Chemical purity using quantitative $^1$H-nuclear magnetic resonance: a hierarchical Bayesian approach for traceable calibrations

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 Metrologia 53 1193
(http://iopscience.iop.org/0026-1394/53/5/1193)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 129.6.179.122
This content was downloaded on 28/09/2016 at 14:54

Please note that terms and conditions apply.
Chemical purity using quantitative $^1$H-nuclear magnetic resonance: a hierarchical Bayesian approach for traceable calibrations

Blaza Toman, Michael A Nelson and Katrice A Lippa

National Institute of Standards and Technology, US Department of Commerce, Gaithersburg, MD, USA

E-mail: toman@nist.gov

Received 27 July 2016
Accepted for publication 25 August 2016
Published 28 September 2016

Abstract

Chemical purity assessment using quantitative $^1$H-nuclear magnetic resonance spectroscopy is a method based on ratio references of mass and signal intensity of the analyte species to that of chemical standards of known purity. As such, it is an example of a calculation using a known measurement equation with multiple inputs. Though multiple samples are often analyzed during purity evaluations in order to assess measurement repeatability, the uncertainty evaluation must also account for contributions from inputs to the measurement equation. Furthermore, there may be other uncertainty components inherent in the experimental design, such as independent implementation of multiple calibration standards. As such, the uncertainty evaluation is not purely bottom up (based on the measurement equation) or top down (based on the experimental design), but inherently contains elements of both. This hybrid form of uncertainty analysis is readily implemented with Bayesian statistical analysis. In this article we describe this type of analysis in detail and illustrate it using data from an evaluation of chemical purity and its uncertainty for a folic acid material.

Keywords: measurement uncertainty, observation equation, measurement equation, internal calibration, Markov chain Monte Carlo

(Some figures may appear in colour only in the online journal)
audio frequency signal and digitized, then lastly Fourier transformed to obtain NMR spectral data in the frequency domain. Each individual signal in the spectrum therefore represents a structurally distinct molecular moiety of the measured sample.

The amplitude of each spin component of the FID is directly proportional to the number of corresponding resonant nuclei, and thus the signal intensities of the Fourier transformed spectra can be used to directly infer the ratio amount of nuclei for each unique resonance [11–13]. It is for this reason that qNMR, when performed with proper experimental conditions, is considered a direct primary ratio measurement method for quantitative chemical analysis. The common isotope of hydrogen, $^1$H ($\approx 99.9\%$ abundance, $s = 1/2$, 2 spin states), is the usual nucleus for high-precision qNMR as it provides the highest analytical sensitivity and excellent linearity of signal intensity with respect to $^1$H concentration ($R \geq 0.999$) [12, 14].

Figure 1 illustrates the qNMR measurement of folic acid (FA) dissolved in deuterated aqueous (D$_2$O) phosphate buffer.

Precise and unbiased quantity determinations of the purity of (soluble) liquid or solid chemical materials may be made via ratio reference of mass and NMR signal intensity to those of primary chemical standards of known purity [15]. The main component of these standards must contain the appropriate reference nuclei, as well as a distinct spectral peak and common solubility, but they are otherwise not compound-specific. The $^1$H-qNMR assessment described in this article is an evaluation of the mass purity, $P_{\text{PC}}$, of a neat chemical folic acid material. The measurement equation used to derive this quantity is presented as equation (1).

$$P_{\text{PC}} = \left( \frac{N_{\text{IS}}}{N_{\text{PC}}} \right) \times \left( \frac{M_{\text{PC}}}{M_{\text{IS}}} \right) \times \left( \frac{A_{\text{PC}}}{A_{\text{IS}}} \right) \times \left( \frac{m_{\text{IS}}}{m_{\text{PC}}} \right) \times P_{\text{IS}} \quad (1)$$

where:
- $N_{\text{PC}}$ multiplicity (# H/peak) of primary chemical species signal
- $N_{\text{IS}}$ multiplicity (# H/peak) of internal reference standard species signal
- $M_{\text{PC}}$ relative molar mass (molecular weight, g mol$^{-1}$) of primary species
- $M_{\text{IS}}$ relative molar mass (molecular weight, g mol$^{-1}$) of the internal standard species
- $A_{\text{PC}}$ integrated area of primary species signal
- $A_{\text{IS}}$ integrated area of the internal reference standard species signal
- $m_{\text{PC}}$ mass (g) of the composite material weighed for sample solution, adjusted for buoyancy effects
- $m_{\text{IS}}$ mass (g) of the internal reference standard weighed for sample solution, adjusted for buoyancy effects
- $P_{\text{IS}}$ known purity (g g$^{-1}$) of the internal reference standard

Due to chemical and spectroscopic limitations of the measurement system, purity determinations were first made for select secondary reference materials that were suitable for measurement with the folic acid species. For this evaluation, two high-purity ($P_{\text{IS}} \geq 99.99\%$) primary reference standards, Standard Reference Material (SRM) 350b Benzoic acid (BA) and SRM 84k Potassium Hydrogen Phthalate (KHP), were used as internal calibrators to determine the mass purity of two chemically-distinct secondary reference materials ($P_{\text{IS}} > 99.9\%$): methylsulfonylethane (MSM, also known as dimethyl sulfone) and 2,2-dimethylpropanedioic acid (Me$_2$PDA, also known as dimethylmalonic acid). With this experimental design, the evaluated mass purity of folic acid is traceable to the certified mass fraction purity of both primary standards.

Section 2 describes the statistical model used to evaluate purity of the primary chemical species based on a qNMR analysis. Section 3 describes the implementation of this model during the investigation performed to evaluate the purity of a neat folic acid material. Conclusions are given in section 4.

2. The measurement model in terms of observation equations

The measurement equation (equation (1)) is based on four different measured quantities: $A_{\text{PC}}$, $A_{\text{IS}}$, $m_{\text{PC}}$, and $m_{\text{IS}}$. The remaining input quantities are either known constants, $N_{\text{PC}}$, $N_{\text{IS}}$, $M_{\text{PC}}$, and $M_{\text{IS}}$, or quantities with a mean and uncertainty specified by a certificate, $P_{\text{IS}}$. Though molecular masses ($M_{\text{PC}}$ and $M_{\text{IS}}$) are imperfectly known due to natural variation in isotopic abundances, the slight degree of uncertainty is relatively inconsequential. Given such data, the most common approach to the evaluation of purity and its uncertainty would be to apply a standard Guide to the expression of uncertainty in measurement (GUM) [16] analysis to each individual sample’s inputs and propagate the uncertainties through the measurement equation. This would produce a purity estimate and associated uncertainty for each sample. GUM-style analyses are readily accomplished ‘by hand’ or using software such as the NIST Uncertainty Machine [17].

One potentially critical factor for chemical mass purity evaluations is the need to constrain the result to lie within the interval (0 g g$^{-1}$, 1 g g$^{-1}$). The usual GUM analysis does not naturally preserve this constraint, which is especially relevant for highly pure materials ($P_{\text{PC}} > 99\%$). But whether or not a constraint is necessary, in order to capture the between sample variability (reproducibility uncertainty) of multiple samples, the individual results must be combined to obtain a final purity estimate and uncertainty. When multiple samples with different internal standards are used, as was the case for the evaluation that we describe in detail in the following sections, accounting for all sources of uncertainty in a rigorous manner is not straightforward since the outputs from the measurement equation for single samples may not be statistically independent due to shared sources of uncertainty attributable to the experimental design. Thus, the evaluation of measurement uncertainty of the purity estimate includes both bottom-up elements, that is, uncertainty in the inputs to the measurement equation propagated through the function, and top-down elements, being sources of uncertainty due to factors that are part of the experimental design and whose contribution to uncertainty is accounted for using the variability of the measurements.

A statistically rigorous method of combining these bottom-up and top-down elements of the uncertainty analysis, while also preserving any natural physical and
chemical constraints, is best achieved using an observation equation [18] as an integral part of a Bayesian statistical model [19]. An observation equation is a statistical model for the measurements, which in this case are $A_{PC}$, $A_{IS}$, $m_{PC}$, and $m_{IS}$. While uniquely able to account for all of the various constraints and uncertainty sources present in a measurement process, Bayesian analysis requires so called prior distributions for all inputs that are not known constants. These prior distributions summarize all of the information about such quantities that is available before the measurements are obtained.

In order to specify the statistical model for the area measurements, we first note that equation (1) can be written as

$$P_{PC} \times \frac{N_{PC}}{M_{PC}} \times m_{PC} = P_{IS} \times \frac{N_{IS}}{M_{IS}} \times m_{IS} = K.$$
It then follows that:

$$P_{IS} \times \frac{N_{IS}}{M_{IS}} = K \times \frac{A_{IS}}{m_{IS}} \tag{2}$$

Since $N_{IS}$ and $M_{IS}$ are known constants, the constant $K$ converts the ratio of the measurement of the integrated area of the signal to the mass of the material in the sample solution into a measurement of chemical purity. Since the purity of the internal standard is known up to an uncertainty, it is possible to estimate $K$ and use it to estimate the purity of the analyte.

To estimate $K$ we first note that

$$A_{IS} = \left( P_{IS} \times N_{IS} \times \frac{m_{IS}}{M_{IS}} \right) \frac{1}{K},$$

a relationship which will define the expected value of the measurement of $A_{IS}$. We term the uncertainty for this measurement, $u_{A_{IS}}$, NMR area measurements can generally be represented by a Gaussian distribution, or when $u_{A_{IS}}$ is based on known degrees of freedom, by a Student $t$ distribution. Measurements of sample mass, $m_{IS}$, can also be represented as Gaussian or Student $t$ distributions with mean $\mu_{IS}$ and uncertainty $u_{m_{IS}}$. Assuming Gaussian distributions, the observation equations for $A_{IS}$ and $m_{IS}$ are therefore:

$$A_{IS} \sim N\left( \frac{P_{IS} \times N_{IS} \times \frac{\mu_{IS}}{M_{IS}}}{K}, \frac{1}{K} \right),\tag{3}$$

and

$$m_{IS} \sim N(\mu_{IS}, u_{m_{IS}}^2),$$

where the symbol ‘$\sim$’ is interpreted ‘distributed as’ and $N(\cdot)$ specifies a Gaussian distribution with given mean and variance. When the uncertainties $u_{A_{IS}}$ or $u_{m_{IS}}$ are based on known degrees of freedom, Student $t$ distributions, $t(\cdot)$, should be used instead of $N(\cdot)$.

The main objective of these internal standard-based observation equations is to obtain information about $K$ to be used in the analysis of similar observation equations for the primary chemical species. Generally, the information about the purity of the internal standard is given in a certificate, value $\mu_{IS}$ and uncertainty $u_{IS}$, and can be transformed into a probability distribution. If no constraint on purity is necessary, a Gaussian distribution with the given mean and variance, $P_{IS} \sim N(\mu_{IS}, u_{p_{IS}}^2)$, may be used. If purity must be constrained to lie between 0 and 1 then the beta distribution

$$P_{IS} \sim \text{Beta}(a, b)$$

where

$$a = \mu_{PC} \left[ \frac{\mu_{PC}(1 - \mu_{PC})}{u_{PC}^2} - 1 \right],$$

$$b = (1 - \mu_{PC}) \left[ \frac{\mu_{PC}(1 - \mu_{PC})}{u_{PC}^2} - 1 \right]$$

is appropriate.

Without additional information, a rectangular distribution on the interval $(0, c)$ for some constant $c$ can be used as a prior for $K$: $K \sim R(0, c)$. Again without additional information, a Gaussian distribution with mean 0 and a large variance is appropriate for $\mu_{IS}$: $\mu_{IS} \sim N(0, \text{large})$. In our application we found the results to be robust with respect to the choice of the variance of $\mu_{IS}$.

While the posterior distribution, which summarizes all the information resulting from a Bayesian analysis, of $K$ cannot be obtained in closed form, a Markov chain Monte Carlo (MCMC) [19] analysis using the free software OpenBUGS [20] is straightforward and produces a sample of random draws from this posterior distribution. The OpenBUGS code is given in the appendix.

The main objective of the analysis described in this article is the estimation of the purity of the primary chemical species. This is accomplished through Bayesian analysis of two observation equations for $A_{PC}$ and $m_{PC}$:

$$A_{PC} \sim N\left( \frac{P_{PC} \times N_{PC} \times \mu_{PC}}{M_{PC}} \frac{1}{K}, \frac{u_{2_{PC}}^2}{K} \right), \tag{4}$$

and

$$m_{PC} \sim N(\mu_{PC}, u_{PC}^2),$$

where the quantity of interest now is not $K$, but the measurand $P_{PC}$.

As above, all unknown constants need to have prior probability distributions. The analysis of equation (3) produces a random sample from the posterior distribution for $K$ which the OpenBUGS program can use as a prior distribution. A suitable prior distribution for $m_{PC}$ is again Gaussian distribution with mean 0 and a large variance $\mu_{PC} \sim N(0, \text{large})$.

A prior distribution for the measurand $P_{PC}$ needs to be defined expressly for each application, as this is where the specifics of the experimental design are expressed in the form of constraints and/or additional structure of the prior distribution. For example, purity of different samples analyzed using a single standard may be somewhat different due to slightly different measurement conditions or heterogeneity of the material, but the differences would usually be smaller than differences in purity of samples analyzed using two different internal standards. Such details can be modeled using a second level in the prior distribution of $P_{PC}$ creating a so called hierarchical structure.

Evaluation of equations (3) and (4) performed using MCMC preserves all correlations present in the statistical model. The resulting posterior distribution of $P_{PC}$ can be summarized in terms of a mean, standard deviation, and 95% uncertainty interval to produce the desired purity estimate.

3. Examples of purity determination

This section shows application of equations (3) and (4) to the specific case of purity evaluation of folic acid. The OpenBUGS code used for the analysis is given in the appendix.

3.1. Experimental design

For the evaluation of each secondary reference material, four samples containing BA as internal standard and three containing KHP were dissolved in per-deuterated dimethyl
sulfoxide, DMSO-d$_6$ (99.9% D atom purity; Cambridge Isotope Laboratories (CIL), Cambridge, MA). These secondary reference materials were evaluated for use in folic acid (FA) measurements because they have $^1$H resonance signals that do not interfere with those of FA and are mutually soluble in aqueous solutions. For purity determination of FA, four samples containing the MSM and three samples containing the Me$_2$PDA secondary standard were prepared in a K$_2$HPO$_4$/KD$_2$PO$_4$ aqueous (99.99% D atom purity D$_2$O, CIL) buffer solution. Approximately 4 mg–15 mg masses of neat FA and internal standard materials were weighed using an ultramicrobalance (Mettler Toledo UMX5; Columbus, OH).

Experimental NMR data was acquired by a Bruker Avance 600 MHz spectrometer equipped with a 5 mm broadband inverse (BBI) detection probe and operating with Topspin (Version 3.2) software. Experiments were performed with 128 scan repetitions (60 s delay time), a spectral width of 20.0276 ppm, and transmitter frequency offset (O1) of 6.175 ppm. A 90° excitation pulse width was calibrated for each sample. Data acquisition time was 5.45 s for each scan and 131 072 data points were collected for each FID. The probe temperature was 298 K. No $^{13}$C decoupling was executed during data acquisition. Processing of the Fourier transformed $^1$H spectra, including baseline corrections, phase adjustment, and signal integration, was performed manually.

3.2. Purity of the secondary internal standards MSM and Me$_2$PDA

During this investigation and for the reasons described herein, purity of FA was evaluated using two different internal standards, MSM and Me$_2$PDA. The inputs for the purity evaluation of MSM are given in table 1. Masses and areas are based on three measurements, and therefore the uncertainties of the measurements are associated with 2 degrees of freedom. Three or four different samples containing neat chemical and internal standard materials were prepared for each of the purity assessments.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 4</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{\text{PC}}$ (g mol$^{-1}$)</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
</tr>
<tr>
<td>$M_{\text{IS}}$ (g mol$^{-1}$)</td>
<td>122.12</td>
<td>122.12</td>
<td>122.12</td>
<td>122.12</td>
<td>122.12</td>
<td>204.22</td>
<td>204.22</td>
<td>204.22</td>
<td>204.22</td>
<td>204.22</td>
</tr>
<tr>
<td>$N_{\text{PC}}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{IS}}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{\text{PC}}$</td>
<td>1411977.105</td>
<td>975476.967</td>
<td>1460479.190</td>
<td>1044400884</td>
<td>2049421640</td>
<td>1735241388</td>
<td>1751551403</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{A_{\text{PC}}}$</td>
<td>413545.106</td>
<td>144615.57</td>
<td>246609.241</td>
<td>104452.893</td>
<td>243774.687</td>
<td>580978.61</td>
<td>364045.672</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{\text{IS}}$</td>
<td>750028.937</td>
<td>623003.497</td>
<td>518822.717</td>
<td>671799.916</td>
<td>456562.022</td>
<td>814572.865</td>
<td>747400.179</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{A_{\text{IS}}}$</td>
<td>73363.6013</td>
<td>349365.9</td>
<td>116331.207</td>
<td>165393.397</td>
<td>409770.0251</td>
<td>36063.2556</td>
<td>18405.9296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{\text{IS}}$ (mg)</td>
<td>9.2704</td>
<td>7.6297</td>
<td>6.6060</td>
<td>8.3029</td>
<td>5.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{m_{\text{PC}}}$ (mg)</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{m_{\text{IS}}}$ (mg)</td>
<td>9.2704</td>
<td>7.6297</td>
<td>6.6060</td>
<td>8.3029</td>
<td>5.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{\mu_{\text{PC}}}$ (g g$^{-1}$)</td>
<td>0.999 978</td>
<td>0.999978</td>
<td>0.999978</td>
<td>0.999978</td>
<td>0.999978</td>
<td>0.999978</td>
<td>0.999978</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{\mu_{\text{IS}}}$ (g g$^{-1}$)</td>
<td>0.000044</td>
<td>0.000044</td>
<td>0.000044</td>
<td>0.000044</td>
<td>0.000044</td>
<td>0.000044</td>
<td>0.000044</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We specified the prior for $P_{\text{PC}}$ as $P_{\text{PC}} \sim N(\mu_{\text{PC}}, u_{\mu_{\text{PC}}}^2)$, assigning $\mu_{\text{PC}}$ the non-informative prior $\mu_{\text{PC}} \sim N(0, d)$ with $d$ set to a large value. We assign $u_{\mu_{\text{PC}}}$ a non-informative gamma distribution prior, $u_{\mu_{\text{PC}}} \sim \text{Gamma}(e, f)$ with $e$ and $f$ set as large numbers, that captures the non-negativity of standard deviations. The calculated purities of the secondary internal standards. $P_{\text{IS}}$ in corresponding calculations of folic acid purity, were not constrained to a maximum of 1. Given that qNMR is a relative measurement, the effective purity of the internal standard is a normalized ratio of select resonant $^1$H content with respect to $m_{\text{IS}}$, rather than an absolute chemical mass fraction. Therefore, the $P_{\text{IS}}$ may exceed the natural limit of MSM or Me$_2$PDA purity. This may be the case if an assumed $M_{\text{IS}}$ is not consistent with the actual isotopic composition, or if there is a systematic bias of $A_{\text{IS}}$ associated with an indiscernible chemical impurity. Several qNMR evaluations of the high purity MSM secondary standard over multiple years and by multiple analysts indicate that $P_{\text{IS}}$ is greater than 1 if a molecular weight of 94.13 g mol$^{-1}$ is assumed.

The MCMC posterior distribution of $P_{\text{PC}}$ was summarized as a mean and standard deviation. The resulting values for each sample and internal standard are given in table 2.

The observation equation model as specified in equations (3) and (4) represents only the bottom-up elements of the analysis, and the results in table 2 are essentially identical to the usual uncertainty analysis obtained with the GUM procedure [16].
The top-down elements in the uncertainty quantification, that is, the between sample and between internal standard variability, can be accounted for in a separate layer of the observation equation model. This is preferable to simply averaging or otherwise combining the entries in table 2 because the values of purity for the samples based on the same internal standard are correlated, and the value of this correlation would need to be estimated to produce an accurate uncertainty for the final purity estimate. A hierarchical model [21] on $P_{PC}$ accomplishes this naturally, without the need for separate estimation of this correlation. We define

$$PN_i^{1}, \ldots, i = 1, 2; j = 1, \ldots, 3 \text{ or } 4 \quad (5)$$

and

$$\alpha_i \sim N(\mu_{P}, \omega^2)$$

where $i = 1$ for BA, and $i = 2$ for KHP. Here the parameters $\tau_j$ quantify the between-sample variability for each internal standard and the $\omega$ quantifies the between-internal standard variability, which is also the basis of the correlation between the sample values. Note that this is the usual random effects model [22], widely used to account for additional uncertainty in interlaboratory studies.

The final estimate of purity is the posterior mean of $\mu_{P}$. For MSM this is 1.0009, with standard uncertainty $= 0.0009$.

### Table 3. Inputs for secondary internal standard Me$_2$PDA.

<table>
<thead>
<tr>
<th>Internal standard</th>
<th>SRM 350b benzoic acid (BA)</th>
<th>SRM 84 k potassium hydrogen phthalate (KHP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Sample 1</td>
<td>Sample 2</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>$M_{PC}$ (g mol$^{-1}$)</td>
<td>132.11</td>
<td>132.11</td>
</tr>
<tr>
<td>$M_{IS}$ (g mol$^{-1}$)</td>
<td>122.121</td>
<td>122.121</td>
</tr>
<tr>
<td>$N_{IS}$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$N_{PC}$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$A_{PC}$</td>
<td>640.246.178.3</td>
<td>966.957.581</td>
</tr>
<tr>
<td>$u_{AE}$</td>
<td>93.106.606.06</td>
<td>24.507.84</td>
</tr>
<tr>
<td>$A_{IS}$</td>
<td>815.288.040.8</td>
<td>$1.108 \times 10^9$</td>
</tr>
<tr>
<td>$u_{AE}$</td>
<td>108.299.626.1</td>
<td>148.379.41</td>
</tr>
<tr>
<td>$m_{PC}$ (mg)</td>
<td>8.5701</td>
<td>5.9678</td>
</tr>
<tr>
<td>$u_{m_{PC}}$ (mg)</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>$m_{IS}$ (mg)</td>
<td>10.0868</td>
<td>6.3157</td>
</tr>
<tr>
<td>$u_{m_{IS}}$ (mg)</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>$\mu_{P}$ (g g$^{-1}$)</td>
<td>0.999978</td>
<td>0.999978</td>
</tr>
<tr>
<td>$u_{P}$ (g g$^{-1}$)</td>
<td>0.000044</td>
<td>0.000044</td>
</tr>
</tbody>
</table>

### Table 4. Summary outputs for purity of secondary internal standard Me$_2$PDA.

<table>
<thead>
<tr>
<th>Internal standard</th>
<th>Sample</th>
<th>Posterior mean (g g$^{-1}$)</th>
<th>Posterior standard deviation (g g$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA</td>
<td>1</td>
<td>0.9999</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.9991</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9992</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9998</td>
<td>0.0013</td>
</tr>
<tr>
<td>KHP</td>
<td>1</td>
<td>1.0040</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.9997</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9995</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9989</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

The top-down elements in the uncertainty quantification, that is, the between sample and between internal standard variability, can be accounted for in a separate layer of the observation equation model. This is preferable to simply averaging or otherwise combining the entries in table 2 because the values of purity for the samples based on the same internal standard are correlated, and the value of this correlation would need to be estimated to produce an accurate uncertainty for the final purity estimate. A hierarchical model [21] on $P_{PC}$ accomplishes this naturally, without the need for separate estimation of this correlation. We define

$$P_{PC_i} \sim N(\alpha_i, \tau^2_i), \quad i = 1, 2; j = 1, \ldots, 3 \text{ or } 4 \quad (5)$$

and

$$\alpha_i \sim N(\mu_{P}, \omega^2)$$

where $i = 1$ for BA, and $i = 2$ for KHP. Here the parameters $\tau_j$ quantify the between-sample variability for each internal standard and the $\omega$ quantifies the between-internal standard variability, which is also the basis of the correlation between the sample values. Note that this is the usual random effects model [22], widely used to account for additional uncertainty in interlaboratory studies.

The final estimate of purity is the posterior mean of $\mu_{P}$. For MSM this is 1.0009, with standard uncertainty $= 0.0009$. For the evaluation of folic acid purity these are interpreted as $P_{IS} \sim N(1.0009, 0.0009^2)$. Figure 2 shows the posterior distribution of $\mu_{P}$. The same procedure was followed to obtain a purity estimate for the second internal standard, Me$_2$PDA, using inputs given in table 3. Bayesian analysis of equations (3) and (4), with these data and the same prior distributions as before, produced the values in table 4.

These estimates were combined using the hierarchical model of equation (5) to obtain the posterior distribution of $\mu_{P}$, the purity of Me$_2$PDA. The posterior mean and standard deviation were 0.9994 and 0.0008, respectively. Figure 3 shows the posterior distribution of Me$_2$PDA.

### 3.3. Folic acid purity

To estimate the purity of the folic acid material, multiple samples were evaluated using each of the two internal standards. The inputs are given in table 5. As in the purity assessments described above, the analysis proceeds from the same four observation equations, here using...
Table 5. Inputs for folic acid (FA) samples.

<table>
<thead>
<tr>
<th>Internal standard</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 4</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mem (g mol⁻¹)</td>
<td>441.4</td>
<td>441.4</td>
<td>441.4</td>
<td>441.4</td>
<td>441.4</td>
<td>441.4</td>
<td>441.4</td>
</tr>
<tr>
<td>MS (g mol⁻¹)</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>94.13</td>
<td>132.11</td>
<td>132.11</td>
<td>132.11</td>
</tr>
<tr>
<td>NIS (mg)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>N(PC)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>AreaPC (µ)</td>
<td>435.884911</td>
<td>546.673820</td>
<td>310.049348</td>
<td>445.682930</td>
<td>627.834226</td>
<td>575.519508</td>
<td>479.363762</td>
</tr>
<tr>
<td>u(Apc)</td>
<td>13411.114</td>
<td>1163.9939</td>
<td>656.77749</td>
<td>1062.08287</td>
<td>862.113073</td>
<td>2004.86269</td>
<td>616.232246</td>
</tr>
<tr>
<td>AreaIS (µ)</td>
<td>11537.8716</td>
<td>1.88 × 10⁶</td>
<td>746.379745</td>
<td>1492.16752</td>
<td>579.567035</td>
<td>606.371428</td>
<td>854.524945</td>
</tr>
<tr>
<td>u(Apc)</td>
<td>576893.568</td>
<td>937811.13</td>
<td>373.189873</td>
<td>746.083.776</td>
<td>289.788.017</td>
<td>303.185.714</td>
<td>427.262.472</td>
</tr>
<tr>
<td>m(PC) (mg)</td>
<td>5.0559</td>
<td>5.2872</td>
<td>4.7875</td>
<td>4.7024</td>
<td>4.4486</td>
<td>5.0308</td>
<td>4.4013</td>
</tr>
<tr>
<td>u(mPC) (mg)</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>u(mIS) (mg)</td>
<td>2.5868</td>
<td>3.5075</td>
<td>2.21</td>
<td>3.0396</td>
<td>1.112</td>
<td>1.4288</td>
<td>2.1288</td>
</tr>
<tr>
<td>u(mIS) (mg)</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

Table 6. Summary outputs for purity of the folic acid (FA) samples.

<table>
<thead>
<tr>
<th>Internal standard</th>
<th>Sample 1</th>
<th>Posterior mean (g g⁻¹)</th>
<th>Posterior standard deviation (g g⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSM</td>
<td>1</td>
<td>0.9068</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.9069</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9062</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9064</td>
<td>0.0021</td>
</tr>
<tr>
<td>Me2PDA</td>
<td>1</td>
<td>0.9044</td>
<td>0.0024</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.9051</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9058</td>
<td>0.0021</td>
</tr>
</tbody>
</table>

The information about the two internal standards (probability distributions of $\mu_{p1}$ and $\mu_{p2}$ as prior distributions) obtained in the previous section as

$$\begin{align*}
A_{IS} & \sim t\left(\mu_{IS} \times N_{IS} \times \frac{1}{M_{IS}^2}, 1/K_i, u_{A_{IS}^2}, 2\right); \\
m_{IS} & \sim N(\mu_{IS}, u_{m_{IS}}^2), \quad i = 1, 2
\end{align*}$$

(6)

where $i = 1$ for BA and $i = 2$ for KHP. The area measurements were based on 3 replicates and so a Student $t$ distribution with 2 degrees of freedom was used. Bayesian analysis of equation (6) leads to a posterior distribution for each $K_i$ which can then be applied in the analysis of the observation equations for $A_{PC}$ and $m_{PC}$:

$$\begin{align*}
A_{PC} & \sim t\left(\mu_{PC} \times N_{PC} \times \frac{1}{M_{PC}^2}, 1/K_i, u_{A_{PC}^2}, 2\right); \\
m_{PC} & \sim N(\mu_{PC}, u_{m_{PC}}^2), \quad i = 1, 2
\end{align*}$$

(7)

One difference between this analysis and the two described above is that the purity of folic acid must be constrained to the interval (0 g g⁻¹, 1 g g⁻¹). To produce purity estimates separately for the individual samples, this constraint is applied using a rectangular prior distribution: $P_{PC} \sim R(0, 1)$. MCMC analysis produced the results in table 6.

Since somewhat different folic acid purity values were inferred from the two internal standards, the corresponding sample sets were combined separately. The constraint to lie in

the interval (0 g g⁻¹, 1 g g⁻¹) requires a different version of the hierarchical model for the purity values:

$$P_{PC_i} \sim Beta(a_i, b_i), \quad i = 1, 2; \quad j = 1, \ldots, 3 \quad \text{or} \quad 4 \quad (8)$$

and

$$a_i = \frac{\mu_{PC_i}^2}{\sigma_i^2}; \quad b_i = \frac{1 - \mu_{PC_i}}{\sigma_i^2}; \quad \mu_{PC_i} \sim R(0.6, 1); \quad \sigma_i \sim R(0, 0.1).$$

The analysis results were not sensitive to the ranges of the uniform distributions for the parameters $\mu_{PC}$ and $\sigma_i$, but convergence was faster with the slightly informative prior distributions given here. As there were different distributions for the measurand depending on the internal standard $i$, there were two different sets of posterior means and standard deviations as given in table 7.

It is required to arrive at a single posterior distribution of purity. There are various statistical approaches for combining posterior distributions of a measurand, the most conservative in the sense of producing the largest uncertainty in the estimate is the linear pool method [21]. This treats the two posterior distributions of $PC$ (call them $p_1$ and $p_2$) as equally likely to be correct and combines them to produce a single probability distribution $p$: $p(x) = \frac{p_1(x) + p_2(x)}{2}$ for any value $x$ of $PC$.

The value and uncertainty for the purity of FA, $P_{FA}$, produced in this way was the mean and standard deviation of the posterior distribution of $p$: 0.9056 g g⁻¹ and 0.0039 g g⁻¹, respectively. Considering this uncertainty to be too conservative, we used an alternative approach that assumes the purities of the samples using either of the two internal standards are related through the same parameters of the hierarchical model, that is:

Table 7. Summary outputs for purity for the two internal standards.

<table>
<thead>
<tr>
<th>Internal standard</th>
<th>Posterior mean (g g⁻¹)</th>
<th>Posterior standard deviation (g g⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSM</td>
<td>0.9065</td>
<td>0.0027</td>
</tr>
<tr>
<td>Me2PDA</td>
<td>0.9048</td>
<td>0.0047</td>
</tr>
</tbody>
</table>
This model does not account for the between internal standard variability directly, rather it becomes part of the repeatability uncertainty component. Using this model, the estimate of purity is the mean and standard deviation of the posterior distribution of $\mu_{PC}$: 0.9058 g g$^{-1}$ and 0.0011 g g$^{-1}$, respectively. Figure 4 shows the four posterior distributions.

4. Conclusions

We have developed an observation equation model for chemical purity determinations based on qNMR measurements. The advantage of this approach is that it ensures that any constraints on the purity estimate (for example that the mass-fraction purity should not exceed 1) are satisfied and incorporates both top-down and bottom-up uncertainty evaluations in the same statistical model, thus naturally including correlations of the measurements due to the experimental design. We illustrated Bayesian analysis of this model on measurements of folic acid (FA) purity using two different internal standards. We showed how the purity of the internal standards can first be evaluated and then used as input into the FA analysis. The OpenBUGS code used for the analysis is given in the appendix. This model was developed to be implemented as part of a direct measurement approach for traceable chemical purity assessments via qNMR, and allows observation of natural limits and evaluation of realistic uncertainty intervals.

Acknowledgments

The authors would like to thank David Duewer and Antonio Possolo for many helpful comments and suggestions that greatly improved this article. Certain commercial equipment, instruments, or materials are identified in this paper to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Appendix. OpenBUGS code and data file for evaluation of purity of folic acid

```plaintext
{## constants
# M_{IS1} is the relative molar mass of MSM
# M_{PC1}, M_{PC2} is the relative molar mass of FA
# M_{IS2} is the relative molar mass of Me2PDA
## inputs
# m_{IS1} is the mass measurement of MSM
# u_{m_{IS1}} is the uncertainty of the mass of the MSM weighed for sample solution
# m_{PC1} is the mass of FA in samples analyzed using MSM
# m_{IS2} is the mass measurement of Me2PDA
# u_{m_{IS2}} is the uncertainty in the mass of the Me2PDA weighed for sample solution
```
# $m_{PC2}$ is the mass of FA in samples analyzed using Me$_2$PDA
# $A_{IS1}$ integrated area of the MSM signal
# $u_{A_{IS1}}$, uncertainty in the integrated area of the MSM signal
# $A_{PC1}$ integrated area of FA signal in samples analyzed using MSM
# $u_{A_{PC1}}$, uncertainty in the integrated area of FA signal in samples analyzed using MSM
# $A_{IS2}$ integrated area of the Me$_2$PDA signal
# $u_{A_{IS2}}$, uncertainty in the integrated area of the Me$_2$PDA signal
# $A_{PC2}$ integrated area of FA signal in samples analyzed using Me$_2$PDA
# $u_{A_{PC2}}$, uncertainty in the integrated area of FA signal in samples analyzed using Me$_2$PDA

## parameters of the probability distributions of the mass measurements
# $m_{m_{IS1}}$ is $\mu_{IS1}$
# $m_{m_{PC1}}$ is $\mu_{PC1}$
# $m_{m_{IS2}}$ is $\mu_{IS2}$
# $m_{m_{PC2}}$ is $\mu_{PC2}$

## prior parameters
# $\mu_{P1}$ is the purity of MSM
# $\mu_{P2}$ is the purity of Me$_2$PDA

## the measurand
# $\mu$ is $\mu_{PC}$, i.e. the mean of the FA purity

\[
\begin{align*}
M_{IS1} &< - 94.13 \\
M_{PC1} &< - 441.4 \\
p_{rcm_{IS1}} &< - 1/(u_{m_{IS1}} * u_{m_{IS1}}) \\
M_{IS2} &< - 132.11 \\
M_{PC2} &< - 441.4 \\
p_{rcm_{IS2}} &< - 1/(u_{m_{IS2}} * u_{m_{IS2}})
\end{align*}
\]

# define the parameters of the Beta prior distribution for the purity of FA. The mean of this
# distribution $\mu$ is the measurand.

\[
\begin{align*}
a &< - \mu u_{sd} \\
b &< - (1 - \mu)u_{sd} \\
\mu &\sim uu(0.6, 1) \\
\mu_{sd} &\sim uu(0, 0.1)
\end{align*}
\]

# define the distribution of the purity of MSM and of Me$_2$PDA, as obtained in section 3.2.
# Variability is given in terms of precision, that is, as $1/0.0009^2$ for MSM and $1/0.0008^2$ for
# Me$_2$PDA.

# $\mu_{P1}$ ~ $\mathcal{N}(1.0009, 1234.568)$
# $\mu_{P2}$ ~ $\mathcal{N}(0.9994, 1562.500)$

## specify the prior distributions for the 4 samples analyzed using MSM

\[
\begin{align*}
\text{for}(i \in 1:4) \{ p_{PC1}[i] \sim dbeta(a, b) \\
\quad m_{m_{m_{IS1}}}[i] \sim dnorm(0, 1.0 \times 10^{-5}) \\
\quad m_{m_{m_{PC1}}}[i] \sim dnorm(0, 1.0 \times 10^{-5}) \}
\end{align*}
\]

## specify the observation equations for the mass measurements for the 4 samples analyzed using # MSM

\[
\begin{align*}
\text{for}(i \in 1:4) \{ m_{IS1}[i] \sim dnorm(m_{m_{m_{IS1}}}[i], prec_{m_{IS1}}) \\
\quad m_{PC1}[i] \sim dnorm(m_{m_{m_{PC1}}}[i], prec_{m_{IS1}}) \}
\end{align*}
\]

## specify the observation equation for the area measurements of MSM

\[
\begin{align*}
\text{for}(i \in 1:4) \{ k_{1}[i] \sim uu(0, 0.01) \\
\quad \text{mean}_{A_{IS1}}[i] &< - \mu_{P1} * m_{m_{m_{IS1}}}[i]/M_{IS1}/k_{1}[i] \\
\quad \text{prec}_{A_{IS1}}[i] &< - 1/(u_{A_{IS1}}[i]^2 * u_{A_{IS1}}[i]) \}
\end{align*}
\]
for( i in 1:4){
    A_{IS1}[i] \sim \text{dt}(\text{mean } A_{IS1}[i], \text{prec } A_{IS1}[i], 2) 
    
    #
    # specify the observation equation for area of FA for the 4 samples analyzed using MSM
    
    for( i in 1:4){
        k.cut1[i] <- \text{cut}(k_1[i])
        \text{mean } A_{PC1}[i] <- \frac{p_{PC1}[i] \times mu_{rec}[i]}{M_{PC1}/k.cutt1[i]}
        \text{prec } A_{PC1}[i] <- \frac{1}{(A_{rec}[i] \times A_{rec}[i])}
    }

    for( i in 1:4){
        A_{PC1}[i] \sim \text{dt}(\text{mean } A_{PC1}[i], \text{prec } A_{PC1}[i])
    }
}

# specify the prior distributions for the 3 samples analyzed using Me2PDA

for( i in 1:3){
    p_{PC2}[i] \sim \text{dbeta}(a, b)
    \text{mum } PC2[i] ~ \text{dtnorm}(0, 1.0 \times 10^{-5})
    \text{mum } IS2[i] ~ \text{dtnorm}(0, 1.0 \times 10^{-5})
}

# specify the observation equations for the mass measurements for the 3 samples analyzed using Me2PDA

for( i in 1:3){
    m_{IS2}[i] ~ \text{dtnorm}(\text{mum } IS2[i], \text{prec } m_{IS2})
    m_{PC2}[i] ~ \text{dtnorm}(\text{mum } PC2[i], \text{prec } m_{IS2})
}

# specify the observation equation for the area measurements of Me2PDA

for( i in 1:3){
    k_2[i] ~ \text{dunif}(0, 0.01)
    \text{mean } A_{IS2}[i] <- \frac{p_{PC2}[i] \times mu_{rec}[i]}{M_{IS2}/k_2[i]}
    \text{prec } A_{IS2}[i] <- \frac{1}{(A_{rec}[i] \times A_{rec}[i])}
}

for( i in 1:3){
    A_{IS2}[i] \sim \text{dt}(\text{mean } A_{IS2}[i], \text{prec } A_{IS2}[i], 2)
}

for( i in 1:3){
    k.cut2[i] <- \text{cut}(k_2[i])
    \text{mean } A_{PC2}[i] <- \frac{p_{PC2}[i] \times mu_{rec}[i]}{M_{PC2}/k.cutt2[i]}
    \text{prec } A_{PC2}[i] <- \frac{1}{(A_{rec}[i] \times A_{rec}[i])}
}

for( i in 1:3){
    A_{PC2}[i] \sim \text{dt}(\text{mean } A_{PC2}[i], \text{prec } A_{PC2}[i])
}
}

## Data file:

list(m_{IS1}=c(0.00258680,0.00350750,0.002210,0.00303960),
    m_{PC1}=c(0.00505592,0.00528720,0.00478750,0.00470240),
    um_{IS1}=0.00000005,
    A_{IS1}=c(0.00153787136,0.00187600,0.007463797450,0.014921675520),
    u_{AIS1}=c(0.00000000797460,0.00000000638230,0.00000000766960,0.00000000731310),
    A_{PC1}=c(0.00043588491110,0.005466738200,0.003120493480,0.000445682930),
    u_{APC1}=(0.000013141111,0.000011639930,0.000006567770,0.000001062083),
    m_{IS2}=c(0.001112,0.00142880,0.00212880),
    m_{PC2}=c(0.0004486600,0.00050388,0.00044013),u_{m_{IS2}}=0.00000005,
    A_{IS2}=c(0.00057957603500,0.006063714280,0.008545249450),
    u_{AIS2}=(0.00000012225400,0.000000134470,0.000000190030),
    A_{PC2}=c(0.006278342260,0.005755195080,0.004793673620),
    u_{APC2}=(0.00000086211300,0.000020048630,0.000000162320))
References

[18] Possolo A and Toman B 2007 Metrologia 44 464–75