CCQM-K147

Comparison of value-assigned CRMs for niacin (vitamin B_3) in milk powder and infant formula matrices

Key Comparison
Track A Model 2

Final Report
June 2019

Measurement Laboratory

Claudia Marcela Salazar Arzate
Centro Nacional de Metrología (CENAM) and
National Institute of Standards and Technology (NIST)

Data Analysis Laboratory

Hugo Gasca Aragón
Centro Nacional de Metrología (CENAM)

Coordinating Laboratories

Claudia Marcela Salazar Arzate, Hugo Gasca Aragón, and Mariana Arce Osuna
Centro Nacional de Metrología (CENAM)
Querétaro, 76246 México

Carolyn Q. Burdette and Katrice A. Lippa
National Institute of Standards and Technology (NIST)
Gaithersburg, MD 20899 USA

With contributions from:

Laura Regalado Contreras
Centro Nacional de Metrología (CENAM), México

Joonhee Lee, Sunyoung Lee, Song-Yee Baek, Byungjoo Kim
Korea Research Institute of Standards and Science (KRISS), Republic of Korea

Ting Huang, Zhen Guo, Hong Mei Li, and Qing He Zhang
National Institute of Metrology (NIM), China

Melissa M. Phillips, Blaza Toman, and David L. Duewer
National Institute of Standards and Technology (NIST), USA

DISCLAIMER

Certain commercial materials, instruments, software, and equipment are identified in this report to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by NIST or CENAM, nor does it imply that the material, instrument, software, or equipment are necessarily the best available for the purpose.

ABSTRACT

The 2017 CCQM-K147 “Comparison of value-assigned CRMs for niacin (vitamin B_3) in milk powder and infant formula matrices” is the first Key Comparison directly testing the chemical measurement services provided to customers by National Metrology Institutes (NMIs) and Designated Institutes (DIs) through certified reference materials (CRMs). CRMs certified for vitamin B_3 (as niacin or niacinamide) content in milk powder and infant formula matrices were compared using measurements made on these materials under repeatability conditions. Five NMIs/DIs submitted seven CRMs certified for niacinamide and two CRMs certified for niacin. These materials represent most of the higher-order reference materials available in 2017 for this important nutrimental measurand.

Generalized Gauss Markov Regression (GGMR) and Bayesian methods were used to establish the Key Comparison Reference Function (KCRF) relating the CRM certified values to the repeatability measurements. The niacinamide and niacin results for all nine CRMs were deemed equivalent at the 95 % level of confidence and were used to define the KCRF for vitamin B_3 (as niacinamide).

Monte Carlo methods and Bayesian methods were used to estimate 95 % level-of-confidence coverage intervals for the relative degrees of equivalence of materials, $\%d \pm U_{95}$ ($\%d$), and of the participating NMIs/DIs, $\%D \pm U_{95}$ ($\%D$). The Bayesian method estimates were selected as the final DoE values. For the niacinamide and niacin materials, all of the $\%D \pm U_{95}(\%D)$ intervals, were within (-10 to 10) % of the consensus results and all of these are statistically equivalent. These results demonstrate that the participating institutions can value-assign CRMs for niacinamide and/or niacin in milk powder and infant formula matrices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

ANOVA	analysis of variance
CCQM	Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology
CENAM	Centro Nacional de Metrología, NMI: México
CMC	Calibration and Measurement Capability
CRM	certified reference material
CV	coefficient of variation, expressed in %: CV = 100·s/$\bar{x}$
DI	designated institute
DoE	degrees of equivalence
GDR	generalized distance regression
GGMR	generalized Gauss Markov regression
ID	isotope dilution
IRMM	Institute for Reference Materials and Measurements: Belgium
JCTLM	Joint Committee for Traceability in Laboratory Medicine
JRC	Joint Research Centre, NMI: European Union
KC	Key Comparison
KCRF	Key Comparison Reference Function
KCRV	Key Comparison Reference Value
KRISS	Korea Research Institute of Standards and Science, NMI: Republic of Korea
LAI	leave all in
LC-IDMS/MS	liquid chromatography isotope dilution tandem mass spectrometry
LOO	leave one out
MRM	multiple reaction monitoring
MS	mass spectrometry
m/z	mass to charge ratio
NIM	National Institute of Metrology, NMI: China
NIST	National Institute of Standards and Technology, NMI: USA
NMI	national metrology institute
OAWG	Organic Analysis Working Group
PBMC	parametric bootstrap Monte Carlo
PI	participating institution
PT	proficiency test

SYMBOLS

~	signifies “distributed as”
ˆ	signifies an estimated quantity when above a symbol
α	intercept
β	slope
ε_i	residual difference for a given material
E	residual random error
γ_j	between-unit differences for a given material
δ_{jk}	within-unit differences of the k^{th} aliquot of the j^{th} unit of a given material
μ	mean
$\sigma_{a\%}$	between-aliquot imprecision/within-material heterogeneity
$\sigma_{c\%}$	between-campaign imprecision/between-unit material heterogeneity
$\sigma_{char\%}$	characterization uncertainty
$\sigma_{r\%}$	repeatability imprecision
$\%\sigma$	Standard deviation expressed as percent relative to value x: $\%\sigma=100*\sigma/x$
ν	degrees of freedom
d_i	degree of equivalence for the i^{th} material
D	degree of equivalence for one NMI/DI in a KC
$\%d_i$	degree of equivalence for the i^{th} material as a percentage of the value
$\%D$	degree of equivalence for one NMI/DI in a KC of the value
i	index over materials
j	index over campaigns and/or units
k	index over replicates or aliquots
l	index over replicates
N(.,.)	Normal distribution
n_a	number of aliquots of each unit
n_c	number of measurement campaigns
n_m	number of materials
n_r	number of replicates of each aliquot or unit
n_u	number of units of each material
R	generic representation of “instrument response”
R_{jkl}	instrument response for the l^{th} replicate of the k^{th} aliquot of the j^{th} unit of a given material
SIGN	signum or “sign” (±1) of a value
$t_{0.05,\nu}$	two-tailed Student’s t critical value for 95 % confidence with ν degrees of freedom
u	standard uncertainty
u_∞	“large sample” standard uncertainty
U_{95}	half-width of the 95 % level of confidence expanded uncertainty
V	generic representation of “assigned value”
V_i	assigned value for the i^{th} material

1.0 INTRODUCTION

1.1 Historical Background

The CCQM-K147 Track A Model 2 comparison of value-assigned materials was intended to compare reference material-based measurement services for vitamin B_3 in milk powder or infant formula matrices as they are delivered to customers and stakeholders. All national metrology institutes (NMIs) and designated institutes (DIs) that deliver measurement services via one or more value-assigned certified reference materials (CRMs), PT materials, or accuracy quality controls for vitamin B_3 (as niacin or niacinamide) in milk powder or infant formula matrices were invited to participate.

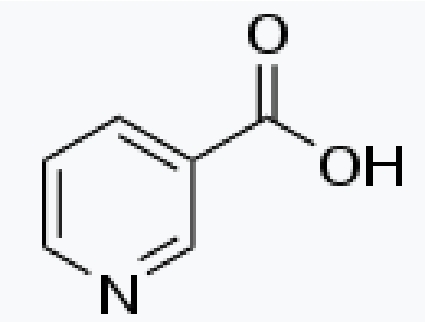
Participation in CCQM-K147 was accomplished by providing the study's Measurement Laboratory with materials that the participating institute value-assigned, kept in storage, and shipped to customers. The Centro Nacional de Metrología (CENAM) and the National Institute of Standards and Technology (NIST) volunteered to serve as Co-Coordinating and Measurement Laboratories for this study.

All comparison measurements were made at NIST's Gaithersburg, MD, laboratory location with an isotope dilution liquid chromatography tandem mass spectrometry (ID-LC-MS/MS) validated methodology [1] under repeatability conditions. These measurements were made by a CENAM analyst who was seconded to NIST for much of 2017.

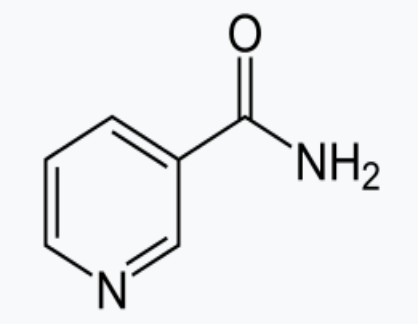
1.2 Measurands

Figure 1 displays the molecular structure and mass for niacin and niacinamide measurands.

Figure 1 Molecular structure and mass for the two measurands.



NIACIN
CAS Number: 59-67-6
MW: 123.1094 g/mol



NIACINAMIDE
CAS Number: 98-92-0
MW: 122.13 g/mol

Niacinamide, also known as nicotinamide, is a form of vitamin B_3 found in food and used as a dietary supplement to prevent and treat mouth lesions, pellagra, and acne. Niacinamide occurs as a component of a variety of biological systems. It is a critically

important part of the structures of NADH and NAD^+. Foods that contain niacinamide include yeast, meat, milk and green vegetables.

Niacin is obtained in the diet from a variety of whole and processed foods, with highest contents in fortified packaged foods, tuna, vegetables and other animal sources. Medication and supplemental niacin are primarily used to treat high blood cholesterol and pellagra. Insufficient niacin in the diet can cause nausea, skin and mouth lesions, anemia, headaches, and tiredness.

1.3 Comparison Design Background

This study design is based on the experimental design described at some length in the companion report "Comparison of value-assigned CRMs and PT materials: experimental design and data evaluation" [2]. It follows on the heels of previous Track B comparisons: CCQM-K79 (ethanol in water), CCQM-K80 (creatinine in serum) and CCQM-K142 (urea/uric acid in serum/plasma).

For the design of the comparison the coordinators considered the analyte levels for each potential participant. The expected vitamin B_3 mass fraction range was anticipated to be between 4 mg/kg and 110 mg/kg.

A target date for supplying those materials to the Measurement Laboratory was set and the materials were stored under the conditions specified in their Certificates until measurements were made. The measurements were made under repeatability conditions. The measurement result and the uncertainty for each material were determined.

Considering the wide range of mass fraction content of niacin (as niacinamide) present in the CCQM-K147 samples, two materials (high and low mass fraction level) were used as daily method validation/ control materials. Independent units of each control were used each day.

The probable heterogeneity effects of the materials due to the various fat-contents and types of matrix were considered in the dispersion of the measurement data, while keeping the type I and II errors below an upper bound of 5% and 9.5 % respectively, under the typical statistical assumptions (i.e., normality, balanced designs, independence, and unbiased methods). Type I error is the rejection of a true null hypothesis when in fact it is true, also known as "false positive" finding. In this case, the error of claiming the measured material is heterogeneous when it is not. Type II error is the rejection of the alternative hypothesis when in fact it is true, also known as "false negative" finding. In this case, the error of claiming the measured material is homogeneous when it is not.

A consensus model that related to the assigned and measured values, using a technique that considers the uncertainties and correlation on both the assigned and measured values, was adopted. The difference between the assigned and measured value for each material and the value predicted from the consensus model was estimated, considering the uncertainties on the definition of the model, as well as those on the observed values. The differences were then converted into degrees of equivalence.

2.0 STEP 1: DESIGN OF THE STUDY

2.1 Timeline

Table 1: Timeline

Date	Action
October 2016	Key/Pilot study design presentation at OAWG
March 2017	OAWG authorization; protocol approved
01 April 2017	Call for participation to OAWG members
07 July 2017	Study sign up deadline
04 August 2017	CRM unit collection deadline
31 December 2017	Measurement campaign completed
01 October 2018	Submission of Draft A Report
01 April 2019	Submission of Draft B Report

2.2 Participating Institutes (PIs)

Table 2: Participating Institutes

Acronym	Participating Institute	Country	Remarks
CENAM	Centro Nacional de Metrología	México	Niacinamide Reference Value
KRISS	Korea Research Institute of Standards and Science	Korea	Niacin Reference Value
NIM	National Institute of Metrology	China	Niacin and Niacinamide Reference Values
NIST	National Institute of Standards and Technology	USA	Niacinamide Reference Value

2.3 Materials

Milk powder and infant/adult formula-based materials with valid certified values and uncertainties were eligible for inclusion in CCQM-K147. To ensure that the required repeatability measurements could be made on at least three units of each material,

the participating institutes (PIs) were asked to provide four units of each material. Table 3 summarizes the certification information as provided by the participating institutes for the niacinamide and niacin materials.

For each material, Table 3 lists the PI, the certified value "*V*," the uncertainty on the certified value "$U_{95}(V)$" at a 95 % level of confidence, and the units of certification, the table also lists the auxiliary information deemed useful for evaluating the materials' suitability for inclusion in the comparison and for the measurement design.

Table 3: CRMs from participants

NMI	ID	Name	Matrix	g	Fat content (%)	Vitamin	Certified value *V*(mg/kg)	$U_{95}V$ (mg/kg)
CENAM	DRM-486b	Leche semidescremada en polvo	Milk powder	130	14	Niacinamide	**4.51**	0.22
CENAM	DMR-274g	Leche entera en polvo	Milk powder	130	26	Niacinamide	**5.52**	0.26
NIST	SRM 1549a	Whole Milk Powder	Milk powder	10	30	Niacinamide	**5.91**	0.39
CENAM	DMR-82c	Leche descremada en polvo	Milk powder	130	0.7	Niacinamide	**8.83**	0.41
NIM	GBW(E)100227	Vitamin B1, B2, Nicotinic acid and 9 inorganic Elements in Infant formula	Infant formula, milk based	2	25	Niacin	**39.8**	2.7
KRISS	108-02-003	Infant formula (for the analysis of organics nutrients)	Infant formula, milk based	1	24	Niacin	**60.6**	1.3
NIM	GBW10037	Nicotinamide in Infant Formula	Infant formula, milk based	1	24	Niacinamide	**65.0**	5.6
NIST	SRM 1869	Infant/Adult Nutritional Formula II	Infant formula, milk, whey and soy-based	10	19	Niacinamide	**97.0**	4.00
NIST	SRM 1849a	Infant/Adult Nutritional Formula	Infant formula, milk based	10	30	Niacinamide	**108.0**	10.0

Matrix is the form of the material; **g** is the content of material per unit

All materials were analyzed on an as-received basis. One of the materials, KRISS 108-02-003, is certified on a dry-mass basis. To assure that all observed values are compared on the same scale, moisture analysis was performed on units of this material after the ID-LC-MS/MS niacin/niacinamide determinations. The remaining material from each packet was weighed, dried for 2 h at 80 °C in a forced-air oven, and weighed again.

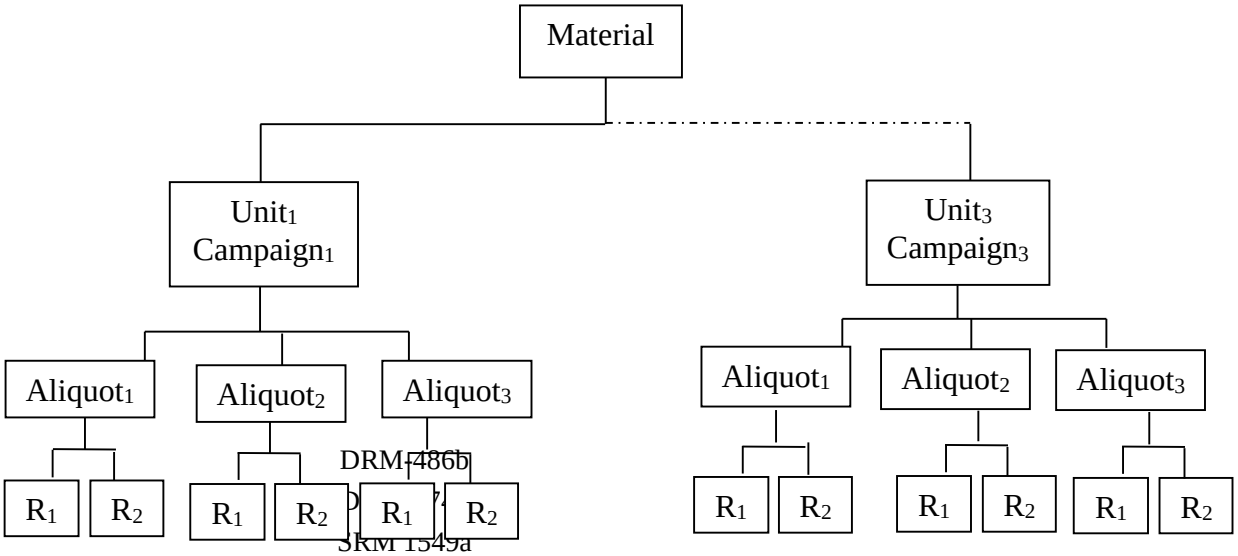
3.0 STEP 2: MEASUREMENTS

3.1 Measurement Design

According to the CCQM-K147 protocol, four units of each certified material were sent to NIST to participate in the comparison. At least three of the CRM units were used for the analysis. Each sample was prepared in triplicate and injected in duplicate. Given the number of materials, the time to receive them and the time required for each analysis, the measurements were made in three measurement campaigns (runs) conducted by the same analyst.

In the three campaigns, two measurements were made on three independently prepared aliquots from one randomly selected unit of each material. Figure 2 summarizes this three-level nested design. Campaigns did not match with current measurement dates for the materials.

Figure 2: Repeatability Measurement Design



Measurements on the comparison materials were performed following a randomized block design with blocking on aliquot and replicate. Quality control materials were interspersed at regular intervals in the measurements. All measurements within each campaign were made under repeatability conditions. No intentional changes were made to the equipment, reagents, or quality control materials between campaigns.

The above design confounds between-unit and between-campaign sources of measurement imprecision. Hence, the measurements made for this study cannot be used to estimate between-unit inhomogeneity for any of the study materials.

The following measurement equation describes the measurand:

$$Mass\ fraction = \frac{(A_{a,s})(A_{IS,c})(m_{IS,s})(m_{a,c})(p_a)}{(A_{IS,s})(A_{a,c})(m_{IS,c})(m_s)} \quad (1)$$

$A_{a,s}$ peak area of the analyte in the sample
$A_{IS,c}$ peak area of the internal standard in the calibrant
$m_{IS,s}$ mass of the internal standard in the sample
$m_{a,c}$ mass of the analyte in the calibrant
p_a purity of the analyte
$A_{IS,s}$ peak area of the internal standard in the sample
$A_{a,c}$ peak area of the analyte in the calibrant
$m_{IS,c}$ mass of the internal standard in the calibrant
m_s mass of the sample

Each observation has attached some characterization uncertainty according to this measurand model.

3.2 Analytical Method

All materials were analyzed under repeatability conditions using a validated ID-LC-MS/MS methodology [1]. This method was designed for analysis of niacin in Milk Powder and Infant Formula Matrices, with niacin levels of 4 mg/kg to 100 mg/kg. All materials were screened for the presence of interferences prior to quantitative analysis. Each measurement using this methodology required a minimum of 1.0 g of powder sample.

The Experimental details are provided in Appendix A. Quantification was based on the relative peak areas for niacin (m/z 124→80) and $^{2}H_4$-Niacin (m/z 128→84), and for niacinamide (m/z 123→80) and 2,4,5,6-$^{2}H_4$-nicotinamide (m/z 127→80). Tables 4, 4a, 5, 5a, 6 and 6a list all the measurement results for the CCQM-K147 materials.

Table 4: Vitamin B_3 as Niacinamide Measurements, Campaign 1

ID	Observed Value, Campaign 1					
	Value	Value	Value	Value	Value	Value
DMR-486b	4.449	4.433	4.083	4.554	4.646	4.210
DMR-274g	5.433	5.334	5.417	5.459	5.445	5.488
SRM 1549a	6.031	5.827	6.024	6.027	5.539	5.721
DMR-82c	8.724	9.207	8.916	8.853	8.952	8.636
GBW10037	65.32	64.56	65.48	65.25	63.64	64.71
SRM 1869	98.59	96.57	96.76	96.83	100.47	100.75
SRM 1849a	100.10	98.51	99.05	101.62	100.24	100.30

Table 4a: Vitamin B3 as Niacin Measurements, Campaign 1

ID	Observed Value, Campaign 1					
	Value	Value	Value	Value	Value	Value
GBW(E)100227	36.49	36.94	37.98	36.81	37.50	37.45
108-02-003	60.53	59.99	61.79	61.63	58.69	60.85

Table 5: Vitamin B_3 as Niacinamide Measurements, Campaign 2

ID	Observed Value, Campaign 2					
	Value	Value	Value	Value	Value	Value
DMR-486b	4.684	4.806	4.491	4.501	4.486	4.471
DMR-274g	5.598	5.352	5.893	5.692	5.282	5.450
SRM 1549a	5.592	5.986	6.221	5.981	5.096	4.983
DMR-82c	9.189	9.117	9.179	8.553	9.043	8.979
GBW10037	61.48	63.33	63.02	63.49	65.14	65.15
SRM 1869	98.84	98.59	96.48	94.42	95.75	97.97
SRM 1849a	100.33	100.29	101.97	99.54	101.29	101.78

Table 5a: Vitamin B_3 as Niacin Measurements, Campaign 2

ID	Observed Value, Campaign 2					
	Value	Value	Value	Value	Value	Value
GBW(E)100227	35.01	35.56	37.79	36.14	35.29	36.92
108-02-003	61.81	60.23	60.11	61.07	61.43	59.44

Table 6: Vitamin B_3 as Niacinamide Measurements, Campaign 3

ID	Observed Value, Campaign 3					
	Value	Value	Value	Value	Value	Value
DMR-486b	4.637	4.528	4.692	4.540	4.501	4.620
DMR-274g	5.575	5.436	5.609	5.586	5.189	5.554
SRM 1549a	5.369	5.394	5.973	5.909	5.556	5.762
DMR-82c	8.609	8.539	8.549	8.421	8.621	8.166
GBW10037	64.66	64.80	64.42	63.86	63.82	63.54
SRM 1869	100.27	96.67	97.02	97.21	100.69	98.80
SRM 1849a	99.89	99.20	98.11	98.74	100.74	99.48

Table 6a: Vitamin B_3 as Niacin Measurements, Campaign 3

ID	Observed Value, Campaign 3					
	Value	Value	Value	Value	Value	Value
GBW(E)100227	39.25	39.49	37.58	37.20	39.14	37.95
108-02-003	59.16	59.99	59.68	59.58	60.83	58.72

3.2.1 Measurement Quality Assurance

In addition to the measurements made on the CCQM-K147 materials, a control solution was analyzed at regularly spaced intervals within each campaign.

3.3 Frequentist Estimation of Value and Uncertainty

The three-level nested measurement design for the CCQM-K147 materials addresses instrumental, sample preparation and between-campaign sources of measurement variability by making two measurements on three independent aliquots of three different units of each material. The least complex model for describing measurements made using this design is:

$$R_{jkl} \sim N\left(\mu + \gamma_j + \delta_{jk}, \sigma_r^2\right)$$

where "~" indicates "is distributed as", N(p,q^2) defines a normal distribution with mean p and standard deviation q, j indexes the units, k indexes the aliquots, l indexes the replicates per aliquot, μ is the population mean, γ_j are between-campaign differences, δ_{jk} are between-aliquot differences, and σ_r is the limiting ID-LC-MS/MS imprecision for the material. The γ_j and δ_{jk} are assumed to be

$$\gamma_j \sim N\left(0, \sigma_c^2\right) \text{ and } \delta_{jk} \sim N\left(0, \sigma_a^2\right)$$

where σ_c reflects the true between-campaign and/or between-unit variability and σ_a reflects the true between-aliquot and/or within-unit variability.

3.3.1 Value

The repeatability measurement for each material, R, can be estimated as the mean of the individual measurements:

$$R = \frac{1}{(N_c \times N_a \times N_r)} \sum_{i=1}^{N_c} \sum_{j=1}^{N_a} \sum_{k=1}^{N_r} R_{ijk}$$

where N_c is the number of measurement campaigns, N_a is the number of aliquots taken from each campaign, and N_r is the number of replicates of each aliquot. For all niacinamide and niacin materials in CCQM-K147: N_c =3, N_a =3, and N_r = 2.

3.3.2 Measurement Standard Uncertainty

The usual estimate of the standard uncertainty of this mean is:

$$u_{\mathrm{DOE}}(R) = \sqrt{\frac{N_a \times N_r \times \sigma_c^2 + N_r \times \sigma_a^2 + \sigma_r^2}{N_c \times N_a \times N_r}}$$

This is recognized as the estimable uncertainty from the experimental design, and we can reduce it by increasing the sampling effort. In addition, characterization uncertainty related to the measurement model may remain constant regardless of the applied sampling effort; the total standard uncertainty is

$$u(R) = \sqrt{u_{\mathrm{DOE}}^2 + \sigma_{\mathrm{char}}^2}$$

The characterization uncertainty, σ_{char}, for each material was estimated at NIST based on the measurand model described by Eq. 1 using the NIST Uncertainty Machine [3].

The raw data were independently reanalyzed at CENAM and results compared to the original summary results. Typographical and formula errors were corrected and σ_{char} estimated using CENAM's own uncertainty estimation system [4], considering all the known sources of uncertainty including repeatability contributions. Uncertainty due to "inhomogeneity" was assessed and considered accordingly; no evidence was found for the presence of "heterogeneity" (confound between-campaigns and between-units uncertainty) in any of the materials; evidence suggests microheterogeneity (confound preparation and within-unit uncertainty) was present in three materials (SRM 1549a, SRM 1869, GBW 10037). Potential correlation in the response values was estimated based on Eq.1 and the known shared sources of uncertainty; actual correlation was assessed from observed values and no significant correlation was found. Certified values were considered independent. Uncertainty was estimated using GUM [5] techniques and confirmed by GUM-S1 techniques [6].

All other standard deviations were estimated from the data, most practically calculated with linear mixed model statistical analysis systems [2]. Tables 7 and 7a list the estimated standard deviation for niacinamide and niacin, expressed as percent relative values:

$$\%\sigma = 100 \times \sigma / R.$$

Tables 7 and 7a also list the relative standard uncertainties of the certified values expressed as percent:

$$\%u(V) = 100 \times u_{\infty}(V)/V$$

where $u_{\infty}(V)$ is the "large sample" standard uncertainty and is equal to one-half of the certified $U_{95}(V)$

$$u_{\infty}(V) = U_{95}(V)/2 \ .$$

Note that σ_{r} estimates are just the instrumental precision, independent of within- and between-unit sample preparation and/or heterogeneity issues. The pooled relative instrumental precision, $_{\%}\sigma_{\mathrm{r}}$, is 2.3 % for all the measurements. The $_{\%}\sigma_{\mathrm{a}}$ and $_{\%}\sigma_{\mathrm{c}}$ estimates are not easily interpreted since σ_{a} combines all aliquot preparation-related differences with within-unit heterogeneity, σ_{c} combines all time-related differences in the method with between-unit heterogeneity. All other uncertainty sources involved in the measurement model are combined in σ_{char}.

3.3.3 Large-Sample Standard Uncertainties

Ideally the $u(R)$ should be representative of the material rather than just the specific units of the material used in the study. As discussed in [2], one approach to accomplishing this is to first expand the estimated standard uncertainty by the appropriate two-tailed Student's t 95% level of confidence factor

$$U_{95}(R) = \sqrt{\left(t_{0.05,\nu} \times u_{\mathrm{DOE}}(R)\right)^2 + (2\sigma_{\mathrm{char}})^2}$$

where ν is the number of degrees of freedom associated with $u(R)$, and then divide the expanded uncertainty by the conventional metrological large-sample coverage factor of 2, giving a "large sample" standard uncertainty:

$$u_{\infty}(R) = U_{95}(R)/2$$

Unfortunately, determining ν is problematic. The usual interpretation of the analysis of variance (ANOVA) results presented in Tables 7 and 7a provide $\nu = N_r \times N_a \times N_c - 1 = 17$ when both σ_a and σ_c are statistically insignificant (here, when $_{\%}\sigma_a$ and $_{\%}\sigma_c$ are zero), $\nu = N_a \times N_c - 1 = 8$ when just σ_c is insignificant, and $\nu = N_c - 1 = 2$ when σ_c is significant (here, when $_{\%}\sigma_c$ is greater than zero). Under this interpretation, $t_{0.05,\nu} / 2$ for the different materials is ≈1.05 when ν is 17, ≈1.15 when ν is 8, and ≈2.15 when ν is 2.

This interpretation considers both the evidence of the measurements and the information about the materials inherent in the uncertainty assigned to the certified values, $u(V)$. For 6 out of the 9 materials, the estimated $_{\%}u(R)$ is less than the certified $_{\%}u(V)$ suggesting that any additional within- and between-unit heterogeneity sources of variability were recognized and accounted for during certification. The estimated $_{\%}u(R)$ is greater than the certified $_{\%}u(V)$ only for 2 materials, DMR-82c and 108-02-003. This suggests that certified uncertainty for these materials may be underestimated. Expanding the $u(R)$ to be greater than $u_{\infty}(V)$ for these materials yields $u_{\infty}(R)$ that are unreasonably large.

For the frequentist analysis discussed in Section 4, based on this insight we assert that the "real" ν for all the materials is "large" and therefore:

$$u_{\infty}(R) \cong u(R).$$

Table 7: Measurement Summary for Vitamin B_3 (as niacinamide) *

PI	CRM	R	$\%\sigma_r$	$\%\sigma_a$	$\%\sigma_c$	$\%\sigma_{char}$	$\%u(R)$	$\%u(V)$
CENAM	DMR-486b	4.52	3.7	0	0	2.01	2.2	2.4
CENAM	DMR-274g	5.49	2.9	0	0	2.00	2.1	2.4
NIST	SRM 1549a	5.72	2.4	2.7	0	2.10	2.8	3.3
CENAM	DMR-82c	8.79	2.3	0	1.9	2.05	3.3	2.3
NIM	GBW 10037	64.3	0.9	0.6	0	2.00	2.1	4.3
NIST	SRM 1869	97.7	1.32	0.68	0	2.06	2.1	2.1
NIST	SRM 1849a	99.8	1.1	0	0	2.00	2.1	4.6

Table 7a: Measurement Summary for Vitamin B_3 (as niacin) *

PI	CRM	R	$\%\sigma_r$	$\%\sigma_a$	$\%\sigma_c$	$\%\sigma_{char}$	$\%u(R)$	$\%u(V)$
NIM	GBW (E) 100227	37.3	2.4	0	1.9	1.80	3.2	3.4
KRISS	108-02-003	60.3	1.7	0	0	1.80	1.9	1.1

* Table Legend

R Mean of repeatability measurements, arbitrary units

$\%\sigma_r$ Relative within-replicate precision, expressed as % of R

$\%\sigma_a$ Relative between-aliquot precision, expressed as % of R

$\%\sigma_c$ Relative between-campaign precision, expressed as % of R

$\%\sigma_{char}$ Relative characterization precision, expressed as % of R

$\%u(R)$ Relative standard uncertainty of measurements, expressed as % of R

$\%u(V)$ Relative standard uncertainty of certification, expressed as % of V

3.3.4 Data Used in the Frequentist Analyses

Table 8 summarizes the certified values and measured values for the study materials used in the frequentist analysis of niacinamide and niacin. In this Table, the materials are sorted in order of increasing certified value, V. Each material is assigned a one-character identifying code to simplify graphical presentation.

Table 8: Data Used in the Analysis of Vitamin B_3 Materials

Code	NMI	ID	Analyte	Certified Value mg/kg		Observed Value mg/kg	
				Value	U95%	Value	U95%
A	CENAM	DMR-486b	Niacinamide	4.51	0.22	4.52	0.20
B	CENAM	DMR-274g	Niacinamide	5.52	0.26	5.49	0.23
C	NIST	SRM 1549a	Niacinamide	5.91	0.39	5.72	0.32
D	CENAM	DMR-82c	Niacinamide	8.83	0.41	8.79	0.58
E	NIM	GBW(E)100227	Niacin	39.8	2.7	37.3	2.36
F	KRISS	108-02-003	Niacin	60.6	1.3	60.3	2.26
G	NIM	GBW10037	Niacinamide	65.3	5.6	64.3	2.68
H	NIST	SRM 1869	Niacinamide	97	4	97.7	8.36
I	NIST	SRM 1849a	Niacinamide	108	10	99. 8	8.18

3.4 Bayesian Estimation

Bayesian analysis is based on a somewhat different definition of probability than the usual frequentist interpretation underpinning classical statistical inference. Under the Bayesian paradigm, parameters such as the measurand value and variance components have probability distributions that quantify our knowledge about them. The estimation process starts with quantification of prior knowledge about the parameters followed by specification of the statistical model that relates the parameters to the data.

The components of the model specified in Section 3.3 are combined via Bayes Theorem to obtain posterior distributions for the parameters. These distributions update our knowledge about the parameters based on the evidence provided by the data. This analysis can produce a probability distribution for each μ (the true value of analyte quantity estimated by the measurement mean, R) which encompasses all information and variability present in the data but is confined by bounds based on prior knowledge. The process yields a probability interval which is interpretable as an uncertainty interval. Markov Chain Monte Carlo (MCMC) empirical Bayesian methods enable computation of coverage intervals. The OpenBUGS software system [7] that implements this analysis is freely available and (relatively) easy to use.

Ideally, Bayesian analysis can proceed using very conservative, minimally-informative priors (e.g., very broad Gaussian distributions) and let the data mostly determine the posterior distribution of the measurand. Unfortunately, somewhat informative priors are required with small degrees of freedom. However, when these priors are carefully defined the analysis can validly produce probability distributions for the μ which encompass the available information on the materials and all the variability present in the data.

Code provided from the previous OAWG Track A Model 2 comparison CCQM-K142 was reused; modifications were required to account the characterization uncertainty from the measurand model in addition to the variance structure from the experimental design.

3.4.1 Differences Between the Implementations

Based again on the insight that the "real" ν for the niacinamide/niacin materials is "large", the Bayesian OpenBUGS codes developed for this study assign an informative prior to each material's between-unit/campaign standard deviation, σ_c. For all materials where $u_\infty(V)$ is as large or larger than the ANOVA estimate for $u_\infty(R)$, the prior is $u_\infty(V)$.

The frequentist ANOVA analysis estimates a different σ_r, for every material. However, the relative estimates, ${}_{\%}\sigma_r = 100\times\sigma_r/R$, are approximately constant for both

measurands. Based on this observation, for each measurand the OpenBUGS codes estimate a common $_{\%}\sigma_r$ for all materials.

3.4.2 Data Used in the Bayesian Analyses

The Bayesian OpenBUGS codes developed for this study use the V and $u_\infty(V)$ values listed in Tables 7, 7a and 8 and the "raw" measurement results listed in Tables 4, 4a, 5, 5a. 6 and 6a. The complete OpenBUGS code and data for niacinamide/niacin materials are listed in Appendix B.

4.0 STEP 3: DEFINE A CONSENSUS MODEL

4.1 The Key Comparison Reference Function (KCRF)

In analogy to the "Key Comparison Reference Value (KCRV)" used with Model 1 comparisons, whatever approach is used to characterize the relationship between the certified values, $V \pm u_\infty(V)$, and the measured summary values, $R \pm u_\infty(R)$, we term the "Key Comparison Reference Function (KCRF)" for the comparison.

Since a definitive method was used for the measurements, a linear relationship is expected between the certified and measured values. Figure 3 provides an overview of the relationship between the certified and measurement values of niacin/niacinamide. The closeness of the values to the lines confirms that the relationship for the measurand, and thus its KCRF, is assumed linear.

4.1.1 Linear Models

A linear relationship can be modelled as:

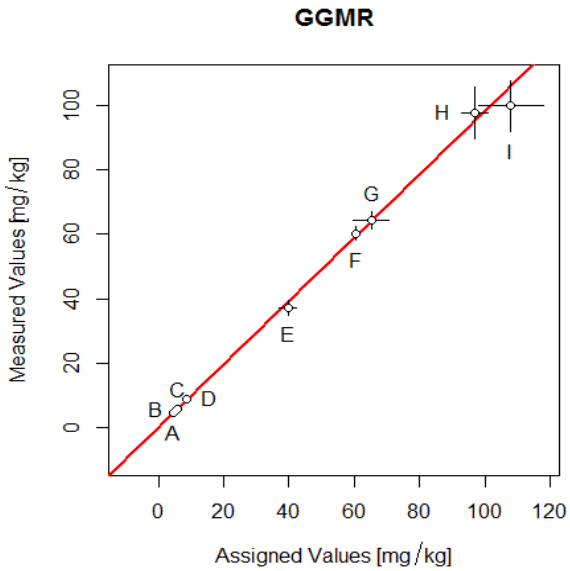
$$R = \alpha + \beta V + E \qquad (2)$$

where α is the intercept, β is the slope, and E is the residual random error. Alternatively, if α is asserted to be zero, then the relationship can be modelled as:

$$R = \beta V + E \qquad (3)$$

The number of degrees of freedom for the model is the number of materials used to parameterize to model, N_m, minus the number of adjustable parameters in the model. Two such parameters are needed for Eq. 2, α and β; only one, β, is needed for Eq. 3. In consequence, if α is truly zero then the uncertainty in the estimate of β should be smaller using Eq. 3 rather than Eq. 2. However, should α be erroneously asserted to be zero then use of Eq. 3 will result in a biased model.

Figure 3: Certified Vs Measured Values



Each cross denotes the {$V \pm 2\times u_\infty(V)$, $R \pm 2\times u_\infty(R)$) for one material. The line represents exact equality between the certified and measured values: $R = V$. The crosses are labelled in order of increasing V. The materials are labeled from A to I; A, B, C, D, G, H and I correspond to niacinamide, while E and F correspond to niacin. Refer to Table 8 for the association between the code and the materials.

4.1.2 Generalized Gauss Markov Regression (GGMR)

Ordinary least squares regression is not an appropriate approach to estimating the parameters of Equations 2 or 3 since both the certified and measurement results have known and non-negligible uncertainty [2]. However, generalized distance regression (GDR) and generalized Gauss Markov regression [8] (GGMR) provide appropriate parameters by iteratively minimizing the total uncertainty-scaled residual distances. The GGMR is a generalization of GDR in the sense that it allows for covariance among the observed values, among the reference values and across reference and observed values. Since there was some caution about correlation among the observed values, GGMR was selected from the beginning. The sum of square errors to be minimized is:

$$E = \sum_{i=1}^{N_\mathrm{m}} \varepsilon_i^2; \qquad \varepsilon_i^2 = \left(\frac{V_i - \widehat{V}_i}{u_\infty(V_i)}\right)^2 + \left(\frac{R_i - \widehat{R}_i}{u_\infty(R_i)}\right)^2; \qquad \widehat{R}_i = \hat{\alpha} + \hat{\beta}\widehat{V}_i$$

where i indexes the materials, N_m is the number of materials, and $\widehat{V}_i$, $\widehat{R}_i$, $\hat{\alpha}$, and $\hat{\beta}$ are predicted estimates for the parameters. Note that the residual uncertainty-weighted distance for a given material, ε_i is symmetric in V_i and R_i.

There are several available Frequentist implementations of GDR [2]. In this report, these results were obtained using CENAM's own GGMR system [9]. There is at least

one available implementation of GGMR [10]. Further analysis using NIST's RegViz [11] system was not pursued.

4.1.3 Parametric Bootstrap Monte Carlo Uncertainty Evaluation

Like the RegViz system, CENAM's GGMR system incorporates a parametric bootstrap Monte Carlo (PBMC) technique that facilitates the estimation of the variability for all quantities estimated with GDR. With PBMC, the entire set of V_i and R_i values used in the GDR/GGMR analysis are repeatedly replaced with corresponding "pseudo-values" randomly drawn from each of the $\mathrm{N}\left(V_i, u_\infty^2(V_i)\right)$ and $\mathrm{N}\left(R_i, u_\infty^2(R_i)\right)$ normal kernels. The parameters and associated quantities are stored and, once a suitably large number have been generated, approximate 95% expanded uncertainty intervals are estimated from the percentiles of the empirical distributions. Since only the central 95% of the distributions are of interest, relatively few pseudo-sets are required for stable estimates.

4.1.4 Bayesian GDR

The OpenBUGS Bayesian codes were reused from CCQM-K142 and modified for this project, treating both the *V* and *R* values as distributions rather than fixed values. As such, they inherently produce result distributions that can be summarized as GDR parameter and parameter uncertainty estimates.

4.2 Graphical Analyses Using the CENAM's GGMR System

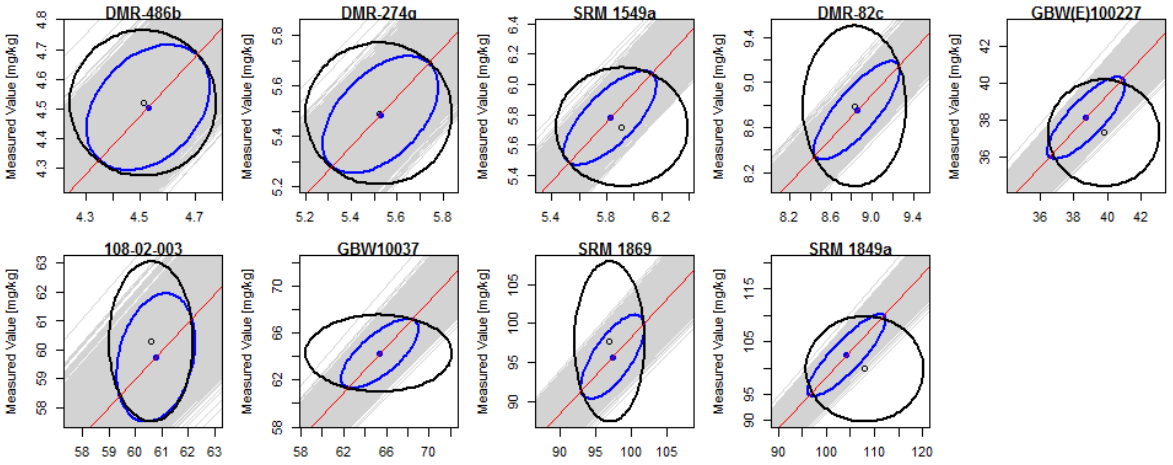
4.2.1 Overview

Figure 4 displays summary GGMR results for all the materials. Panel A displays the results based on the $R = \alpha + \beta V$ model (non-zero intercept model) and panel B displays the results for $R = \beta V$ model (zero intercept model). The graphical resolution required for simultaneously displaying all materials in single scatterplot is insufficient for adequately visualizing the analyses. Figure 4 displays each material in a series of high-resolution individual "thumbnail" scatterplots.

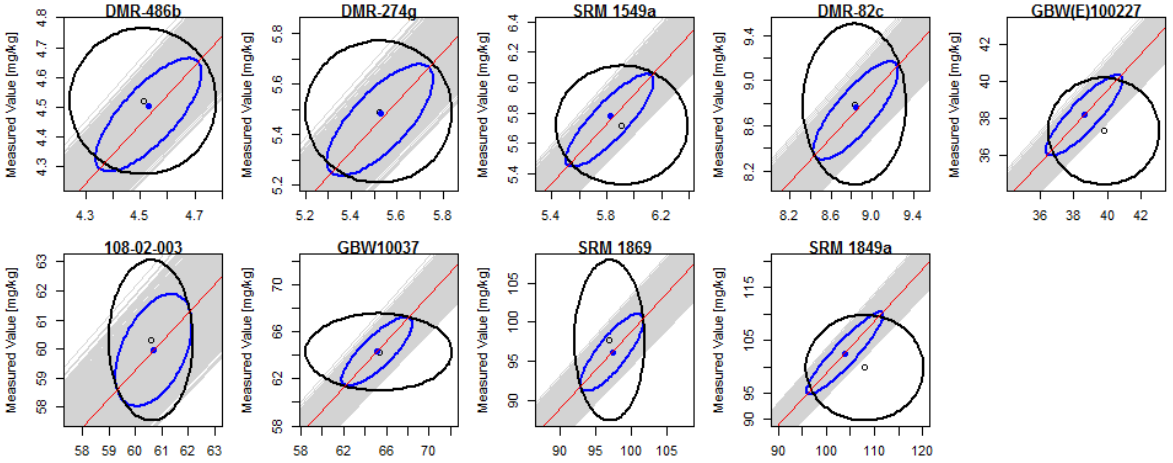
The black ellipses bound all points that are within the 95 % confidence region around the $\{V_i, R_i\}$. The square of the GGMR uncertainty-weighted residuals, ε_i^2, are expected to be distributed as χ^2 with two degrees of freedom. Therefore, ε_i less than the value for the 95th percentile expected from this distribution, $\sqrt{5.99} \approx 2.45$, indicate that the uncertainties adequately cover the difference between the estimated and observed values at the 95 % level of confidence. The best estimates for the corresponding points on the KCRF are known with uncertainty; this is represented by the blue ellipses. Ellipses (black and blue) that overlap mutually suggest that the observed values $\{V_i, R_i\}$ are consistent with the KCRF. Ellipses that do not substantially overlap mutually suggest that the observed values $\{V_i, R_i\}$ are not consistent with the KCRF.

Figure 4: GGMR Results for Vitamin B_3 Materials

A: *R = α+βV*



B: *R = βV*



Each thumbnail plots the certified value of a given material along the horizontal axis and the results of the repeatability measurements along the vertical. Each thumbnail is centered at $\{V_i, R_i\}$. The thumbnails are arranged in order of increasing Vi. All thumbnails for a given measurand have the same relative scale. The red lines represent the candidate KCRF. The gray lines are approximate 95% level of confidence intervals on the candidate KCRF, U_{95}(KCRF). As expected, the KCRF±U_{95}(KCRF) confidence regions are somewhat narrower for the R = βV models.

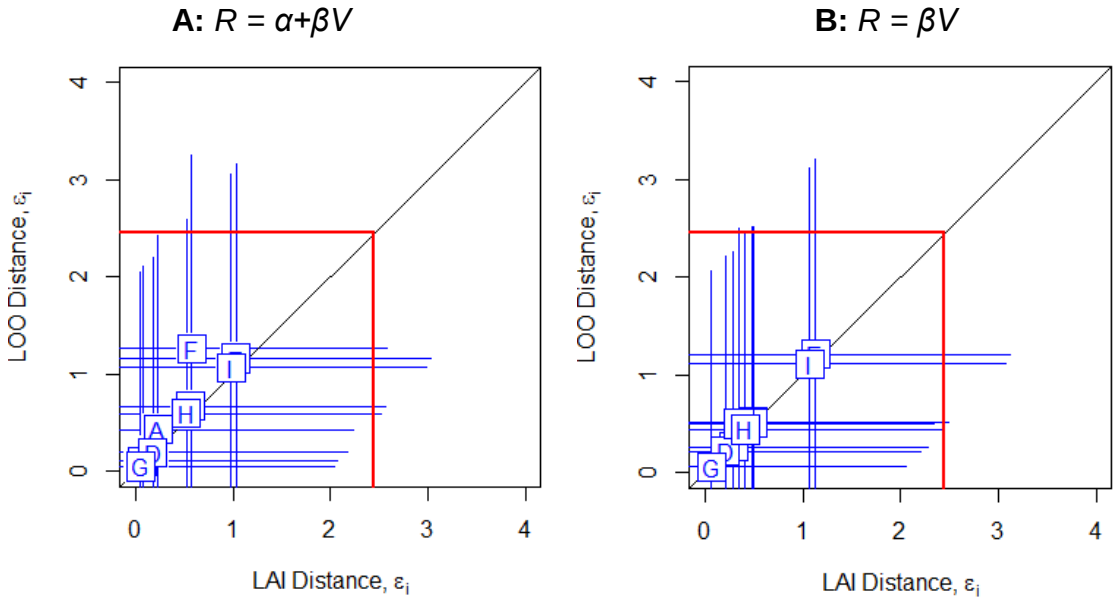
4.2.2 Identifying Influential Materials

The GGMR solution can be strongly influenced by materials having small $u_\infty(V_i)$ and/or $u_\infty(R_i)$. The magnitude of this influence depends not only on the magnitudes of the uncertainties but also on where the $\{V_i, R_i\}$ pair is located relative to the other materials.

Leave-one-out (LOO) validation is an efficient approach to establishing which, if any, materials are sufficiently influential to distort the consensus estimation of the candidate KCRF. A LOO analysis proceeds by excluding each material in turn from its own

evaluation. For all the materials, this involves 10 GGMR analyses: one solution with all 9 materials included in the analysis and 9 solutions each with one material excluded. Figure 5 displays the "exact" ε_i, calculated using all the materials with their LOO-estimated analogues. The A panel displays results for the $R = \alpha+\beta V$ model and the B panel displays results for the $R = \beta V$ model.

Figure 5: Strongly Influential Vitamin B_3 Materials



The open squares represent estimates for individual materials; the crosses represent the PBMC-estimated 95% level of confidence intervals on the estimates. Results inside the red lines indicate materials that are consistent with the consensus GDR/GGMR solution. The diagonal line represents equality between the two estimates. Results far from the diagonal line indicate materials that strongly influence the consensus solution. Results near or beyond the red square border indicate materials that are nearly or strongly inconsistent with the GDR/GGMR solution. A, B, C, D, G, H and I correspond to niacinamide, while E and F correspond to niacin.

Since there is no evidence of any material being strongly influential for the consensus value, the LAI solution is preferred.

4.3 Parameter Values for the Candidate KCRFs

In addition to identifying materials that could distort the consensus GDR/GGMR solution, LOO-PBMC enables a more robust estimate of the variability of the GDR/GGMR parameters. The LOO estimates are influenced by biases (systematic differences in the GDR/GGMR solutions with-and-without each material in the models) that are not present when all materials are included ("Leave-All-In" or "LAI") in the model. Thus, the LOO-PBMC parameter uncertainties are constrained to be somewhat larger than those determined with LAI-PBMC analysis. Another probable scenario is that all the materials may be consistent, hence the solutions with-and-without each material in the model will lead to parameter estimates with a reduced variability when compared to the estimates obtained by including all the materials in

the model. In this case, uncertainty is underestimated due to over sampling from the same cases and LAI-PBMC estimates may be preferable.

Table 9 lists the consensus solution parameters for all the materials based on the $R = \alpha + \beta V$ and $R = \beta V$ models as estimated using the frequentist approach and the Bayesian OpenBUGs systems. The slightly smaller LOO-based asymptotic standard uncertainty estimates obtained from the frequentist approach provide more optimistic coverage than do the LAI-based estimates. The OpenBUGs implementations that were developed for this study do not provide LOO estimates.

Table 9: Model Parameter Estimates

			$u(\hat{\alpha})^a$			$u(\hat{\beta})^b$	
Model	Method	$\hat{\alpha}^a$	LAI[c]	LOO[d]	$\hat{\beta}^b$	LAI[c]	LOO[d]
$R = \alpha+\beta V$	GGMR	0.17	0.28	0.16	0.978	0.035	0.020
	BUGs	0.04	0.11		0.988	0.012	
$R = \beta V$	GGMR				0.987	0.012	0.012
	BUGs				0.990	0.0079	

a Intercept and its uncertainty estimates are expressed in mg/kg
b Slope and its uncertainty estimates are of dimension 1
c Standard deviation of Leave-All-In PBMC parameter estimates where all eligible materials were included in model
d Standard deviation of Leave-One-Out PBMC parameter estimates where one eligible material is in turn left out of the model in each set

Since there is no evidence for the presence of a non-zero intercept (α) in the model, i.e., its confidence interval does contain zero. Hence the zero intercept LAI model is preferred. GGMR and BUGS estimates are virtually the same (the respective confidence intervals overlap), however BUGS estimates intrinsically consider correlations among the input data and the β coefficient, which will be required for estimating the degrees of equivalence. Using the BUGS parameter estimates and the LAI estimates of parameter uncertainty, the candidate KCRF for all the materials is:

- $\hat{R} = (0.990 \pm 0.0079)\hat{V}$

4.4 BUGS Predicted Values

Table 10 lists the Bayesian approach estimates for the assigned values and repeatability measurements along with their LAI-estimated asymptotic standard uncertainties for all the materials.

Table 10: BUGS Predicted Values for Vitamin B_3 Materials, these represent the final KCRF values for the comparison.

Code	CRM	mg/kg		mg/kg	
		$\hat{V}_i$	$u(\hat{V}_i)$	$\hat{R}_i$	$u(\hat{R}_i)$
A	DMR-486b	4.55	0.10	4.50	0.10
B	DMR-274g	5.52	0.13	5.47	0.12
C	SRM 1549a	5.91	0.17	5.85	0.16
D	DMR-82c	8.94	0.22	8.81	0.21
E	GBW(E)100227	37.84	0.81	37.51	0.82
F	108-02-003	60.6	1.2	60.2	1.2
G	GBW10037	65.3	1.5	64.6	1.4
H	SRM 1869	98.4	2.2	97.3	2.2
I	SRM 1849a	101.3	2.3	100.3	2.2

5.0 STEP 4: DEGREES OF EQUIVALENCE

5.1 Degrees of Equivalence for Materials

An appropriate definition for the degrees of equivalence for materials in the present comparison is the percent relative signed orthogonal distance [1]:

$$\%d_i = 100 \times SIGN\left(V_i - \hat{V}\right) \times \frac{\sqrt{\left(V_i - \hat{V}\right)^2 + \left(\left(R_i - \hat{R}_i\right)/\hat{\beta}\right)^2}}{\left(V_i + R_i/\hat{\beta}\right)/2}$$

where the measurement-related terms are transformed to have the same scale as the assigned values. The function SIGN returns the sign (±1) of its argument and defines whether the observed $\{V_i, R_i\}$ pair is "above" or "below" the candidate KCRF. In practice $\left(V_i, R_i, \hat{V}, \hat{R}, \hat{\beta}\right)$ are not independent and their covariance must be considered in estimating the uncertainty of $\%d_i$.

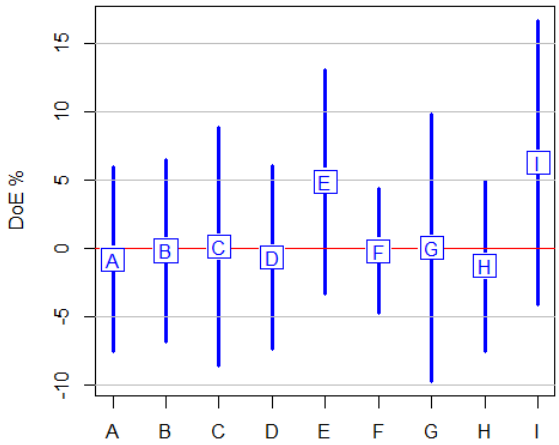
5.1.1 Degree of Equivalence Uncertainty for Individual Materials

The $d_i \pm U_{95}(\%d_i)$ can be estimated from the empirical distribution of the $\%d_i$values calculated for each set of PBMC pseudo-values, using the LAI analysis to make the uncertainty estimates less robust to each material's "self-referential" influence. The $U_{95}(\%d_i)$ for each material can be estimated from the distribution of the $\%d_i$ calculated when its own values are used in the KCRF solution. While requiring less calculations, these LAI-PBMC estimates are not free of correlation between each material's observed values and the candidate KCRF.

5.1.2 Graphical Representation of Degrees of Equivalence for Materials

Figure 6 displays OpenBUGS estimates of $\%d_i \pm U_{95}(\%d_i)$ for the niacin/niacinamide materials in dot-and-bar format. The red line denotes zero bias relative to the KCRF; the $\%d_i$ for materials with bars that cross this line are consistent with the consensus model with about a 95% level of confidence. The horizontal axis in these figures displays the V_i of each material.

Figure 6: DoE for Vitamin B_3 Materials



A, B, C, D, G, H and I correspond to niacinamide, while E and F correspond to niacin.

5.2 Degrees of Equivalence for Participating Institutes

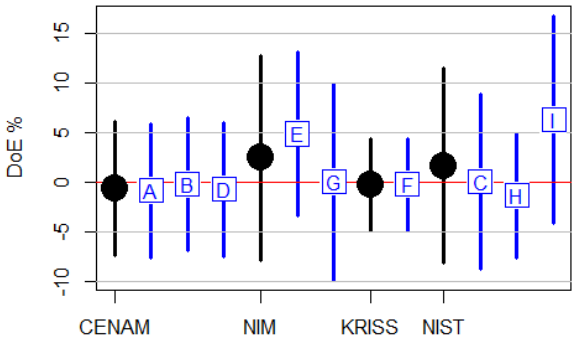
All the PIs in CCQM-K147 are represented by one or more materials. The results for all the materials from each PI contributing more than one material must therefore be combined in some way to provide the desired goal of the comparison: the expected degrees of equivalence of the PIs, $\%D$.

For the BUGS estimates, the $\%D$ are estimated from the median and empirical 95 % confidence interval of the probability density function produced by combining the $N(\%d_i, (U_{95}(\%d_i)/2)^2)$ kernels of each material. This method is described as the "Mixture Model Median" in [10] and the "Linear Pool" in [3,11].

5.2.1 Graphical Representation of Degrees of Equivalence for PIs

Figure 7 displays the OpenBUGS estimates of $\%D \pm U_{95}(\%D)$ and $\%d_i \pm U_{95}(\%d_i)$ in dot-and-bar format for niacin/niacinamide. The thick black bars and black solid dots represent the $\%D$ and thin blue bars and blue open dots the $\%d_i$. The PIs are arranged in alphabetical order.

Figure 7: DoE for PIs That Submitted Vitamin B_3 Materials, Zoom in



A, B, C, D, G, H and I correspond to niacinamide, while E and F correspond to niacin.

5.3 Tabular Presentation of Degrees of Equivalence

Table 11 lists the estimates of the degrees of equivalence for each material and for their submitting PIs, considering correlation between the observed values and the KCRF is present.

Table 11: Vitamin B_3 DoEs

PIs				Materials						
	%*D*, percent					$\%d_i$, percent			V_i	
PI	Value	*u*	U_{95}	Material	Code	Value	*u*	U_{95}	mg/kg	$\%U_{95}$[a]
CENAM	-0.7	3.4	6.7	DMR-486b	A	-1.0	3.3	6.7	4.51	4.88
				DMR-274g	B	0.0	3.3	6.7	5.52	4.71
				DMR-82a	D	-1.3	3.4	6.7	8.83	4.64
NIM	2.5	4.9	10.3	GBW(E)100227	E	5.0	4.0	8.2	39.8	6.8
				GBW 10037	G	0.0	4.9	9.8	65.3	8.6
KRISS	0.0	2.3	4.6	108-02-003	F	0.0	2.3	4.6	60.6	2.1
NIST	1.7	4.9	9.8	SRM 1549a	C	0.0	4.4	8.7	5.91	6.60
				SRM 1869	H	-1.4	3.1	6.3	97.0	4.1
				SRM 1849a	I	6.3	5.2	10.4	108.0	9.3

[a] Percent relative expanded uncertainty, $100\% \times U_{95}(V_i)/V_i$

A, B, C, D, G, H and I correspond to niacinamide, while E and F correspond to niacin.

5.4 Choice of Model for Degrees of Equivalence

GGMR approach for $d_i \pm U_{95}(\%d_i)$, and $\%D \pm U_{95}(\%D)$ provides asymptotic uncertainty estimates (assumes an infinite experimental effort), while PBMC approach intrinsically considers the relatively small applied effort and the covariance presence. The selected version of the DoE estimates is the BUGS estimates, shown in Table 11.

The GGMR and BUGS systems provide essentially equivalent KCRFs,$d_i \pm U_{95}(\%d_i)$, and $\%D \pm U_{95}(\%D)$ values for $R = \beta V$ model. While relatively unfamiliar within the chemical metrology community, the Bayesian approach implemented by the BUGS models is statistically sound, explicitly identifies its assumptions, facilitates exploring those assumptions, is computationally efficient, and can performed using freely accessible and well-supported software.

The more familiar frequentist approach introduced to the OAWG in the previous Excel-based RegViz system is now implemented in the GGMR system. GGMR supports specialized data visualization tools, is computationally efficient, is implemented in the freely accessible R language [12], and is supported by a group of users at CENAM. CENAM and NIST encourages the use of the Bayesian approach for estimating parameters of interest in this and future Model 2 comparisons either as the main

estimation method or as a validating method. In this case of CCQM_K147 the BUGS data in Table 11 are the final DoE values for the comparison.

5.5 CONCLUSIONS

Generalized Gauss Markov Regression (GGMR) and Bayesian statistics were used to establish the Key Comparison Reference Function (KCRF) relating the CRM certified values to the repeatability measurements. The niacinamide and niacin results for all nine CRMs were deemed equivalent at the 95 % level of confidence and were used to define the KCRF for vitamin B_3 (as niacinamide).

Monte Carlo methods based on GGMR and Bayesian methods were used to estimate 95 % level-of-confidence coverage intervals for the relative degrees of equivalence of materials, $\%d \pm U_{95}(\%d)$, and of the participating NMIs/DIs, $\%D \pm U_{95}(\%D)$. The GGMR method alone does not provide a straight way to estimate the uncertainty of the degrees of equivalence hence the Bayesian method was selected as the formal data for the comparison.

For the niacinamide and niacin materials, all of the $\%D \pm U_{95}(\%D)$ intervals, were within (-10 to 10) % of the consensus results and all of these are statistically equivalent. These results of CCQM-K147 Comparison of value-assigned CRMs, Track A Model 2, demonstrate that the participating institutions can value-assign CRMs for niacinamide and/or niacin in milk powder and infant formula matrices.

6.0 REFERENCES

1 Phillips, M. M. Liquid chromatography with isotope-dilution mass spectrometry for determination of water-soluble vitamins in foods. Anal Bioanal Chem. 2015 Apr;407(11):2965-74. doi: 10.1007/s00216-014-8354-y.

2 Duewer DL, Gasca-Aragon H, Lippa KA, Toman B. Experimental design and data evaluation considerations for comparisons of reference materials. Accred Qual Assur 2012;17:567-588. https://doi.org/10.1007/s00769-012-0920-4

3 NIST Uncertainty Machine (NUM). https://uncertainty.nist.gov/

4 Gasca, A. Hugo. (2017). Sistema para Evaluación de Incertidumbre, versión 3. Centro Nacional de Metrología. Registro Público del Derecho de Autor: 03-2018-012512305800-01. URL www.cenam.mx/incertidumbre

5 JCGM 100:2008. Evaluation of measurement data — Guide to the expression of uncertainty in measurement https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

6 JCGM 101:2008. Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method https://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf

7 Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing, 10:325-337. URL: http://www.openbugs.net/w/FrontPage

8 ISO/TS 28037:2010 (E). Determination and use of straight-line calibration functions. Geneva

9 Available on request from hgasca@cenam.mx

10 NPL Mathematics and Modelling for Metrology (2014). Software to Support ISO/TS 28037: 2010 (E) http://www.npl.co.uk/science-technology/mathematics-modelling-and-simulation/mathematics-and-modelling-for-metrology/software-to-support-iso/ts-28037-2010(e)

11 Available on request from david.duewer@nist.gov

12 Duewer DL. A comparison of location estimators for interlaboratory data contaminated with value and uncertainty outliers. Accred Qual Assur 2008;13:193–216. https://doi.org/10.1007/s00769-008-0360-3

13 Toman B. Bayesian Approach to Assessing Uncertainty and Calculating a Reference Value in Key Comparison Experiments. J Res Natl Inst Stand Technol 2005;110(6):605–612. https://doi.org/10.6028/jres.110.085
See also: Koepke A, LaFarge T, Possolo A, Toman B. NIST Consensus Builder User's Manual. May 13, 2017. https://consensus.nist.gov/NISTConsensusBuilder-UserManual

14 R Core Team (2018). R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org

APPENDIX A: SOURCES OF INFORMATION

A.1 Reagents and Materials

Niacinamide (Lot #N0E024) was obtained from the U.S. Pharmacopeia (Rockville, MD). Purity of these reference standards has been evaluated in ROA 839.02-10-014 and ROA 02-17-029 by LC-UV, DSC and qNMR, respectively. All mass fractions reported herein have been corrected for the purity of the reference standard. Niacinamide (2,4,5,6-$^{2}H_{4}$, Lot #DS2-2005-202A1) was obtained from Isosciences (King of Prussia, PA). All solvents used were HPLC grade. All other salts and acids used in sample and mobile phase preparation were reagent grade.

A.2 Calibration and Internal Standard Solutions

All stock, mix, diluted, and calibration solutions were prepared gravimetrically in 0.1 mol/L ammonium acetate in water, adjusted to pH 2.6 with hydrochloric acid. All solution preparation was conducted under reduced lighting to minimize potential vitamin degradation and stored in the refrigerator (4 °C) when not in use.

Mass fractions of niacinamide in the samples were bracketed with calibration solutions. The analysis set contained two or three injections of each of calibration solutions A and B, two injections of each of three subsamples from each unit of CCQM-K147 samples, two injections from DMR-486b as a control and two injections of SRM 1869 as a control, and three blanks (mobile phase).

A response factor was calculated for the main transition and an average response factor of at least three calibrant solutions was determined considering at least 6 injections per day for each one.

Approximately 5 mg of niacinamide and 1 mg of niacin were weighed into amber vial and about 25 mL of 0.1 mol/L ammonium acetate was added to the vial. For the internal standards, approximately 5 mg of labeled niacinamide and 5 mg of labeled niacin were weighed into amber vials and about 25 to 35 mL of 0.1 mol/L ammonium acetate was added to the vial. The final mass fraction was determined. This gave solutions of niacinamide and labeled niacinamide with mass fractions of about 200 µg/g and solutions of niacin and labeled niacin with mass fractions of about 40 µg/g and 130 µg/g respectively. For both analytes, the solutions were combined to yield three calibration blends with isotope mass ratio close to 1.

A.3 Sample Preparation

The samples received were stored following the directions of their certificates and allowed to warm to room temperature before the analysis. During the process, samples were prepared according to the final conditions of the methodology developed by Melissa Phillips (single extraction with approximately 30 mL of 1 % acetic acid).

A portion of 2.5 g sample of each CRM was accurately weighed in triplicate in 50 mL polyethylene centrifuge tube. An aliquot of internal standard of niacinamide or niacin (exact mass known) and a portion (≈ 25 mL) of extraction solvent (1 % acetic acid in water) were added to each sample. The total extraction volume ranged from 25 mL to 30 mL. Samples were extracted by sonication for 30 min at room temperature, and then were centrifuged for 15 min at 3000 rpm, and an aliquot of the supernatant was removed and filtered through a 0.45 µm regenerated cellulose (RC) filter into an autosampler vial. Extracts were analyzed by ID-LC-MS/MS and the resulting vitamin mass fractions compared. All sample preparation was conducted under reduced lighting to minimize potential vitamin degradation.

A.4 Quality Control

Considering the wide range of mass fraction content of niacin (as niacinamide) in the CCQM-K147 samples, two materials (high and low mass fraction level) were used as daily method validation/ control materials. For both controls, independent units of the samples designated to be evaluated, were used.

A portion of DMR-486b, Leche Semidescremada en Polvo and SRM 1869 Infant/Adult Nutritional Formula II were used as control samples (low and high mass fractions of niacinamide) and were prepared and analyzed as described above for the rest of CRM samples. The resulting mass fractions of niacinamide were compared.

A.5 Instrumentation

Samples were analyzed by using an Agilent Series 1290 LC equipped with an Agilent Series 6410 Triple Quadrupole MS with electrospray ionization in the positive ion mode. The system was composed of a mobile phase degasser, binary pump, autosampler, and mass selective detector. The relevant instrument parameters included:

- Nebulizer pressure 15 psig
- Drying gas flow 11 L/min
- Drying gas temperature 350 °C
- Capillary voltage 4000 V
- Dwell time 100 s

A Cadenza CD-C18 column (250 × 4.6 mm i.d., 3 µm particles, Serial PF20HUD) from Silvertone Sciences (Philadelphia, PA) was used for the analyses without a guard cartridge. Mobile phase A consisted of 20 mmol/L ammonium formate in water adjusted to pH 4.0 with formic acid, and mobile phase B was methanol. A 10 µL injection volume was used for all samples. The following gradient elution program was used with a flow rate of 0.8 mL/min: (0 to 10) min, 100 % A; ramping to 50 % A at 45

min; (45 to 55) min, 50 %A; ramping to 100 % A at 57 min; (57 to 67) min, 100 % A. 10 µL injection volume was used.

Quantification was performed in multiple reactions monitoring (MRM) mode using the timetable, transitions, fragmentation voltages, and collision energies listed in Table A1 for the vitamins and their respective internal standards.

Table A1. Conditions for multiple reaction monitoring (MRM) for analysis of niacinamide by ID-LC-MS/MS Units identified as daily control materials

Vitamin	Precursor (m/z)	Product (m/z)	Fragmentation E° (V)	Collision E° (V)
B3-NH_2	123	80	20	20
$^2H_4$,B3-NH2	127	80	20	20

APPENDIX B: OpenBUGS Analysis Code

B.1 Vitamin B_3 Materials

```
# OPENBUGS code
# Scalars
# a..... intercept
# b..... slope
# n0..... number of materials (here, 9)
# n1..... number of units per material (here, 3)
# n2..... number of aliquots per unit materials (here, 3)
# n3..... number of repeats per aliquot (here, 2)
# pmthd.. instrumental 1/(relative variance)
# smthd.. instrumental relative SD
# pmodel. Measurand model precision
# smodel. Measurand model imprecision
#
# Vectors
# doe[n0].... degree of equivalance
# prept[n0].. instrumental 1/variance
# pVhat[n0].. 1/(uVda2 * uVda2)
# pVtru[n0].. 1/(uVda1 * uVda1)
# uVda1[n0].. reported uncertainties (certified and Informative)
# uVda2[n0].. same as uVda1 (replacing the Informative cases with its
corresponding value Vda1[])
# Vda1[n0]... reported values (certified and Informative)
# Vda2[n0]... identical to Vdat1
# uRda1[n0].. measurand model uncertainty estimates
# Rhat[n0]... conditional predicted R values
# srept[n0].. instrumental SDs
# Vhat[n0]... conditional V values
# Rhatun[n0]. Unconditional predicted R values
# Vtruun[n0]. Unconditional reference V values
#
# Matrices
# dlta[n0,n1,n2].... unit-related bias
# gmma[n0,n1]....... aliquot-related bias
# pdlta[n0,n1,n2]... unit-related 1/variance
# pgmma[n0,n1]...... aliquot-related 1/variance
# Rdat[0,n1,n2,n3].. individual R measurements
model {
#######################################################
# Regression parameters: you must de-comment one of the two "a"
# definitions
#######################################################
#a~dnorm(0, 1.0E-5)  #Remove the initial “#” for R=a+bV
#a<-0               #Remove the initial “#” for R=bV
b~dnorm(1, 1.0E-5)
#######################################################
# Instrumental variability-related parameter & distributions
# & model unconditional variability
#######################################################
pmthd~dgamma(1.0E-5, 1.0E-5)
smthd<- 100/sqrt(pmthd)
for(i in 1:n0) {
             prept[i]<- pmthd/pow(uVda1[i], 2)
             srept[i]<- 1/sqrt(prept[i])
}
for(i in 1:n0) {
             pmodel[i]<- 1/pow(uRda1[i], 2)
```

```
            smodel[i]~dnorm(0, pmodel[i])
}
#####################################################
# Certified value-related distributions
#####################################################
for(i in 1:n0){
            Vtru[i]~dnorm(0, 1.0E-5)
            pVtru[i]<-1/pow(uVda1[i], 2)
            Vda1[i]~dnorm(Vtru[i], pVtru[i])
}
for(i in 1:n0){
            Vhat[i]~dnorm(0, 1.0E-5)
            pVhat[i]<-1/pow(uVda2[i], 2)
            Vda2[i]~dnorm(Vhat[i], pVhat[i])
}
#####################################################
# Regression-related predictions
#####################################################
for(i in 1:n0){Rhat[i]<- a+b*Vhat[i]}
for(i in 1:n0){Rhatuc[i]<- Rhat[i]+smodel[i]}
for(i in 1:n0){Vhatuc[i]<- (Rhatuc[i]-a)/b}
#####################################################
# Measurement/ANOVA-related distributions
#####################################################
for(i in 1:n0){for(j in 1:n1){
            pgmma[i,j]~dgamma(1.0E-5, 1.0E-5)
            gmma[i,j]~dnorm(Rhat[i], prept[i])}}
for(i in 1:n0){for(j in 1:n1){for(k in 1:n2)        {
            pdlta[i,j,k]~dgamma(1.0E-3, 1.0E-3)
            dlta[i,j,k]~dnorm(gmma[i,j],pgmma[i,j])}}}
for(i in 1:n0){for(j in 1:n1){for(k in 1:n2){for(l in
1:n3){Rdat[i,j,k,l]~dnorm(dlta[i,j,k],pdlta[i,j,k])}}}}
#####################################################
# doe estimation
#####################################################
for(i in 1:n0){doe[i]<-200*(Vtru[i]-Vhatuc[i])/(Vtru[i]+Vhatuc[i])}
}

#############################################################
## first data set                                          ##
## CRMs & measurand characterization uncertainty           ##
## the uncertainty of any informative value is assigned    ##
## the same reported value instead of excluding them       ##
#############################################################
Vda1[]        uVda1[]       Vda2[]
              uVda2[]       uRda1[]       #PI           CRM
4.51          0.11          4.51          0.11          0.09          #CENAM
              DMR-486b
5.52          0.13          5.52          0.13          0.11          #CENAM
              DMR-274g
5.91          0.195         5.91          0.195         0.12          #NIST
              SRM-1549a
8.83          0.205         8.83          0.205         0.18          #CENAM
              DMR-82c
39.8          1.35          39.8          1.35          0.66          #NIM
              GBW(E)100227
60.6          0.65          60.6          0.65          1.10          #KRISS
              108-02-003
```

```
65.3          2.8           65.3          2.8           1.30          #NIM
              GBW10037
97            2.00          97            2.00          2.01          #NIST
              SRM-1869
108           5.00          108           5.00          2.03          #NIST
              SRM-1849a
END



###################################
## second dataset                ##
## Experimental Measurements     ##
###################################
list(n0=9, n1=3, n2=3, n3=2,
Rdat=structure(.Data=c(
4.449, 4.433, 4.083, 4.554, 4.646, 4.210,
4.684, 4.806, 4.491, 4.501, 4.486, 4.471,
4.637, 4.528, 4.692, 4.540, 4.501, 4.620,
5.433, 5.334, 5.417, 5.459, 5.445, 5.488,
5.598, 5.352, 5.893, 5.692, 5.282, 5.450,
5.575, 5.436, 5.609, 5.586, 5.189, 5.554,
6.031, 5.827, 6.024, 6.027, 5.539, 5.721,
5.592, 5.986, 6.221, 5.981, 5.096, 4.983,
5.369, 5.394, 5.973, 5.909, 5.556, 5.762,
8.724, 9.207, 8.916, 8.853, 8.952, 8.636,
9.189, 9.117, 9.179, 8.553, 9.043, 8.979,
8.609, 8.539, 8.549, 8.421, 8.621, 8.166,
36.49, 36.94, 37.98, 36.81, 37.50, 37.45,
35.01, 35.56, 37.79, 36.14, 35.29, 36.92,
39.25, 39.49, 37.58, 37.20, 39.14, 37.95,
60.53, 59.99, 61.79, 61.63, 58.69, 60.85,
61.81, 60.23, 60.11, 61.07, 61.43, 59.44,
59.16, 59.99, 59.68, 59.58, 60.83, 58.72,
65.32, 64.56, 65.48, 65.25, 63.64, 64.71,
61.48, 63.33, 63.02, 63.49, 65.14, 65.15,
64.66, 64.80, 64.42, 63.86, 63.82, 63.54,
 98.59, 96.57, 96.76, 96.83, 100.47, 100.75,
 98.84, 98.59, 96.48, 94.42,  95.75,  97.97,
100.27, 96.67, 97.02, 97.21, 100.69,  98.80,
100.10,  98.51,  99.05, 101.62, 100.24, 100.30,
100.33, 100.29, 101.97,  99.54, 101.29, 101.78,
 99.89,  99.20,  98.11,  98.74, 100.74,  99.48
), .Dim=c(9, 3, 3, 2) ))

#############
## Inits  ##
#############
list(pmthd=1,
pgmma=structure(.Data=c(
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1),
.Dim=c(9, 3)),
pdlta=structure(.Data=c(
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1),
.Dim=c(9, 3, 3)))
```

```
######################################################################
## after this we must accept gen inits for the uninitialized
## parameters
##
## then proceed to model/update and inference,
##
## finally consider burning out the initial few thousands of
## simulated values.
######################################################################
```