Alternative designs for the assessment of homogeneity: use of characterisation data

Thomas P. J. Linsinger¹ · Thomas Bacquart^{1,2}

Received: 5 October 2018 / Accepted: 4 April 2019 / Published online: 24 April 2019
© The Author(s) 2019

Abstract

In order to reduce analysis cost and save samples, the homogeneity study of two titanium-certified reference materials was combined with the interlaboratory characterisation of that material. Samples were selected using a random-stratified sampling scheme and were sent to nine and 11 laboratories, respectively. The data were normalised to screen for trends in the production sequence, outliers and normality of the distribution of sample means. The non-normalised data were screened for outlying variances, and the uncertainty of homogeneity was evaluated using two-way analysis of variance. The uncertainty contributions determined were in all cases significant, but did usually not dominate the uncertainty of the assigned value. The findings were confirmed by a homogeneity study with a limited scope on one of the two materials. While the assessment of homogeneity using the approach was successful, the goal of reducing the consumption of samples and reducing analytical cost was only partially realised: a similar number of samples were consumed and 5 % – 10 % lower analytical cost than in a classical, separate homogeneity and characterisation study set-up was incurred.

Keywords Certified reference material · Homogeneity assessment · ISO Guide 35

Introduction

Homogeneity is one of the defining characteristics of reference materials (RMs). Consequently, ISO 17034 [1] requires the assessment of homogeneity. While this does not necessarily require actual measurements (exceptions could be, for example, for RM production by validated processes that are under good statistical control), in most cases, and certainly for every new material, measurements are performed that allow confirmation of “sufficient homogeneity”. A dedicated homogeneity study requires using several samples from the production batch (reduction in the sellable samples), time and investment to perform all the necessary analysis. Therefore, optimisation of the set-up of the homogeneity study may be seen as a potential way to lower the final RM price. ISO Guide 35 [2] gives several examples on how the requirements of ISO 17034 can be fulfilled. They all rely on an

appropriate selection of a sufficient number of samples of the whole batch and subsequent analysis in one or several runs by one laboratory. In the first step, the data set should be checked for analytical problems like analytical drift leading to trends in the measurement sequence and outliers. After investigation and resolution of all analytical problems, the analytically valid data set should be checked for trends in the processing sequence before the quantification of the between-unit variation is performed. ISO Guide 35 recommends at least nine degrees of freedom for the final standard deviation between units (s_{bb}). Ideally, all measurements are performed under repeatability conditions, as this ensures the lowest possible analytical variation and therefore the best possibility to quantify the between-unit variation. ISO Guide 35 also foresees the possibility that not all measurements can be performed in one analytical run: in this case, the use of randomised block designs or balanced nested designs is recommended.

One of the possible characterisation approaches for certified reference materials is to use “two or more methods of demonstrable accuracy in one or more competent laboratories” (clause 7.12.3 in ISO 17034). This is often implemented as an interlaboratory comparison among laboratories of demonstrated competence. As each of the laboratories

✉ Thomas P. J. Linsinger
thomas.linsinger@ec.europa.eu

¹ European Commission, Joint Research Centre (JRC), Geel, Belgium

² Present Address: National Physical Laboratory, Teddington, UK

measures different samples, it should be possible to use the data from such a characterisation exercise also for the assessment of homogeneity. Such an approach could have two advantages:

- Fewer samples are consumed, as combining homogeneity and characterisation study eliminates the need for reserving a separate set of samples. This can be relevant for small production batches.
- This combination can lead to cost savings. These, however, will usually be relatively minor, as the relatively large measurement series result in relatively low analytical costs per unit. At the JRC Geel, the costs for the homogeneity studies are usually below 10 % of the total project cost.

Combining the homogeneity with the characterisation study will not necessarily lead to shorter time to the market, as a separate homogeneity study can also be run parallel to the characterisation study.

In this manuscript, we describe the approach taken to use the characterisation data for ERM-EB090a and ERM-EB090b, two titanium RMs to be certified for the trace metal mass fractions for homogeneity assessment.

Material

Titanium ingots were produced by conventional alloying of titanium with 25 elements (Al, B, Ce, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, N, Nb, Ni, O, Pd, Ru, Si, Sn, Ta, V, W, Y and Zr) at mass fractions of 0.1 g/kg – 3 g/kg. The ingots were subjected to vacuum arc remelting to improve homogeneity and forged into smaller diameter round bars. These forged bars themselves were machined into rods of a final diameter of 63 mm and used as electrodes that were then melted in the plasma rotating electrode powder process to obtain a powder.

The powder was sieved to remove particles larger than 0.8 mm, and the final mass of powder produced was 114 kg. The powder was mixed and filled into cans made of low-carbon steel with inner diameters of 63 mm. The cans were then subjected to hot isostatic pressing for 240 min at 1750 °C and 101 MPa. This resulted in eight dense bars in a mantle of low-carbon steel. The diameter of the titanium core was 54 mm, the outer diameter 63 mm. The length of the bars varied between 1147 mm and 1352 mm.

The top and bottom 20 mm of each bar were discarded. Of each bar, 42–50 discs of 20 mm height were obtained by water jet cutting. Of each of the discs, an inner core with a diameter of 40 mm was cut by water jet cutting. Each of the discs was engraved with the CRM code and a running

number representing the position in the bar. In total, only 400 discs were produced. This rather low number made the combination of characterisation and homogeneity study interesting.

The remaining parts of the bars were cut into slices of about 2 mm thickness. Chips of 6 mm diameter, each weighing about 250 mg, were punched out of these slices. All chips were mixed, and 7 g of chips each was filled into amber glass bottles. In total, 765 bottles produced.

Study set-up

Laboratory selection

Eleven laboratories that had demonstrated their competence in the analysis of metals in titanium or titanium alloys were invited to participate in the study. Ten of the laboratories hold formal accreditation to ISO/IEC 17025, but only three of them are actually accredited for the analysis of elements in titanium. Adherence to the provisions of ISO/IEC 17025 for the eleventh laboratory was checked using a questionnaire in which the laboratory confirmed that it runs a quality system and also confirmed that procedures for the technical requirements of ISO/IEC 17025 are in place.

Sample selection and measurement set-up

Twenty-four discs were selected using a random-stratified sampling scheme covering the complete batch. The samples were sent to six different laboratories that provided in total nine independent data sets, each data set consisting of two measurement results per element for each of three samples (six results per data set). As not every data set contained all elements, laboratories were asked to return the samples after analysis and six samples were sent to one of the laboratories for analysis for the omitted elements to achieve at least nine degrees of freedom for the between-unit variation (s_{bb}). This laboratory performed duplicate analysis for Hf, La, Nb and Ta under repeatability conditions. The set-up is shown schematically in Table 1.

Thirty-nine bottles of chips were selected using a random-stratified sampling scheme covering the complete batch. The samples were distributed to 11 laboratories that provided in total 14 independent data sets. The laboratories were the same as for the discs plus two additional laboratories. Again, each data set consisted of two results performed on chips of three bottles each (six results per data set) and all laboratories were requested to return their samples after analysis. To resolve inconsistencies in laboratory data, one laboratory performed quadruplicate analysis for all elements of six samples that were returned by the laboratories under

Table 1 Overview of samples and measurements performed

Laboratory	Samples received	Measurements performed
1	Three discs	Al, B, Co, Cr, Cu, Fe, Mn, Mo, Ni, Sb, Si, Sn, V, W, Zr by ICP-MS
	Three discs	Al, B, Co, Cr, Cu, Fe, Mn, Mo, Ni, Sb, Si, Sn, V, W, Zr by ICP-OES
	Three bottles of chips	Al, B, Co, Cr, Cu, Fe, Mn, Mo, Ni, Sb, Si, Sn, V, W, Zr by ICP-MS
	Three bottles of chips	Al, B, Co, Cr, Cu, Fe, Mn, Mo, Ni, Sb, Si, Sn, V, W, Zr by ICP-OES
2	Three discs	Al, B, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Y, Zr by ICP-MS
	Six discs from re-testing	Hf, La, Nb, Ta by ICP-MS; tests on 1 day
	Three bottles of chips	Al, B, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Y, Zr by ICP-MS
	Three bottles from re-testing	Al, B, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Y, Zr by ICP-MS; tests on 1 day
3	Three discs	Al, B, Cr, Cu, Fe, Mn, Mo, Nb, Ni, Pd, Ru, Si, Sn, V, Y, Zr by ICP-MS
	Three bottles of chips	Al, Cr, Cu, Fe, Mn, Mo, Nb, Ni, Si, Sn, V, Zr by ICP-MS
4	Three discs	Al, B, Co, Cr, Cu, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Zr by GD-MS
	Three discs	Fe by AAS
	Three bottles of chips	Al, B, Co, Cr, Cu, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Zr by GD-MS
	Three bottles of chips	Fe by AAS
5	Three discs	Al, B, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Y, Zr by ICP-OES
	Three bottles of chips	Al, B, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Nb, Ni, Pd, Ru, Sb, Si, Sn, Ta, V, W, Y, Zr by ICP-OES
6	Three discs	Co, Cr, Fe, Hf, Mo, Ru, Sb, Sn, Ta, W, Zr by NAA
	Three bottles of chips	Co, Cr, Fe, Hf, Mo, Ru, Sb, Sn, Ta, W, Zr by NAA
7	Three bottles of chips	Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Pd, Ru, Sb, Sn, V, W, Zr by NAA
8	Three bottles of chips	Al, Co, Cr, Cu, Fe, Mn, Mo, Nb, Ni, Sn, V, Y by ICP-MS
9	Three bottles of chips	Al, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Pd, Ru, Sb, Sn, Ta, V, W, Zr by NAA
10	Three bottles of chips	Al, B, Bi, Ce, Co, Cr, Cu, Fe, Hf, La, Mn, Mo, Nb, Ni, Sn, Ta, V, W, Y, Zr by ICP-OES
		Ni by photometry
		Co, Cu, Mn by AAS
	Three bottles of chips	Hf, La, Mo, Nb, Sb, Sn, Ta, V, W, Y, Zr by ICP-MS
		Fe, Mn by photometry
11	Three bottles of chips	Co, Cr, Fe, Hf, La, Mn, Mo, Ru, Sb, Sn, Ta, V, W, Zr by NAA

ICP-MS Inductively coupled plasma mass spectrometry, *ICP-OES* inductively coupled plasma optical emission spectrometry, *AAS* atomic absorption spectrometry and *NAA* neutron activation analysis

repeatability conditions. The set-up is shown schematically in Table 1.

The six measurements per data set were spread over at least 2 days to ensure conditions of intermediate precision. Laboratories were free to use their method of choice. Inductively coupled plasma mass spectrometry (ICP-MS) and optical emission spectrometry (ICP-OES), k_0 -neutron activation analysis (k_0 NAA) and glow discharge mass spectrometry (GD-MS) were used by the participants.

Statistical analysis

Independent data sets from different laboratories or different methods are usually affected by an unknown bias due to calibration and other systematic variations in the procedure. In addition, data from different laboratories or methods often show different precisions, which further complicates the analysis by analysis of variance (ANOVA). Different

statistical approaches were applied for the investigation of analytical trends, outliers and position in the bars on the one hand and for the quantification of s_{bb} . The data analysis was performed independently for discs and chips, respectively.

For the investigation of trends and outliers, each individual datapoint was normalised to the average of the respective data set. The normalised data were tested for trends in the production sequence by testing the slope of the linear regression result versus sample number for significance on a 95 % confidence level. The data set was tested for individual outliers and outlying sample means using the Grubbs test on a 99 % confidence level. All data sets were tested for normality using normal probability plots.

Datasets without trends or outliers were used in the second step of the evaluation. The non-normalised data for each element were tested for outlying variances using the Cochran procedure on a 99 % confidence level. This test was applied on the standard deviation of each data. Data sets with outlying variances were removed as ANOVA requires

equal variances in each cell. This standard deviation also includes a contribution of the between-unit variation, but this is expected to be the same for all laboratories. In addition, this contribution was expected to be small compared to the repeatability standard deviation, so there was little risk that this test eliminated the between-unit variation that should be assessed.

Quantification of between-unit heterogeneity was undertaken by two-way ANOVA without interaction of the original, non-normalised data using Statistica 13 (Dell Software, Round Rock, USA). This separates s_{bb} from the within-unit variation (s_{wb}). The latter is equivalent to the method intermediate precision. u_{bb}^* , heterogeneity hidden by method precision, was calculated as shown in the equation below [3].

$$u_{bb}^* = \sqrt{\frac{MS_{\text{within}}}{n}} \sqrt{\frac{2}{\nu_{\text{within}}}}$$

In this equation, MS_{within} is the mean of squares within-unit from an ANOVA, n the number of replicates per unit (2) and ν_{within} the degrees of freedom of MS_{within} . The larger value of s_{bb} and u_{bb}^* was used as estimate of the between-unit variation.

Results and discussion

Homogeneity assessment of Ti-discs

None of the elements show a trend in the position in the bars. Individual outliers were detected for Hf and La. These individual outliers also resulted in the respective average being an outlier for Hf, but not for La. Normal probability plots of the individual results were in line with the assumption of normally distributed data for all elements except Hf, La and Sb. While the deviation for La was small, the one for Hf was substantial enough to preclude evaluation by ANOVA. While the data for Sb did not follow a normal distribution, they did at least follow a unimodal distribution, thus making evaluation by ANOVA possible. The normal probability plots of the unit means were for all elements except Hf in agreement with a normal distribution. Based on this evaluation, the data of Hf were excluded from the further evaluations.

The Cochran test flagged several data sets as outliers. One data set each for B, Mo, Si and Ta and two data sets for Sb were excluded from the further evaluation. These data sets came from various laboratories and reflect the suitability of method principles or laboratory internal procedures for the determination of certain elements.

The results of the two-way ANOVA are shown in Table 2. s_{bb} of Ce, Pd, Sb and Y still have fewer than nine degrees of

freedom. Disagreement between results precluded certification of Ce and Sb anyway, and insufficient data sets were available for a certified value for Pd and Y. As certification was impossible for these elements, it was decided not to include them in the re-testing.

Repeatability standard deviations were, with the exception of Cu and Si, between two and four per cent, showing that the methods are sufficiently precise to allow a meaningful assessment of homogeneity. Despite the good precision, MS_{between} was in most cases smaller than MS_{within} and s_{bb} could therefore not be calculated. As expected from the careful manufacturing process, homogeneity between discs was very good with the conservative estimates for between-disc variation for most elements between one and two per cent. Standard deviations between-laboratory means ranged from 2.2 % to 9.0 % (data not shown). The estimates for the between-unit homogeneity are therefore significant, but, as shown in Table 2 column Fract, not a dominating factor for the uncertainty of most assigned values. This shows that the approach chosen delivered suitable homogeneity assessments for the production of certified reference materials.

Homogeneity assessment of Ti-chips

None of the elements show a trend in the production sequence. Individual outliers were detected for Cr and Si. These did not lead to outlying unit means, but one outlying unit mean for Co was found, and Co was therefore excluded from further evaluation. For all most data sets, the assumption of normally distributed individual data and bottle means was valid and all data sets followed at least unimodal distributions.

The Cochran test flagged several data sets as outliers. One data set each was excluded for Co, Cr, Cu, Mn, Ni, Pd, Sb, Si, Sn, Ta and V, and two data sets each were removed for Mo, Ru and W. As for the discs, these outlying data sets came from several different laboratories and reflect inherent variation of precision between laboratories.

The results of the two-way ANOVA are shown in Table 3. Repeatability standard deviations are comparable to the ones obtained for the discs. This is only expected as the laboratories providing the data are largely identical. Contrary to the discs, s_{bb} could be quantified for the majority of elements (16 of 22). The estimates for the between-bottle variation are similar to the ones obtained for the discs, with s_{bb} or u_{bb}^* ranging from 0.9 % to 2.9 %.

Comparison of multi-laboratory study versus single-laboratory study

The data from the re-testing of the characterisation samples allow the comparison of a multi-laboratory study evaluated by two-way ANOVA, and a dedicated homogeneity study

Table 2 Results of the homogeneity assessment of titanium discs

Element	$s_{wb,rel}$ (%)	$\nu_{between}$	$s_{bb,rel}$ (%)	$u_{bb,rel}^*$ (%)	$u_{bb,rel}$ (%)	Fract (%)
Al	3.90	10	n.c.	1.67	1.67	28
B	3.33	10	n.c.	1.42	1.42	10
Co	2.70	12	n.c.	1.10	1.10	36
Cr	2.37	14	n.c.	0.93	0.93	64
Cu	5.60	12	n.c.	2.29	2.29	76
Fe	1.82	14	0.83	0.72	0.83	52
La	2.77	12	n.c.	1.19	1.19	6
Mn	2.81	12	n.c.	1.15	1.15	34
Mo	1.98	12	1.27	0.81	1.27	31
Nb	2.56	14	2.69	1.05	2.69	86
Ni	3.37	12	n.c.	1.37	1.37	32
Pd	2.77	8	3.65	1.25	3.65	27
Ru	2.45	10	n.c.	1.05	1.05	22
Si	8.11	10	n.c.	3.47	3.47	7
Sn	3.93	14	n.c.	1.54	1.54	43
Ta	2.19	12	1.08	0.94	1.08	25
V	2.05	12	n.c.	0.84	0.84	29
W	2.21	12	1.35	0.90	1.35	44
Y	3.21	6	n.c.	1.56	1.56	20
Zr	2.50	14	n.c.	0.98	0.98	44

n.c., Cannot be calculated as the mean squared between units from ANOVA $< MS_{within}$; $\nu_{between}$, degrees of freedom of the mean squared between units from ANOVA; Fract, fraction of the homogeneity contribution to the combined uncertainty of the assigned value in %

Table 3 Results of the homogeneity assessment of titanium chips

Element	$s_{wb,rel}$ (%)	$\nu_{between}$	$s_{bb,rel}$ (%)	$u_{bb,rel}^*$ (%)	$u_{bb,rel}$ (%)	Fract (%)
Al	4.11	19	n.c.	1.44	1.44	22
B	2.93	13	1.53	1.11	1.53	12
Cr	1.59	23	n.c.	0.53	0.53	37
Cu	3.07	21	n.c.	1.05	1.05	41
Fe	3.01	27	n.c.	0.98	0.98	61
Hf	2.57	19	1.26	0.90	1.26	22
La	3.05	17	2.86	1.10	2.86	27
Mn	1.80	23	0.49	0.61	0.61	13
Mo	1.58	23	1.57	0.53	1.57	41
Nb	1.67	15	2.54	0.61	2.54	85
Ni	1.61	15	0.78	0.57	0.78	13
Pd	2.44	9	1.54	1.00	1.54	6
Ru	1.53	9	0.70	0.62	0.70	11
Sb	2.67	19	0.65	0.94	0.94	4
Si	6.44	11	n.c.	2.53	2.53	33
Sn	3.73	25	n.c.	1.23	1.23	20
Ta	2.24	17	0.99	0.80	0.99	36
V	1.48	21	0.93	0.51	0.93	33
W	1.59	19	1.05	0.56	1.05	14
Y	2.17	11	1.26	0.85	1.26	41
Zr	2.74	23	0.71	0.92	0.92	22

n.c., Cannot be calculated as $MS_{between} < MS_{within}$; $\nu_{between}$, degrees of freedom of the mean squared between units from ANOVA; Fract, fraction of the homogeneity contribution to the combined uncertainty of the assigned value in %

evaluated in a single laboratory: the 24 results (four results from six bottles each) were evaluated by one-way ANOVA. This data set results in only five degrees of freedom for the between-bottle standard deviation, which is lower than the degrees of freedom from the between-laboratory study, which limits the reliability of the individual data from one-way ANOVA. Figure 1 shows the ratios of the estimates for within-unit standard deviation and u_{bb} for evaluation by two-way ANOVA from the interlaboratory data (enumerator) and one-way ANOVA from the re-testing in one laboratory (denominator). By and large, within-unit standard deviations were lower in the re-testing. Several factors may contribute to this effect. On the one hand, a laboratory with a better than average repeatability was chosen for the re-testing. On the other hand, even if ANOVA aims to separate the between-sample variation from the variations due to the application of different methods, use of different instruments, analysis by different analysts, this separation is not perfect and the estimate for the between-bottle standard deviation can be affected by the variation due to the other factors. An effect of less relevance in the case of these very stable materials is potential degradation during transport: As the transport conditions (time and temperature) may differ for the different laboratories, an additional variation can be introduced for less stable materials. This variation should in theory also be part of the between-laboratory variation. Despite all these factors, the overall estimates for u_{bb} are roughly comparable. On average, u_{bb} estimated from the interlaboratory data is 54 % higher than the one estimated

from the single-laboratory data. This is partly caused by high estimates for La, Pd and Sb: without these elements, u_{bb} from the interlaboratory data is 30 % above the one from the single-laboratory data. In the case of these two materials, the higher uncertainty was still acceptable, as the contribution of u_{bb} to the final uncertainty was minor. This was caused by the larger uncertainty of characterisation. This is not necessarily true for other CRMs. Especially in cases where the uncertainty of characterisation is low, the higher uncertainty of homogeneity may have a significant influence on the final uncertainty.

This shows that the evaluation of homogeneity from characterisation data can yield estimates for the between-unit variation that are realistic and fit for purpose. The data also might indicate a trend to higher estimates of the uncertainty of homogeneity when the homogeneity estimates from characterisation data compared to those from a single-laboratory homogeneity study. However, this assessment has to be taken with caution as, due to the lower number of degrees of freedom of the single-laboratory data, the reliability of the individual estimates was lower in this case.

Evaluation of costs and samples saved

While the assessment of homogeneity using interlaboratory data was successful, the evaluation of the envisaged savings of samples and cost is more ambiguous. Standard procedure at the JRC for characterisation studies is to supply

Fig. 1 Ratio of s_{wb} and u_{bb} estimated from two-way ANOVA from interlaboratory data and from single-laboratory data

each laboratory with two samples which are each analysed in several replicates (often three per sample). In this case, three samples were sent to the laboratories in order to have sufficient degrees of freedom for the estimate of u_{bb} .

For the Ti-discs, this means that in total 24 samples were distributed. This is very close to the number of samples than would have been used in a classical characterisation study for nine data sets (18 samples) and a classical homogeneity study for such a small batch (ten samples). For the Ti-chips, 39 bottles were distributed. This is slightly more than would have been used in a classical characterisation study for 14 data sets (28 samples) and homogeneity study (ten samples). This means that the savings in term of samples did not materialise. In this case, as the samples are exceptionally stable, samples could have been reused for the characterisation. In this way, only 18 samples would have been used for the discs and 28 samples for the chips. (The ten samples from the homogeneity study would have been reused for the characterisation.)

According to offers received from private laboratories, performing classical homogeneity studies on the two materials with 25 elements quantified in duplicate on ten samples each would have cost roughly between EUR 10 000 and 20 000. While these sums are significant, this is only 5 % – 10 % of the total costs incurred for external services and naturally even less if the staff costs for planning, project management and evaluation are taken into consideration. Therefore, the cost saving of avoiding the homogeneity study may decrease the overall cost of the RMs by a few per cent, while the final uncertainty might increase.

Conclusion

The data obtained in the characterisation study of two Ti-materials were successfully used for the assessment of homogeneity. The estimates for the between-sample variation are low enough to make the two materials suitable as certified reference materials. This shows that it is in

principle possible to combine homogeneity assessment and interlaboratory characterisation into one study.

However, the set-up with several laboratories required a more complex statistical evaluation, with testing normalised data for trends and outlying individual data and sample means and subsequent testing of the non-normalised data for outliers of variances and performing the two-way ANOVA. The envisaged gains in terms of consumption of fewer samples were not achieved and the cost savings by skipping a dedicated homogeneity study were minor.

Based on the experience with interlaboratory characterisation studies of a wide variety of matrices and analytes, it is expected that these problems are generic rather than limited to the determination of trace metals in Ti. Careful evaluation of the benefits and costs of combining a homogeneity and an interlaboratory characterisation study is therefore advisable.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. International Organization of Standardization, ISO 17034 (2016) General requirements for the competence of reference material producers. International Organization of Standardization, Geneva
2. International Organization of Standardization ISO Guide 35 (2017) Reference materials: guidance for characterization and assessment of homogeneity and stability. International Organization of Standardization, Geneva
3. Linsinger TPJ, Pauwels J, van der Veen AMH, Schimmel H, Lambert A (2001) Homogeneity and stability of reference materials. *Accred Qual Assur* 6:20–25

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.