17th IMEKO TC 10 and EUROLAB Virtual Conference
"Global Trends in Testing, Diagnostics & Inspection for 2030"
October 20-22, 2020.

Assessment of the homogeneity of repackaged batches of reference materials

Abella, Johanna[1]; Mojica, Andrea[2]; Ahumada, Diego A[1]

[1]*Grupo de investigación en Metrología Química y Bionanálisis, Subdirección de Metrología Química y Biomedicina, Instituto Nacional de Metrología, Bogotá, Colombia, subdirquimica@inm.gov.co, (571) 2542222*
[2]*Grupo de investigación en Metrología Química y Bionanálisis, Subdirección de Innovación y Servicios Tecnológicos, Instituto Nacional de Metrología, Bogotá, Colombia, subdirinnovación@inm.gov.co,(571) 2542222*

***Abstract* – During 2020, the National Metrology Institute of Colombia, organized one proficiency testing (PT) for the quantification of elements in water. In the PT, external reference values were used instead of consensus values derived from the participants' results. The test material used in the PT was a certified reference material from another National Metrology Institute. To maintain the confidentiality of certified values, it was necessary to carry out a repackaging process. Therefore, in this work are presented the results of several studies that sought to establish the best mechanism to carry out this repackaging process. Among the investigations carried out, were evaluated: (i) kind of sampling and (ii) the number of bottles to assess the homogeneity of the new batch. The homogeneity uncertainty was estimated based on the number of bottles and depending on the type of sampling. From the results, it was identified that (i) most of the elements evaluated, the smallest contribution to the uncertainty is obtained when every two units are systematically sampled, and (ii) the estimation of uncertainty due to homogeneity causes an overestimation up to three times when the number of bottles is varied.**

Keywords – Assessment of homogeneity, Uncertainty, Proficiency testing

I. INTRODUCTION

Proficiency testing (PT) are a kind of interlaboratory comparison designed to evaluate the competence of participating laboratories for a specific test [1]. The participation in the proficiency testing constitute an essential tool for laboratories, because can be used to i) assess the performance of analytical methods, ii) identify measurement problems, iii) assess efficacy of analytical quality control, iv) assess measurement uncertainty, v) verify the reliability of their results against the reference values [2][3]. The provider of the PT distributes the test material, which is identified as a comparison item that can be a reference material (RM) or certified reference material (CRM) [4]. From a metrological point of view, the ideal case is when the PT provider distributes between participant laboratories a CRM [5].

CRM is a measurement reference, which has values of one or more specified properties with the associated uncertainties and traceability [6]. The advantages of using a certified reference material are: (i) there is greater independence in the evaluation, (ii) the value is traceable to the international system of units (SI), and (iii) there is no minimum number of participants required to perform the PT. Unfortunately, the use of CRMs is an expensive approach also, and the appropriate reference materials are not often available [3].

In some cases, when the certificates of reference materials are on the producer website, the organizer of the PT has to take the necessary actions to maintain the confidentiality of the material. Some of the actions that the PT provider can choose is to repack or relabel the reference material. If the material is repacked, it is necessary to take all actions to prevent contamination and preserve the homogeneity of it [7].

The homogeneity can refer to the variation of a property value between separate units of the material [8]. This homogeneity is expressed as the uncertainty due to the heterogeneity between bottles (u_{hom}).

This uncertainty is considered as one of the primary sources of uncertainty in the certification of reference material [9]. The magnitude of this uncertainty source can vary widely, depending on the nature of the certified property [8]. In the estimation of u_{hom}, the precision of the measurement method is one of the more critical aspects of homogeneity study [10]. In this context, the aim of this work is to present the results of different studies associated with the use of different sampling strategies in homogeneity experiments, in order to reduce the uncertainty associated with the homogeneity of the new

batches.

II. RELATED RESULTS IN THE LITERATURE

The homogeneity testing, within bottle homogeneity and between bottles homogeneity are two types of homogeneity that contribute to the uncertainty. Most important to carry out a homogeneity study within-bottle with a method that has the least analytical variation, so, the heterogeneity or between bottle effect can be quantified. [11]. Supplementary comparison on the determination of elements in river water was organized as part of the Matrix Reference Materials for Environmental Analysis project. The measurements of the homogeneity of the test material were developed with HR-ICPMS. Between bottle homogeneity for As, Cd, Ni, Pb and, Se were determined into 0.34 % and 0.79 % [12].

Otherwise, the repackaged of a CRM is an option when is necessary to keep the identity hidden of material according to its use. A control composites was developed for the National Food and Nutrient Analysis Program (NFNAP). This control was used to look out the precision and accuracy of laboratories and changes in the analytical methods, in other activities. Control composites was repackaged from certified reference materials. These reference materials were purchased from commercial suppliers. The homogeneity was assessed by analysis of the most representative nutrients. The results for the verification of the homogeneity confirmed the uniformity of the control. The variation between units was less than 1% [7].

Another use of repackaging of reference materials is the certification of the properties values or recertification through an interlaboratory comparison. A total, elemental and isotopic carbon interlaboratory was developed. A batch of SRM 1649 was repackaged and distributed to the laboratories. In this material, total carbon has a contribution of less than 1% error due to sample homogeneity [13].

III. DESCRIPTION OF THE METHOD

A. Repackaged batch of reference material

Certified drinking water reference material was produced by the National Metrology Institute of Colombia. This material was used in this study. One bottle of 125 g of CRM was divided into four portions, which were added directly to clean LDPE containers. On three repackaged bottles, 30 g of the CRM was added. In the fourth bottle, 15 g of the CRM were added, 30 g CRM repackaged were completed by adding 15 g from another bottle of the initial CRM. Fig. 1 shows the repacked scheme for the bottles.

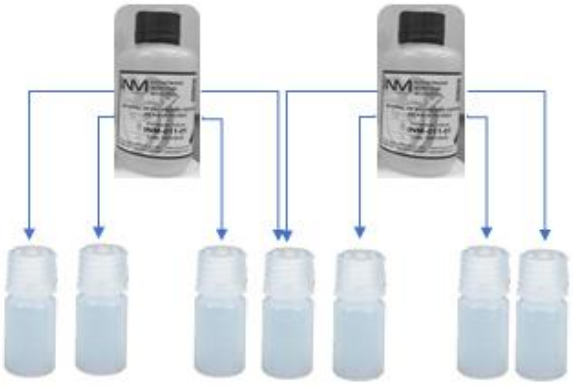

Fig. 1. Repacked scheme for the new batch

In total, six bottles of the CRM were repackaged, generating a new batch of 24 refilled units. The repackaged of each unit was carried out gravimetrically using a XPE 204 Mettler Toledo digital balance.

Prior to the experiment, all bottles and labware were cleaned by soaking in a 4 mol L^{-1} nitric acid solution for one week, followed by rinsing five times with pure water in a cleanroom.

B. Measurement of elements

The analyses for the homogeneity test were performing by using inductively coupled plasma with mass spectrometry detection (ICP-MS) Perkin Elmer NexION 300D instrument. The plasma was generated using Argon (99.990%), a power of 1600 W, gas flow rates of plasma 15 L min^{-1}, auxiliary 1.2 L min^{-1}, and nebulizer 0.52 L min^{-1}. A quartz cyclonic chamber, Meinhard nebulizer, nickel sampler, and skimmer cones were used for measurements. The torch position, gas flow rates, plasma power, and deflection voltage were optimized daily with a tuning solution containing Be, In, U, Ce to 1 ug L^{-1}.

ICP-MS was used to measure ^{97}Mo, ^{98}Mo, ^{54}Fe, ^{25}Mg, ^{43}Ca, ^{75}As, ^{208}Pb, ^{113}Cd, ^{66}Zn, ^{63}Cu and ^{50}Ni. A solution of the internal standard was mixed online with the samples before reaching the plasma.

C. Reagents and reference materials

The samples were diluted using HNO_3 from Sigma-Aldrich (St. Louis, USA) and H_2O_2 (SupraPure) from Merck Chemicals (Darmstadt, Germany). The HNO_3 used was purified by double sub-boiling distillation.

The reference materials were purchased from the National Institute of Standards and Technology (NIST) and Slovak Institute of metrology (SMU).

D. Assessment of homogeneity

A homogeneity test was carried out following the ISO Guide 35:2017. All the bottles of the batch were subjected to measurements by ICP-MS. The measurements were randomized. Twelve replicates of measurement were made for each bottle.

An analysis of the variance (ANOVA) was carried out for the results, and the obtained within-bottle mean square (MS_{within}) and between-bottle mean square ($MS_{between}$) were applied in the equation (1) [8].

$$s_{bb}^2 = u_{bb}^2 = \frac{MS_{between} - MS_{within}}{n_0} \qquad (1)$$

where, s_{bb} represent the between unit variance or between unit homogeneity uncertainty (u_{bb}), $MS_{between}$ and MS_{within} are between-group mean square and withing group mean square respectively, n_o is the number of replicates.

D. Sampling strategies

Considering that all the bottles in the batch were measured, the effect of sampling techniques over uncertainty estimation was simulated. In this study, simple random sampling and systematic random sampling were evaluated.

For simple random sampling, the simulation consisted of using a random number generator to choose the number of the bottles of the batch to evaluate u_{hom}. Afterward, since to each bottle has a certain number of measurement data, these were used to evaluate the homogeneity of the batch, according to equation 1. On the other hand, for the with a random number generator, the effect of the number of bottles used for the assessment of homogeneity in the batch was studied.

Additionally, systematic sampling of units from the batch was evaluated, thus: i) every four units a bottle was selected, completing a total of six units of the batch; ii) every two units a bottle was chosen randomly, completing a total of 12 units; iii) every three units a bottle was randomly selected, completing a total of seven units.

IV. RESULTS AND DISCUSSIONS

A. *Measurement method*

According to the chemical properties of the solutions, the drinking water is considered a homogeneous solution. Hence, the variation within the bottle is due to the variation generated by the measurement method, especially considering that all bottles were analyzed in repeatability conditions.

On the other hand, the measurement method should have adequate repeatability for being applied in the homogeneity evaluation. Thus, the uncertainty u_{rep} in the analytical result due to random effects in the analytical process (e.g., instrumental drift, weighing of internal standard) can be estimated from the repeatability, expressed as the standard uncertainty (see Fig. 2) from ANOVA results.

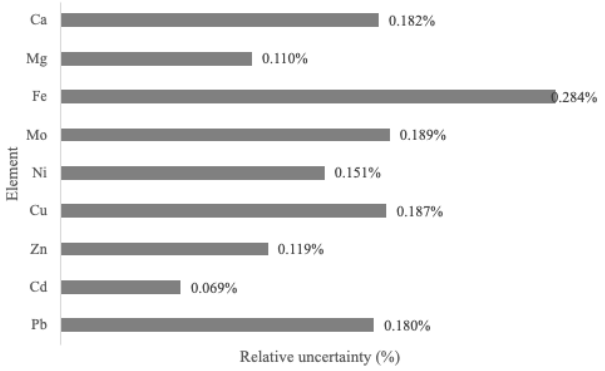

Fig. 2. Method uncertainties estimated for different elements.

According to equation (1), if $MS_{between} < MS_{within}$, it is an indication that the study setup and/or the method repeatability was not sufficiently good. In this context, considering that (i) the target uncertainty is between 1% and 2% [14], (ii) the results of the Fig. 2, and that (iii) twelve replicates of measurement were made per bottle, it is found that the method is fit to the purpose.

B. *Simple random sampling: number of bottles study.*

For the purpose of this research, the estimation of the inhomogeneity component of uncertainty was undertaken using the Guide ISO 35. In this guide, an acceptable estimate of the between bottle variance can be obtained with a minimum of ten units. A good practice is to increase the number of units examined as the total number of units produced. For a quantitative property, the recommended number of units for homogeneity testing is according to equation 2.

$$N_{min} = \max\left(10, \sqrt[3]{N_{prod}}\right) \qquad (2)$$

Where N_{prod} is total units produced in batch.

If the batch size is below 100 units, homogeneity can be assessed with three units minimum or 10% of the total units produced, selected at random.

In this study, 24 units were produced, then according to the recommendation of ISO guide 35, three units minimum should be assessed. Fig. 3. shows uncertainties associated with inhomogeneity for Calcium, Iron, and Magnesium (elements that are on the order of mg kg^{-1}) using a different number of units from the batch. These results are obtained by performing a simple random sampling.

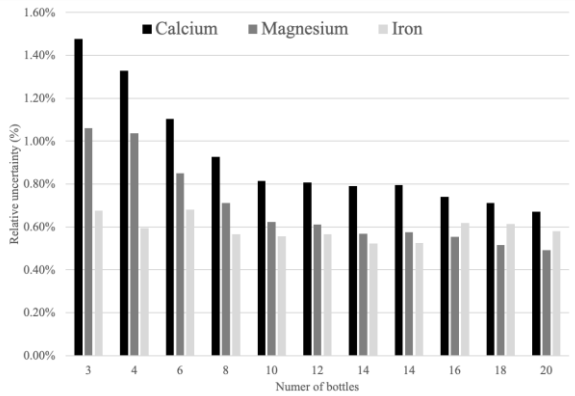

Fig. 3. Variation of the uncertainty related to inhomogeneity (Ca, Mg, and Fe) with the number of bottles in the homogeneity study.

Fig. 4 and Fig. 5 shows the uncertainties associated with inhomogeneity for Cadmium, Nickel, Molybdenum, Copper, Lead, and Zinc (elements that are on the order of μg kg^{-1}) using a different number of units from the batch.

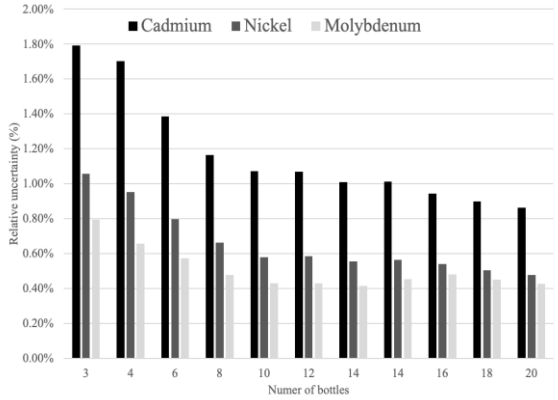

Fig. 4. Variation of the uncertainty related to inhomogeneity (Cd, Ni, and Mo) with the number of bottles in the homogeneity study.

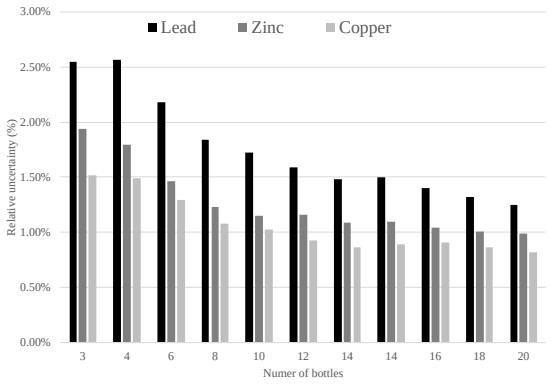

Fig.5. Variation of the uncertainty related to inhomogeneity (Pb, Cu, and Zn) with the number of bottles in the homogeneity study.

Fig. 3 and Fig. 5 shows that uncertainty for homogeneity decreases as the number of bottles used in the study increases, except for Iron (Fe) that had similar uncertainties in all the cases. In general, this decrease is close 50% approximately for the majority of elements evaluated, which reveals the importance of the number of bottles used for the study.

Similarly, as can be observed in the Fig.3 to 5, there is an asymptotic behavior of the uncertainties. Even when evaluating the uncertainty with the entire batch (24 units), the uncertainty value is practically the same as using 20 units. Therefore, it can be concluded that the use of only three units is not convenient, if the uncertainties reported previously reported by other authors for similar reference materials [12,14]. In the same way, as can be seen in Fig. 3 to 5, it is found that taking an excessive number of bottles is not very useful and, on the contrary, can lead to a lot of work.

C. Systematic random sampling: number of bottles and study.

The main advantage of using systematic sampling over simple sampling is its simplicity, because it does not require a random number generator. This strategy of sampling involves selecting items from an ordered population using a frequency or sampling interval. In this study, three sampling intervals were evaluated. Fig. 6. show the variation of the homogeneity uncertainty with sampling interval for some elements.

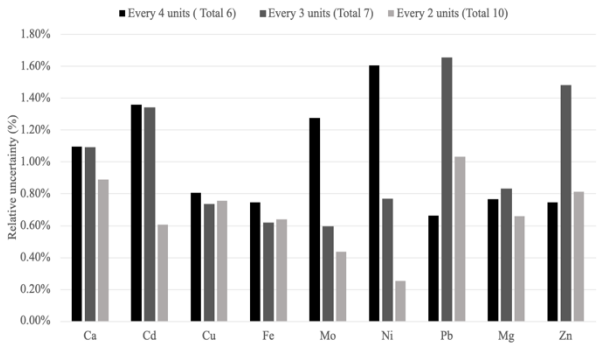

Fig. 6. Variation of homogeneity uncertainty with the sampling interval.

For Cadmium, Molybdenum, and Nickel lowest homogeneity uncertainty were obtained when one bottle was sampled every two units, as indicated in Fig. 6.; likewise, for these elements, it was found the same behavior as that found in the random sampling (see Fig.3-5), where the results showed an inverse-asymptotic relationship between the uncertainty and the total number of bottles.

Furthermore, Iron, Magnesium, and Copper were not influenced by the variation in the sampling interval, since,

there no changes are observed in the uncertainty homogeneity, as can be observed in Fig. 6. Lastly, lead and zinc exhibit unpredictable behavior.

Table 1, present the uncertainties obtained for simple random sampling using 10 bottles, systematic random sampling using ten bottles, and the homogeneity assessments using all the 24 bottles. As can be seen from Table 1, some uncertainties are reduced with the change in the sampling strategy. To give a point of comparison, the uncertainty for each sampling strategy was estimated using ten bottles.

Considering the equation (1), it is possible to assume that the uncertainties estimates for the entire batch (24 bottles) is the reference or even better, it represents the asymptotic value of the Fig. 3 to 5. Accordingly, Table 1 evidences that the systematic strategy is more appropriate, since its uncertainties are closer to those of the reference (complete batch). It is important to note that for simple sampling, between 14 and 16 units are required to obtain similar uncertainties.

Assuming, that the uncertainties are independent random variables, the correlation coefficients obtained between each sampling strategy and the reference uncertainty are presented in Table 1. These correlation coefficients confirm that better sampling strategy is the systematic approach.

Table 1. Summary of uncertainties related to batch inhomogeneities using different sampling strategies.

	u_{hom} simple sampling*	u_{hom}, systematic sampling *	u_{hom}, complete batch
Ca	0.81%	0.89%	0.58%
Cd	1.07%	0.61%	0.17%
Cu	1.02%	0.75%	0.76%
Fe	0.56%	0.64%	0.61%
Mo	0.43%	0.43%	0.41%
Ni	0.58%	0.25%	0.31%
Pb	1.72%	1.03%	1.15%
Mg	0.62%	0.66%	0.68%
Zn	1.15%	0.81%	1.03%
Correlation coefficient	0.61	0.74	

* uncertainties estimated using 10 bottles.

Finally, to uncertainties associated with random sampling errors, systematic sampling errors may also be present. Simple random sampling tends to be inefficient when assessing distributions with the possibility of trends because most of the time is spent evaluating no-representative samples, leading to high error variances.

V. CONCLUSIONS AND OUTLOOK

This paper demonstrates the effectiveness of systematic random sampling by means of a simple experiment that consists of simulating the choice of the bottles in the assessment of the homogeneity of reference materials. The above, considering it was found that the number the bottles was the minimum using systematic sampling (the u_{hom} varied between from 0.23% to 1.03%); while for simple random sampling the uncertainties varied between 0.43% and 1.03%, for the same number of bottles.

Through this simulation, it was found that it is necessary to take about 65% of the bottles in the batch to make an adequate estimate of the uncertainty using simple random sampling, which differs considerably from the recommendation of ISO guide 35 related with small batches.

VI. ACKNOWLEDGMENTS

This research has been conducted with support from the Instituto Nacional de Metrología of Colombia. The authors acknowledge to Paula Sánchez for her helpful review of the manuscript.

REFERENCES

[1] **Kuselman, I.; Pavlichenko, M.**: Designs of the experiment for proficiency testing with a limited number of participants, *Accreditation and Quality Assurance*, Vol. 9., 2004, pp. 387-390.

[2] **Ellison, S.L.R.**; **Hardcastle, W.A.**: Causes of error in analytical chemistry: results of a web-based survey of proficiency testing participants, *Accreditation and Quality Assurance*, Vol. 17., No. 4., 2012.

[3] **Eurachem:** Selection, Use and Interpretation of Proficiency Testing (PT) Schemes. Second Edition. 2011.

[4] **Krska, R.; Welzig, E.**: Certified reference materials and proficiency testing, *Anal Bioanal Chem*, Vol. 376., 2003, pp. 316-318.

[5] **Belli, M.; Ellison, S.L.R.; Fajgelj, A.; Kuselman, I.; Sansone, U.; Wegscheider, W.:** Implementation of proficiency testing schemes for a limited number of participants, *Accreditation and Quality Assurance*, Vol. 12., 2007, pp. 391-398.

[6] **BIPM.:** JCGM 200:2012. International Vocabulary of Metrology – Basic and general concepts and associated terms (VIM). 3rd edition. 2012.

[7] **Phillips, K.M.; Patterson, K. Y.; Rasor, A.S.; Exler, J.; Haytowitz, D.B.; Holden, J. M.; Pehrsson, P. R.:** Quality control materials in the USDA National Food and Nutrient Analysis Program (NFNAP), *Anal Bioanal Chem*, Vol. 384., 2006, pp. 1341-1355.

[8] **International organization for Standardisation (ISO)**. ISO GUIDE 35:2017 Reference materials — Guidance for characterization and assessment of homogeneity and stability. Geneva, 2017.

[9] **Olivares, I.R.B.; Souza, G.B.; Nogueira, A.R.A; Toledo G.T.K.; Marcki, D.C.:** Trends in developments of certified reference materials for chemical analysis – Focus on food, water, soil, and sediment matrices, *Trends in Analytical Chemistry*, Vol. 100., 2018, pp 53-64.

[10] **Zeisler, R.; Murphy, K.E.; Becker, D. A.; Davis, W. C.; Kelly, W. R; Long, S. E.; Sieber, J. R.:** Standar Reference

Materials ® (SRMs) for measurement of inorganic environmental contaminants, Vol. 386., 2006, pp. 1137-1151.

[11] **Lisinger, T.P.J.; Pauwels, J.; van der Veen, A. M.H.; Schimmel, H.; Lamberty, A.:** Homogeneity and stability of rerence materials, *Accreditation and Quality Assurance*, Vol. 6., 2001, pp. 20-25.

[12] **EURAMET.** EURAMET.QM-S11/ EURAMET pilot. Supllementary comparison & pilot study on determination of elements in river water. Technical protocol. 2017.

[13] **Currie, L.A.; Benner Jr, B.A.; Kessler, J.D.; et. al.:** A critical evaluation of interlaboratory data on total, elemental and isotopic carbon in the carbonaceous particle reference material, NIST SRM 1649a, *Journal of Research of the National Institute of Standards and Technology*, Vol. 107., No. 3., pp. 279-198.

[14] **Ahumada, D.A.; Soto, L.L.; Morales, L.V.; Abella, J.P.:** Desarrollo de un material de referencia certificado para análisis elemental de agua potable. *Revista Colombiana de Química*,Vol. 48., No. 3., 2019, pp. 36-44.