Published in final edited form as:

Metrologia. 2016 ; 53(5): 1193–1203. doi:10.1088/0026-1394/53/5/1193.

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

Blaza Toman, Michael A Nelson, and Katrice A Lipka

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

Abstract

Chemical purity assessment using quantitative ^1H -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

1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a prominent measurement technique for chemical analysis. Quantitative solution-state NMR (qNMR) is considered a primary ratio measurement method, through which ratio quantities of chemical substances can be determined directly [1–5]. For this reason, qNMR is highly useful for traceable chemical purity assessments [6–10].

NMR spectroscopy is a measurement method for isotopes that have nonzero nuclear spins, and thus a magnetic moment and angular momentum. Each of these isotopes has opposite spin states ($+$ and $-$) with equivalent energy, dependent upon neutron and proton

composition, that are characterized by a spin quantum number (s). The NMR measurement technique is based on the perturbation of these nuclear magnetic moments when placed within a strong external magnetic field and radiation of a specific (Larmor) radio frequency is absorbed. Electron spins surrounding the nucleus induce opposing magnetic fields, and as such, this resonance is also dependent on the surrounding electronic environment; that is, the molecular structure(s) of the sample material.

After the radiation absorption has ceased, the perturbed magnetic moments relax and precess about the external magnetic field at the corresponding resonance frequencies. The net sum of magnetic precessions induces voltage oscillations that are observable as a time domain signal, known as a free induction decay (FID). This signal is amplified, then converted to an 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, ^1H ($\approx 99.9\%$ 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 ^1H concentration ($R \geq 0.999$) [12, 14]. Figure 1 illustrates the qNMR measurement of folic acid (FA) dissolved in deuterated aqueous (D_2O) 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 ^1H -qNMR assessment described in this article is an evaluation of the mass purity, P_{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_{PC} multiplicity (# H/peak) of primary chemical species signal

N_{IS} multiplicity (# H/peak) of internal reference standard species signal

M_{PC} relative molar mass (molecular weight, g mol^{-1}) of primary species

M_{IS} relative molar mass (molecular weight, g mol^{-1}) of the internal standard species

A_{PC} integrated area of primary species signal

A_{IS} integrated area of the internal reference standard species signal

m_{PC} mass (g) of the composite material weighed for sample solution, adjusted for buoyancy effects

m_{IS} mass (g) of the internal reference standard weighed for sample solution, adjusted for buoyancy effects

P_{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 ($P_{IS} > 99.99\%$) 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 ($P_{IS} > 99.9\%$): methylsulfonylmethane (MSM, also known as dimethyl sulfone) and 2,2-dimethylpropanedioic acid (Me_2PDA , 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_{PC} , A_{IS} , m_{PC} , and m_{IS} . The remaining input quantities are either known constants, N_{PC} , N_{IS} , M_{PC} , and M_{IS} , or quantities with a mean and uncertainty specified by a certificate, P_{IS} . Though molecular masses (M_{PC} , and M_{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 ($P_{PC} > 99\%$). 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

$$P_{PC} \times \frac{N_{PC}}{M_{PC}} \times \frac{m_{PC}}{A_{PC}} = P_{IS} \times \frac{N_{IS}}{M_{IS}} \times \frac{m_{IS}}{A_{IS}} = K.$$

It then follows that:

$$P_{IS} \times \frac{N_{IS}}{M_{IS}} = K \times \frac{A_{IS}}{m_{IS}}. \quad (2)$$

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 mass μ_{IS} and uncertainty $u_{m_{\text{IS}}}$. Assuming Gaussian distributions, the observation equations for A_{IS} and m_{IS} are therefore:

$$A_{\text{IS}} \sim N \left(\left(P_{\text{IS}} \times N_{\text{IS}} \times \frac{\mu_{\text{IS}}}{M_{\text{IS}}} \right) \frac{1}{K}, u_{A_{\text{IS}}}^2 \right) \quad (3)$$

and

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

where the symbol ' $\sim$ ' is interpreted 'distributed as' and $N(\cdot)$ specifies a Gaussian distribution with given mean and variance. When the uncertainties $u_{A_{\text{IS}}}$ or $u_{m_{\text{IS}}}$ 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 μ_{IS} and uncertainty $u_{P_{\text{IS}}}$, and can be transformed into a probability distribution. If no constraint on purity is necessary, a Gaussian distribution with the given mean and variance, $P_{\text{IS}} \sim N \left(\mu_{\text{IS}}, u_{P_{\text{IS}}}^2 \right)$, may be used. If purity must be constrained to lie between 0 and 1 then the beta distribution

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

where

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

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

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 μ_{IS} : $\mu_{\text{IS}} \sim N(0, \text{large})$. In our application we found the results to be robust with respect to the choice of the variance of μ_{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} :

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

and

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

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 $\mu_{PC}; \mu_{PC} \sim N(0, \text{large})$.

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 ^1H 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_2PDA secondary standard were prepared in a $\text{K}_2\text{HPO}_4/\text{KD}_2\text{PO}_4$ aqueous (99.99% D atom purity D_2O , CIL) buffer solution. Approximately 4 mg–15 mg masses of neat FA and internal standard materials were weighed using an ultra-microbalance (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 ^1H spectra, including baseline corrections, phase adjustment, and signal integration, was performed manually.

3.2. Purity of the secondary internal standards MSM and Me_2PDA

During this investigation and for the reasons described herein, purity of FA was evaluated using two different internal standards, MSM and Me_2PDA . 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.

We specified the prior for P_{PC} as $P_{\text{PC}} \sim N(\mu_{\text{PC}}, u_{\mu_{\text{PC}}}^2)$, assigning μ_{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_{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 ^1H content with respect to m_{IS} , rather than an absolute chemical mass fraction. Therefore, the P_{IS} may exceed the natural limit of MSM or Me_2PDA purity. This may be the case if an assumed M_{IS} is not consistent with the actual isotopic composition, or if there is a systematic bias of A_{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_{IS} is greater than 1 if a molecular weight of 94.13 g mol^{-1} is assumed.

The MCMC posterior distribution of P_{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

$$P_{PCij} \sim N(\alpha_i, \tau_i^2), i=1, 2; j=1, \dots, 3 \text{ or } 4 \quad (5)$$

and

$$\alpha_i \sim N(\mu_{P1}, \omega^2)$$

where $i = 1$ for BA, and $i = 2$ for KHP. Here the parameters τ_i quantify the between-sample variability for each internal standard and the ω 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 μ_{P1} . For MSM this is 1.0009, with standard uncertainty = 0.0009. For the evaluation of folic acid purity these are interpreted as $P_S \sim N(1.0009, 0.0009^2)$. Figure 2 shows the posterior distribution of μ_{P1}

The same procedure was followed to obtain a purity estimate for the second internal standard, Me₂PDA, 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 μ_{P1} , the purity of Me₂PDA. The posterior mean and standard deviation were 0.9994 and 0.0008, respectively. Figure 3 shows the posterior distribution of Me₂PDA.

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 the information about the two internal standards

(probability distributions of μ_{P1} and μ_{P2} as prior distributions) obtained in the previous section as

$$\begin{aligned} A_{IS_i} &\sim t \left(\left(\mu_{P_i} \frac{\mu_{IS_i} \times N_{IS_i}}{M_{IS_i}} \right) \frac{1}{K_i}, u_{A_{IS_i}}^2, 2 \right); \\ m_{IS_i} &\sim N \left(\mu_{IS_i}, u_{m_{IS_i}}^2 \right); \quad i=1, 2 \end{aligned} \quad (6)$$

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} :

$$\begin{aligned} A_{PC_i} &\sim t \left(\left(P_{PC_i} \frac{\mu_{PC_i} \times N_{PC_i}}{M_{PC_i}} \right) \frac{1}{K_i}, u_{A_{PC_i}}^2, 2 \right); \\ m_{PC_i} &\sim N \left(\mu_{PC_i}, u_{m_{PC_i}}^2 \right); \quad i=1, 2 \end{aligned} \quad (7)$$

One difference between this analysis and the two described above is that the purity of folic acid must be constrained to the interval $(0 \text{ g g}^{-1}, 1 \text{ g g}^{-1})$. To produce purity estimates separately for the individual samples, this constraint is applied using a rectangular prior distribution: $P_{PC_i} \sim R(0, 1)$. 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 $(0 \text{ g g}^{-1}, 1 \text{ g g}^{-1})$ requires a different version of the hierarchical model for the purity values:

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

and

$$a_i = \frac{\mu_{P_i}}{\sigma_i^2}; \quad b_i = \frac{1 - \mu_{P_i}}{\sigma_i^2}; \quad \mu_{P_i} \sim R(0.6, 1); \quad \sigma_i \sim R(0.0, 1).$$

The analysis results were not sensitive to the ranges of the uniform distributions for the parameters μ_{P_i} and σ_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 P_{PC} (call them p_1 and p_2) as equally likely to be correct and combines them to produce a single probability distribution

$$p:p(x) = \frac{p_1(x) + p_2(x)}{2} \text{ for any value } x \text{ of } P_{PC}.$$

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:

$$P_{PC_i} \sim \text{Beta}(a, b); \quad a = \frac{\mu_{PC}}{\sigma^2}; \quad b = \frac{1 - \mu_{PC}}{\sigma^2} \quad (9)$$

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 μ_{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

```

{## 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 Me2PDA
  #  $A_{IS1}$  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 Me2PDA signal
  #  $u_{A_{IS2}}$  uncertainty in the integrated area of the Me2PDA signal
  #  $A_{PC2}$  integrated area of FA signal in samples analyzed using Me2PDA
  #  $u_{A_{PC2}}$  uncertainty in the integrated area of FA signal in samples analyzed using Me2PDA
## parameters of the probability distributions of the mass measurements
  #  $mu_{m_{IS1}}$  is  $\mu_{IS1}$ 
  #  $mu_{m_{PC1}}$  is  $\mu_{PC1}$ 
  #  $mu_{m_{IS2}}$  is  $\mu_{IS2}$ 
  #  $mu_{m_{PC2}}$  is  $\mu_{PC2}$ 
## prior parameters
  #  $mu_{p1}$  is the purity of MSM
  #  $mu_{p2}$  is the purity of Me2PDA
## the measurand
  #  $mu$  is  $\mu_{PC}$ , i.e. the mean of the FA purity
   $M_{IS1} <- 94.13$ 
   $M_{PC1} <- 441.4$ 

   $prec_{m_{IS1}} <- 1 / (u_{m_{IS1}} * u_{m_{IS1}})$ 

```

```

MIS2 <- 132.11
MPC2 <- 441.4

precmIS2 <- 1 / (umIS2 * umIS2)

# define the parameters of the Beta prior distribution for the purity of FA. The mean of this
# distribution mu is the measurand.
a <- mu/(sd * sd)
b <- (1 - mu)/(sd * sd)
mu ~ dunif(0.6, 1)
sd ~ dunif(0, 0.1)
#

# define the distribution of the purity of MSM and of Me2PDA, as obtained in section 3.2.
# Variability is given in terms of precision, that is, as 1/0.00092 for MSM and 1/0.00082 for
# Me2PDA.
muP1 ~ dhorm(1.0009, 1234 568)
muP2 ~ dhorm(0.9994, 1562 500)
#

# specify the prior distributions for the 4 samples analyzed using MSM
for(i in 1:4){pPC1[i] ~ dbeta(a, b)

  mumIS1 [i] ~ dnorm (0, 1.0 × 10-5)
  mumPC1 [i] ~ dnorm (0, 1.0 × 10-5) }
#

# specify the observation equations for the mass measurements for the 4 samples analyzed using # MSM
for (i in 1:4) {mIS1 [i] ~ dnorm (mumIS1 [i], precmIS1)
  mPC1 [i] ~ dnorm (mumPC1 [i], precmIS1) }
#

# specify the observation equation for the area measurements of MSM
for(i in 1:4){
  k1[i] ~ dunif(0, 0.01)

  meanAIS1 [i] <- muP1 * mumIS1 [i] / MIS1 / K1 [i]
  precAIS1 [i] <- 1 / (uAIS1 [i] * uAIS1 [i]) }

  for (i in 1:4){
    AIS1 [i] ~ d(meanAIS1[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] <- cut(k1[i])

  meanAPC1 [i] <- pPC1 [i] * mumPC1 [i] / MPC1 / k.cut1 [i]
  precAPC1 [i] <- 1 / (uAPC1 [i] * uAPC1 [i])

```

```

}
for(j in 1:4){
  APC1[j] ~ dnorm(meanAPC1[j], precAPC1[j])
  #specify the prior distributions for the 3 samples analyzed using Me2PDA
  for(i in 1:3){pPC2[i] ~ dbeta(a, b)

  mumPC2 [i] ~ dnorm (0, 1.0 × 10-5)
  mumIS2 [i] ~ dnorm (0, 1.0 × 10-5)}

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

for(i in 1:3) {mIS2 [i] ~ dnorm (mumIS2 [i], precmIS2)
  mPC2 [i] ~ dnorm (mumPC2 [i], precmIS2) }

#
# specify the observation equation for the area measurements of Me2PDA
for(j in 1:3){
  k2[j] ~ dunif(0,0.01)

  meanAIS2 [i] <- - muP2 * mumIS2 [i] / MIS2 / k2 [i]
  precAIS2 [i] <- - 1 / (uAIS2 [i] * uAIS2 [i])}

for(i in 1:3){
  AIS2[i] ~ d((meanAIS2[i], precAIS2[i], 2))
# specify the observation equation for area of FA for the 3 samples analyzed using Me2PDA
for(j in 1:3){
  k.cut2[j] <- cut(k2[j])

  meanAPC2 [i] <- - pPC2 [i] * mumPC2 [i] / MPC2 / k.cut2 [i]
  precAPC2 [i] <- - 1 / (uAPC2 [i] * uAPC2 [i])
}

for(i in 1:3){
  APC2[i] ~ dnorm(meanAPC2[i], precAPC2[i])
}

###
Data file:
list(mIS1 = c(0.0025 868,0.0035075,0.00221,0.0030396),
mPC1 = c(0.0050 559,0.0052 872,0.0047 875,0.0047 024),

umIS1 = 0.0000 005,
AIS1 = c(0.011 537 871 36,0.018 76,0.007 463 797 45,0.014 921 675 52),

uAIS1 = c(0.000 000 979 46,0.000 000 638 23,0.000 000 766 96,0.000 000 731 31),
APC1 = c(0.004 358 849 11,0.005 466 738 20,0.003 120 493 48,0.004456829 30),

```

$$\begin{aligned}
 u_{A_{PC1}} &= \alpha(0.000\ 013\ 411\ 11, 0.000\ 011\ 639\ 93, 0.000\ 006\ 567\ 77, 0.000\ 010\ 620\ 83), \\
 m_{IS2} &= \alpha(0.001\ 112, 0.001\ 4288, 0.002\ 1288), \\
 m_{PC2} &= \alpha(0.004\ 4486, 0.005\ 0038, 0.004\ 4013), \quad u_{m_{IS2}} = 0.000\ 0005, \\
 A_{IS2} &= \alpha(0.005\ 795\ 760\ 35, 0.006\ 063\ 714\ 28, 0.008\ 545\ 249\ 45), \\
 u_{A_{IS2}} &= \alpha(0.000\ 000\ 122\ 54, 0.000\ 000\ 134\ 47, 0.000\ 000\ 190\ 03), \\
 A_{PC2} &= \alpha(0.006\ 278\ 342\ 26, 0.005\ 755\ 195\ 08, 0.004\ 793\ 63762), \\
 u_{A_{PS2}} &= \alpha(0.000\ 008\ 621\ 13, 0.000\ 020\ 048\ 63, 0.000\ 006\ 162\ 32)
 \end{aligned}$$

References

1. Quinn TJ. *Metrologia*. 1997; 34:61–5.
2. Malz F, Jancke H. *J Pharm Biomed Anal*. 2004; 38:813–23.
3. Holzgrabe U, Deubner R, Schollmayer C, Waibel B. *J Pharm Biomed Anal*. 2005; 38:806–812. [PubMed: 15893899]
4. Jancke H, Malz F, Haesselbarth W. *Accredit Qual Assur*. 2005; 10:421–9.
5. Pauli GF. *Phytochem Anal*. 2001; 12:28–42. [PubMed: 11704959]
6. DeBièvre P, Dybkaer R, Fajgelj A, Hibbert B. *Pure Appl Chem*. 2011; 83:1873–935.
7. Nelson M, Bedner M, Lang B, Toman B, Lippa K. *Anal Bioanal Chem*. 2015; 407:8557–69. [PubMed: 26345446]
8. Weber M, Hellriegel C, Rueck A, Sauermoser R, Wuethrich J. *Accredit Qual Assur*. 2013; 18:91–8.
9. Saito T, Ihara T, Koike M, Kinugasa S, Fujimine Y, Nose K, Hirai T. *Accredit Qual Assur*. 2009; 14:79–86.
10. Schoenberger T. *Anal Bioanal Chem*. 2012; 403:247–54. [PubMed: 22327969]
11. Jancke H. *NMR Spectroscopy as a primary analytical metho*. CCQM Report. 1998; 98:1–12.
12. Bharti SK, Roy R. *Trends Anal Chem*. 2012; 35:5–26.
13. Wells RJ, Cheung J, Hook JM. *Accredit Qual Assur*. 2004; 9:450–6.
14. Gadape HH, Parikh KS. *J Chem Pharm Res*. 2011; 3:649.
15. Al-Deen TS, Hibbert DB, Hook JM, Wells RJ. *Accredit Qual Assur*. 2004; 9:55–63.
16. Joint Committee for Guides in Metrology. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*. Sevres: International Bureau of Weights and Measures (BIPM); 2008. available at www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf (Accessed: June 2016)
17. Lafarge, T., Possolo, A. *NIST Uncertainty Machine*. Gaithersburg, MD: National Institute of Standards and Technology; 2015. available at <http://uncertainty.nist.gov/> (Accessed: June 2016)
18. Possolo A, Toman B. *Metrologia*. 2007; 44:464–75.
19. Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, D. *Bayesian Data Analysis*. 3rd. Boca Raton, FL: CRC Press; 2013.
20. Lunn DJ, Spiegelhalter D, Thomas A, Best N. *Stat Med*. 2009; 28:3049–82. [PubMed: 19630097]
21. Toman B. *Technometrics*. 2007; 49:81–7.
22. Toman B, Possolo A. *Accredit Qual Assur*. 2009; 14:553–63.

Figure 1.
 ^1H -NMR measurement data for folic acid in D_2O phosphate buffer as: (a) time domain FID and (b) Fourier transformed spectrum.

Figure 2. Posterior distribution of purity (g g^{-1}) for MSM. Open square is the posterior mean, solid circles are the endpoints of a 95% uncertainty interval.

Figure 3. Posterior distribution of purity (g g^{-1}) of Me₂PDA. Open square is the posterior mean, solid circles are the endpoints of a 95% uncertainty interval.

Figure 4. Purity (g g^{-1}) of folic acid based on methylsulfonyl-methane (MSM) (blue), 2,2-dimethylpropanedioic acid (Me_2PDA) (green), the linear pool consensus (red), and the version without the internal standard component (violet). Open squares are the posterior means, solid circles are the endpoints of the 95% uncertainty intervals.

Table 1

Inputs for Secondary Internal Standard MSM.

Internal standard	SRM 350b benzoic acid (BA)			SRM 84 k potassium hydrogen phthalate (KHP)			
Parameters	Sample 1	Sample 2	Sample 3	Sample 4	Sample 1	Sample 2	Sample 3
M_{PC} (g mol ⁻¹)	94.13	94.13	94.13	94.13	94.13	94.13	94.13
M_{IS} (g mol ⁻¹)	122.121	122.121	122.121	122.121	204.22	204.22	204.22
N_{IS}	1	1	1	1	1	1	1
N_{PC}	1	1	1	1	1	1	1
A_{PC}	1411 977 105	975 476 967	1460 479 190	1044 400 884	2049 421 640	1735 241 388	1751 551 403
u_{APC}	413 545.106	144 615.57	246 609.241	104 452.893	243 774.687	580 978.61	364 045.672
A_{IS}	750028 937	623003 497	518822 717	671799 916	456562022	814572865	747400179
u_{AIS}	73363.6013	349365.9	116331.207	165393.397	40977.0251	360632556	18405.9296
m_{PC} (mg)	13.4375	9.2011	14.3194	9.9598	10.3778	8.2405	6.3565
$u_{m_{PC}}$ (mg)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
m_{IS} (mg)	9.2704	7.6297	6.6060	8.3209	5.0207	8.4030	5.8872
$u_{m_{IS}}$ (mg)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
$\mu_{P_{IS}}$ (g g ⁻¹)	0.999 978	0.999 978	0.999 978	0.999 978	0.999 911	0.999 911	0.999 911
$u_{P_{IS}}$ (g g ⁻¹)	0.000 044	0.000 044	0.000 044	0.000 044	0.000 054	0.000 054	0.000 054

Table 2

Summary outputs for purity of secondary internal standard MSM.

Internal standard	Sample	Posterior mean (g g ⁻¹)	Posterior standard deviation (g g ⁻¹)
BA	1	1.001	0.0004
	2	1.0007	0.0024
	3	1.0009	0.0011
	4	1.0011	0.0008
KHP	1	1.0009	0.0004
	2	1.0012	0.0004
	3	1.0004	0.0003

Table 3

Inputs for secondary internal standard Me₂PDA.

Internal standard	SRM 350b benzoic acid (BA)				SRM 84 k potassium hydrogen phthalate (KHP)			
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 1	Sample 2	Sample 3	Sample 4
M_{PC} (g mol ⁻¹)	132.11	132.11	132.11	132.11	132.11	132.11	132.11	132.11
M_{IS} (g mol ⁻¹)	122.121	122.121	122.121	122.121	204.22	204.22	204.22	204.22
N_{IS}	1	1	1	1	1	1	1	1
N_{PC}	1	1	1	1	1	1	1	1
A_{PC}	640246178.3	966957581	966 370 521	864647730	1061975433	1093570394	1097418797	1136552998
$u_{A_{PC}}$	93106.60606	24507.84	53870.701	73390.8488	86891.1374	63973.5968	37111.3776	67927.198
A_{IS}	815288040.8	1.108×10^9	955889852	1079955713	803319978	944110166	613684339	720228806
$u_{A_{IS}}$	1082996261	148379.41	8876.6106	357305.977	59360.1527	96723.3087	61710.4119	100666.621
m_{PC} (mg)	8.5701	5.9678	6.7620	7.0961	5.1465	6.0393	5.8252	5.2730
$u_{m_{PC}}$ (mg)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
m_{IS} (mg)	10.0868	6.3157	6.1778	8.1847	6.0208	8.0579	5.0333	5.1597
$u_{m_{IS}}$ (mg)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
$\mu_{P_{IS}}$ (g g ⁻¹)	0.999 978	0.999 978	0.999 978	0.999 978	0.999 911	0.999 911	0.999 911	0.999 911
$u_{P_{IS}}$ (g g ⁻¹)	0.000044	0.000044	0.000044	0.000044	0.000054	0.000054	0.000054	0.000054

Table 4Summary outputs for purity of secondary internal standard Me₂PDA.

Internal standard	Sample	Posterior mean (g g ⁻¹)	Posterior standard deviation (g g ⁻¹)
BA	1	0.9999	0.0004
	2	0.9991	0.0004
	3	0.9992	0.0001
	4	0.999	0.0013
KHP	1	1.0004	0.0002
	2	0.9997	0.0004
	3	0.9995	0.0003
	4	0.9989	0.0005

Table 5

Inputs for folic acid (FA) samples.

Internal standard	Methylsulfonylmethane (MSM)				2,2-dimethylpropanedioic acid (Me ₂ PDA)			
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 1	Sample 2	Sample 3	Sample 4
M_{PC} (g mol ⁻¹)	441.4	441.4	441.4	441.4	441.4	441.4	441.4	441.4
M_{IS} (g mol ⁻¹)	94.13	94.13	94.13	94.13	132.11	132.11	132.11	132.11
N_{IS}	1	1	1	1	1	1	1	1
N_{PC}	1	1	1	1	1	1	1	1
$Area_{PC}$	435 884 911	546 673 820	312 049 348	445 682 930	627 834 226	575 519 508	479 363 762	616 232 246
u_{APC}	1341 111.45	1163 993.9	656 777.749	1062 082.87	862 113.073	2004 862.69	616 232.246	854 524 945
$Area_{IS}$	1153 787 136	1.88×10^9	746 379 745	1492 167 552	579 576 035	606 371 428	854 524 945	427 262.472
u_{AUS}	576 893.568	937 811.13	373 189.873	746 083.776	289 788.017	303 185.714	427 262.472	4.4013
m_{PC} (mg)	5.0559	5.2872	4.7875	4.7024	4.4486	5.0038	4.4013	0.0005
$u_{m_{PC}}$ (mg)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
m_{IS} (mg)	2.5868	3.5075	2.21	3.0396	1.112	1.4288	2.1288	0.0005
$u_{m_{IS}}$ (mg)	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005

Table 6

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

Internal standard	Sample	Posterior mean (g g ⁻¹)	Posterior standard deviation (g g ⁻¹)
MSM	1	0.9068	0.0023
	2	0.9069	0.0020
	3	0.9062	0.0021
	4	0.9064	0.0021
Me ₂ PDA	1	0.9044	0.0024
	2	0.9051	0.0028
	3	0.9058	0.0021

Table 7

Summary outputs for purity for the two internal standards.

Internal standard	Posterior mean (g g ⁻¹)	Posterior standard deviation (g g ⁻¹)
MSM	0.9065	0.0027
Me ₂ PDA	0.9048	0.0047