignin, sample pretreatment, limited stability of derivatized samples. The paper gives a number of useful recommendations.

Milled wood lignin, HW aMWLs,	Carbonyl (CO) group, OMe, OH, S, G, ArH, total β-O-4, oxygenated aliphatic moieties, and syringyl-to-guaiacyl (S/G) ratio OMe and OH-groups, β-O-4/α-OH, β-β, β-5, COOR and conjugated CO groups (including Ar-CHO), H units, 5-substituted and 5-free OH_{phen} mmol/g	^{13}C NMR 200 mg	Balakshin 2016 A [8]	RSD: 10%…20% for broad and/or minor signals; 0.5%...5% for major and well-defined signals	–	RSD refers to between-instrument reproducibility. In the quantification of wide and/or minor signals, such as carbonyl (CO) groups in general and Aromatic-CHO and spirodienone structures in particular, or COOR and H units, the relative SD (RSD) was substantially higher, in the range of 10 to 20%. Major and well-defined signals, like OMe, ArH, S and G units and clusters in the oxygenated aliphatic regions, have substantially better reproducibility (RSD = 0.5–5%).
Indulin kraft, soda P1000, Alcell, OS-W, OS-P, OS-S	Aliphatic OH, 5-substituted OH, guaiacyl OH, p-hydroxyphenyl OH, total PhOH, COOH, free COOH/tricin in mmol/g; different side chains and aromatic units (S, G & H) in number per 100 aromatic units	HSQC NMR, ^{31}P NMR ^{31}P NMR 40 mg HSQC 200 mg	Constant 2016 Q [31]	Between-lab RSD: ^{31}P NMR 4% to more than 100%; Q-HSQC 0.4% to close to 100%	Error -12%	Replicate analyses conducted in two different laboratories reveal discrepancies of up to around two times. Comparison with the values of model compounds shows bias of-12%, termed as “error”.
A range of lignins with different ether linkage contents (formaldehyde- and propionaldehyde-stabilized, mild dilute acid-catalyzed, organosolv lignin)	Content of β-O-4 linkages, which enable evaluating depolymerization yields (depolymerization mainly proceeds via cleaving these linkages), mmol	Quantitative ^{1}H-^{13}C HSQC NMR, $gsHSQC_0$	Talebi 2019 Q [32]	–	Relative “error” -9.4%…+13.4%	The predicted quantities of different units had relative deviations ranging from -9.4% to +13.4% from values of a synthetic model polymer. A range of NMR-related uncertainty sources are identified (resonance offsets, different T_2 relaxation between different parts of the biopolymer, imperfect pulses, homonuclear coupling, and coupling constant deviations) that are largely the same as in ref [30].

Raspberry dioxane lignin, spruce, birch, wheat straw	Functional OH groups (aliphatic, total phenolic), S, G, H, COOH groups, mmol/g	^{31}P NMR ~15 mg	Popova 2020 Q [33]	RSD 1.7%…11.2%	–	The authors quote "reproducibility" but the RSD most likely refers to repeatability.
Kraft lignin (KL), lignosulfonate (LS), hydrolysis lignin (HL), organosolv lignin (OS) and soda lignin (SL)	Uncondensed aromatics, aliphatic OH, 5-substituted aromatics or condensed aromatics, carboxylic acids, mmol/g	Benchtop ^{31}P NMR 269.57 mg	Gracia-Vitoria 2021 A [34]	0.6%…2% Repeatability RSD	Relative bias: at medium levels (above 5 mmol/g) 2-6% (10% for modified lignins); at low levels (around 1 mmol/g) up to 27.5% (high-field NMR data)	The percentage error (relative error) determined using simple model compounds was, regardless of hydroxyl type 2-6% for technical lignins and above 10% for modified lignins.
Kraft softwood lignin (KSW), organosolv hardwood lignin (OHW), soda lignin (SOD)	Aliphatic OH, 4-O-5' or syringyl OH, 5–5', β-5, guaiacyl OH, p-hydroxyphenyl OH, carboxylic acid OH, mmol/g	Benchtop 24.3 MHz ^{31}P NMR 30 mg	Araneda 2022 A [35]	0.3%…16.1% Repeatability RSD	1%…37% against highe field instrument results	The aim is to demonstrate usefulness of benchtop NMR for quantitative lignin analysis. Precision and bias values refer to results with benchtop NMR instrument.

[a] Article type: "Q" – focus on qNMR in general, "A" – focus on accuracy of qNMR. IS is internal standard, MWL is Milled wood lignin, HW is Hard wood, aMWL is alkaline modified milled wood lignin, AcMWL is Acetylated milled wood lignin, SEAL is steam explosion aspen lignin,, ArH is aromatic hydrogen, OS-W is organosolv lignins-wheat straw, OS-P is organosolv lignins-poplar, OS-S is organosolv lignins spruce, RSD is relative standard deviation, SD is standard deviation, DEPT is Distortionless Enhancement by Polarization Transfer. [b] Range of precision estimates for the whole analysis, if present. In many cases the type of precision was derived by authors of this work on the basis of the data in the original articles. [c] Range of possible bias estimates for the whole analysis, if present. Some bias estimates and types of bias were derived by authors of this work on the basis of the data in the original articles. [d] Comments and estimates of different uncertainty contributions if presented in the original works.

3.2. Accuracy of Determining Different Lignin Parameters

There are certain structural fragments in lignin that show high accuracy in the measurement of their content. These are, for example, the contents of the methoxy group, aliphatic as well as aromatic hydroxyl group, aromatic hydrogens and the contents of the S and G units, as well as the S/G ratio. Conversely, contents of minor moieties such as β–β, β-5 and β-1 are examples of parameters that typically determined with lower accuracy [10]. In order to be accurately measurable, the respective structural feature first of all has to be reasonably abundant and yield strong signals. Secondly, its signals have to be sufficiently separated from signals of other structural fragments. The accuracy decreases with the decreasing content of lignin moieties in the lignin as well as in cases when there is a poor spectral resolution of the specific structural moieties.

The accuracy of lignin parameters also depends on the method used for the measurement. ^{13}C NMR tends to be more accurate than ^{31}P NMR [5]. In ^{31}P NMR, the appropriate choice and stability of internal standards affects the results [5].

3.3. Trends

Early works addressed accuracy/error only very superficially, but in some of the recent works these issues have received a lot of attention. It is perhaps worth to single out here the works by Balakshin and Capanema that have addressed the accuracy of qNMR in lignin analysis very thoroughly [5,8,10].

There is a large diversity of presenting accuracy-related data between papers from different authors. Therefore, it is not easy to spot clear trends in accuracy. Nevertheless, as a very broad generalization, in more recent works, the precision typically has become higher (i.e. lower standard deviation), and bias has become lower. The RSD values observed in the early 1990s articles were quite high compared to the precision, which is found in the recent articles. The same is observed with bias, the bias observed in recent articles is typically lower compared to the articles published in 1980-2000.

3.4. Identified Issues

Here are some issues that were identified as recurring:

(1) The term “lignin” refers to a number of materials differing by the way of extracting, as well as by other treatments. Although the main structural units are the same, different treatments may have modified or degraded some parts of the structure [27]. This has been identified above as an uncertainty source that becomes especially important with the more extensively treated lignins. With such lignins measurand definition [17] can become problematic. In the case of partly decomposed/processed lignins, what is it that was intended to be measured? Was it the ratio of structural units in the original lignin or in the partly decomposed lignin?

(2) In most studies replicate measurements were carried out and thus precision can be estimated. Yet, in most of them it is not specified whether the replicate measurements were made on the same day or on different days, whether from the same subsample of from different subsamples. As a result, the meaning of the precision estimate (expressed as RSD or otherwise) in many cases remains obscure.

(3) The bias estimates (often termed as “error” by authors) were given against different reference values: independently determined lignin composition, known composition of a model mixture, lignin composition determined with a higher field NMR [35]. Two types of issues can be identified. Firstly, it is often unclear, how close is the model system composition to the actual lignin and thus, how relevant are the obtained bias estimates. Secondly, bias refers to systematic effects. So, in order to determine bias, it is not sufficient to just compare a measurement result with a reference value. Instead, it is necessary to carry out replicate measurements in order to suppress the influence of random effects. Depending on how the replicate measurements were made – within a day or over a longer time – the bias estimates are different – within-day, or long-term bias (see Section 6 in the course described in [17]). These bias estimates differ not only by their value but also by their meaning. Within-day bias includes effects that

are systematic within a day but become random over time. Long-term bias does not include such effects. These issues are typically not addressed in papers related to qNMR analysis of lignin.

4. Conclusions and Recommendations

On the basis of the literature analysis two ways for advancing the field could be envisaged. Firstly, the measurand – what exactly is measured – needs to be defined carefully. Some examples follow. Is it the content of some group or moiety in the bulk lignin or in a sample? In the first case the sampling uncertainty is part of the uncertainty budget, in the second case it is not. Is the aim to relate the analysis result to the original untreated lignin or to the lignin that has undergone treatment/degradation?

Secondly, whenever qNMR lignin analysis results are presented, at least the following information should be included:

(1) number of replicate measurements (i.e. number of spectra recorded, not just number of scans within a spectrum);
(2) how were the replicate measurements carried out time-wise (on the same day, within couple of days, over a long term);
(3) which steps of the analysis were repeated – first of all, was a separate subsample used and sample preparation/dissolution done with every replicate;
(4) if comparison with a reference value (e.g. via a well-characterized lignin sample or a mixture of model compounds) is carried out then again, the items 1-3 should be reported, so that the bias and its type could be reasonably evaluated.

Thirdly, there is a shortage of reliable reference values that practitioners could use for comparing their results with. Such reference values could logically be carried by certified reference materials. To the best of our knowledge the COMAR database of certified reference materials does not currently have any materials related to lignin composition. However, as an optimistic note, according to ref [10], the Alcell and Indulin lignin samples from various sources yielded results that were remarkably similar (within the same type of lignin), in spite of originating from different sources. The differences that were observed between different samples of the same type of lignin were similar to the differences between the replicates of the same sample. As a result, there is a small deviation between batches of the same type of lignin that have been distributed within the lignin community [10]. Such materials could be formalized as reference materials (possibly even certified via interlaboratory comparisons) and widely distributed.

Acknowledgements

This research was funded by the Estonian Research Council grant PRG690 and by the EU through the European Regional Development Fund under project TK141 (2014-2020.4.01.15-0011). This study is part of the PhD thesis of Shrikant Shivaji Pawade.

ORCID
Shrikant Shivaji Pawade: 0000-0002-8457-3644
Lauri Toom: 0000-0001-9693-8180
Koit Herodes: 0000-0003-1763-1784
Ivo Leito: 0000-0002-3000-4964

References

[1] W. Boerjan, J. Ralph and M. Baucher (2003). Lignin biosynthesis, *Annu. Rev. Plant Biol.* **54**, 519–546.
[2] I.F. Demuner, J.L. Colodette, A.J. Demuner and C.M. Jardim (2019). Biorefinery review: wide-reaching products through kraft lignin, *BioRes*. **14**, 7543–7581.

[3] Y.-X. An, N. Li, H. Wu, W.-Y. Lou and M.-H. Zong (2015). changes in the structure and the thermal properties of kraft lignin during its dissolution in cholinium ionic liquids, *ACS Sustainable Chem. Eng.* **3**, 2951–2958.
[4] J. R. Obst and L. L. Landucci (1986). Quantitative ^{13}C NMR of lignins – methoxyl : aryl ratio, *Holzforschung* **40**, 87–92.
[5] M. Balakshin and E. Capanema (2015). On the quantification of lignin hydroxyl groups with ^{31}P and ^{13}C NMR spectroscopy, *J. Wood Chem. Technol.* **35**, 220–237.
[6] E.A. Capanema, M.Yu. Balakshin and J.F. Kadla (2005). Quantitative characterization of a hardwood milled wood lignin by nuclear magnetic resonance spectroscopy, *J. Agric. Food Chem.* **53**, 9639–9649.
[7] D. Argyropoulos (1994). Quantitative phosphorus-31 NMR Analysis of lignins, a new tool for the lignin chemist, *J. Wood Chem. Tech.* **14**, 45–63.
[8] M.Yu. Balakshin, E.A. Capanema, R.B. Santos, H. Chang and H. Jameel (2016). Structural analysis of hardwood native lignins by quantitative ^{13}C NMR spectroscopy, *Holzforschung* **70**, 95–108.
[9] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML (2012). JCGM 200:2012. International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3rd ed., Sèvres.
[10] M.Yu. Balakshin and E.A. Capanema (2015). Comprehensive structural analysis of biorefinery lignins with a quantitative ^{13}C NMR approach, *RSC Adv.* **5**, 87187–87199.
[11] Y.-Y. Wang, X. Meng, Y. Pu and A. J. Ragauskas (2020). Recent advances in the application of functionalized lignin in value-added polymeric materials, *Polymers* **12**, 2277.
[12] I. Sulaeva, G. Zinovyev, J.-M. Plankeele, I. Sumerskii, T. Rosenau and A. Potthast (2017). Fast track to molar-mass distributions of technical lignins, *ChemSusChem*. **10**, 629–635.
[13] T. Rundlöf, M. Mathiasson, S. Bekiroglu, B. Hakkarainen, T. Bowden and T. Arvidsson (2010). Survey and qualification of internal standards for quantification by ^{1}H NMR spectroscopy, *J. Pharmaceut. Biomed. Anal*. **52**, 645–651.
[14] G.F. Pauli (2001). qNMR ? a versatile concept for the validation of natural product reference compounds, *Phytochem. Anal.* **12**, 28–42.
[15] Z.-F. Wang, Y.-L. You, F.-F. Li, W.-R. Kong and S.-Q. Wang (2021). Research progress of NMR in natural product quantification, *Molecules* **26**, 6308.
[16] A. Menditto, M. Patriarca and B. Magnusson (2007). Understanding the meaning of accuracy, trueness and precision, *Accred. Qual. Assur.* **12**, 45–47.
[17] I. Leito, I. Helm and L. Jalukse (2015). Using MOOCs for teaching analytical chemistry: experience at University of Tartu, *Anal. Bioanal. Chem*. **407**, 1277–1281.
[18] D.B. Hibbert, E.-H. Korte and U. Örnemark (2021). Metrological and quality concepts in analytical chemistry (IUPAC Recommendations 2021), *Pure and Applied Chemistry*. **93**, 997–1048.
[19] D.B. Hibbert, Ed. (2023). Compendium of terminology in analytical chemistry, 4th ed., The Royal Society of Chemistry.
[20] Stephen.L.R. Ellison and A. Williams, Eds. (2012). Eurachem/CITAC guide: quantifying uncertainty in analytical measurement, 3rd ed., Eurachem/CITAC, Teddington.
[21] L.L. Landucci (1985). Quantitative ^{13} C NMR Characterization of Lignin 1. A Methodology for High Precision, *Holzforschung*. **39**, 355–360.
[22] G. Gellerstedt and D. Robert (1987). Quantitative ^{13}C NMR analysis of kraft lignins., *Acta Chem. Scand.* **41b**, 541–546.
[23] C.-L. Chen and D. Robert (1988). Methods in enzymology, Elsevier, pp. 137–174.
[24] O. Faix, D.S. Argyropoulos, D. Robert and V. Neirinck (1994). Determination of hydroxyl groups in lignins evaluation of ^{1} H-, ^{13} C-, ^{31} P-NMR, FTIR and wet chemical methods, *Holzforschung* **48**, 387–394.
[25] P.M. Froass, A.J. Ragauskas and J. Jiang (1998). Nuclear magnetic resonance studies. 4. Analysis of residual lignin after kraft pulping, *Ind. Eng. Chem. Res.* **37**, 3388–3394.
[26] M. Zawadzki and A. Ragauskas (2001). N-hydroxy compounds as new internal standards for the 31P-NMR Determination of lignin hydroxy functional groups, *Holzforschung* **55**, 283–285.
[27] Z. Xia, L.G. Akim and D.S. Argyropoulos (2001). Quantitative ^{13}C NMR analysis of lignins with internal standards, *J. Agric. Food Chem.* **49**, 3573–3578.
[28] S. Heikkinen, M.M. Toikka, P.T. Karhunen and I.A. Kilpeläinen (2003). Quantitative 2D HSQC (Q-HSQC) via suppression of *J* -dependence of polarization transfer in NMR spectroscopy: application to wood lignin, *J. Am. Chem. Soc*. **125**, 4362–4367.
[29] E.A. Capanema, M.Y. Balakshin and J.F. Kadla (2004). A comprehensive approach for quantitative lignin characterization by NMR spectroscopy, *J. Agric. Food Chem.* **52**, 1850–1860.
[30] L. Zhang and G. Gellerstedt (2007). Quantitative 2D HSQC NMR determination of polymer structures by selecting suitable internal standard references, *Magn. Reson. Chem.* **45**, 37–45.

[31] S. Constant, H.L.J. Wienk, A.E. Frissen, P. de Peinder, R. Boelens, D.S. van Es, R.J. H. Grisel, B.M. Weckhuysen, W.J. J. Huijgen, R.J.A. Gosselink and P.C.A. Bruijnincx (2016). New insights into the structure and composition of technical lignins: a comparative characterisation study, *Green Chem.* **18**, 2651–2665.

[32] M.T. Amiri, S. Bertella, Y.M. Questell-Santiago and J.S. Luterbacher (2019). Establishing lignin structure-upgradeability relationships using quantitative ^{1}H– ^{13}C heteronuclear single quantum coherence nuclear magnetic resonance (HSQC-NMR) spectroscopy, *Chem. Sci.* **10**, 8135–8142.

[33] Y.A. Popova, S.L. Shestakov, A.V. Belesov, I.I. Pikovskoi and A.Yu. Kozhevnikov (2020). Comprehensive analysis of the chemical structure of lignin from raspberry stalks (Rubus idaeus L.), *Int. J. Biol. Macromol*. **164**, 3814–3822.

[34] J. Gracia-Vitoria, M. Rubens, E. Feghali, P. Adriaensens, K. Vanbroekhoven and R. Vendamme (2022). Low-field benchtop versus high-field NMR for routine ^{31}P analysis of lignin, a comparative study, *Ind. Crop. Prod.* **176**, 114405.

[35] J.F. Araneda, I.W. Burton, M. Paleologou, S.D. Riegel and M.C. Leclerc (2022). Analysis of lignins using ^{31}P benchtop NMR spectroscopy: quantitative assessment of substructures and comparison to high-field NMR, *Can. J. Chem.* **100**, 799–808.

[36] S. Westwood, K. Lippa, Y. Shimuzu, B. Lalerle, T. Saito, D. Duewer, X. Dai , S. Davies , M. Ricci and A. Baldan (2023). Methods for the SI-traceable value assignment of the purity of organic compounds (IUPAC Technical Report), *Pure Appl. Chem*. **95**, 1–77.