

The Second EWGRD Round Robin: Inter-Comparison of ^{93m}Nb Measurements

Dean Thornton^{a*}, James Watson^a, Hubert Carcreff^b, Christophe Domergue^c, Tomasz Dziel^d, Ana Fernandes^e, Anita Gerenyi^f, Linda Gravier^g, Vít Klupák^h, Vyacheslav Kochkinⁱ, Marie-Christine Lépy^j, Mladen Mitev^k, Vladimir Radulović^l, Ben Straathof^m, Jan Wagemansⁿ, Eva Zsolnay^f

^a Jacobs Engineering, Gloucester, United Kingdom

^b CEA, Service d'études des réacteurs et de mathématiques appliquées, Université Paris-Saclay, France

^c CEA, DES, IRESNE, DER, Cadarache, F-13108 St-Paul-lez-Durance, France

^d National Centre for Nuclear Research, Radioisotope Centre POLATOM, Poland

^e Instituto Superior Técnico, Universidade de Lisboa, Portugal

^f Budapest University of Technology and Economics, Hungary

^g Laboratoire d'Analyses Radiochimiques et Chimiques, CEA, France

^h Research Centre Řež, Czech Republic

ⁱ Kurchatov Institute, Moscow, Russia

^j Laboratoire National Henri Becquerel, CEA, France

^k Institute for Nuclear Research and Nuclear Energy of the Bulgarian Academy of Sciences, Bulgaria

^l Jožef Stefan Institute, Slovenia

^m NRG, Netherlands

ⁿ SCK-CEN, Belgium

Abstract. Following on from an initial Round Robin inter-comparison of gamma spectrometry measurements reported in 2014, this paper presents the results of the second part of a further Round Robin inter-comparison commissioned in 2018 by the European Working Group on Reactor Dosimetry. The purpose of the exercise was to demonstrate the level of consistency between different organisations' measurements of the $^{93}\text{Nb}(n,n')^{93m}\text{Nb}$ reaction, which plays a key role in underwriting reactor dosimetry assessments. To achieve this, measurements of ^{93m}Nb activity were performed by twelve European organisations on six sets of near identical niobium samples, each having its own geometry and tantalum concentration. The samples were provided by CEA, France and irradiated in the MARIA reactor, operated by National Centre for Nuclear Research, Poland. Participants provided their results to an independent referee who collated and compared the data. The inter-comparison has demonstrated agreement to within standard deviations ranging from $\pm 2.2\%$ to $\pm 7.9\%$, with a tendency for some organisations to measure elevated values. The results of the inter-comparison are presented in an anonymised form together with discussion and conclusions which may be drawn from the exercise.

1 Introduction

Over the last three decades, the neutron activation reaction $^{93}\text{Nb}(n,n')^{93m}\text{Nb}$ has come to play a key role in reactor dosimetry assessments because its energy dependence provides one of the best matches for fast neutron damage in steels. Reliant upon precision measurements of x-ray fluorescence, measurements of the ^{93m}Nb activation product demand specialist

* Corresponding author: dean.thornton@jacobs.com

techniques, sometimes requiring dissolution and hence destruction of some or all of the sample. Hitherto, the international reactor dosimetry community has not had an opportunity to benchmark measurements of this reaction.

This paper describes the results of an inter-comparison of ^{93m}Nb spectrometry measurements commissioned by the European Working Group on Reactor Dosimetry (EWGRD). It was organised as one of two separate inter-comparison exercises comprising the Second EWGRD Round Robin organised to follow the First EWGRD Round Robin reported to the 15th International Symposium on Reactor Dosimetry (ISRD) in 2014. [1]. The allied gamma activation inter-comparison is reported to the 17th International Symposium on Reactor Dosimetry (ISRD) in a separate paper [2].

Twelve organisations from ten European countries took part in the ^{93m}Nb inter-comparison, as follows:

- Budapest University of Technology and Economics, Hungary
- Instituto Superior Técnico, University of Lisbon, Portugal
- Institute for Nuclear Research and Nuclear Energy, Bulgaria
- POLATOM, Poland
- Laboratoire National Henri Becquerel, CEA, France
- Laboratoire d'Analyses Radiochimiques et Chimiques, France
- CEA Cadarache, France
- Kurchatov Institute, Russia
- Jožef Stefan Institute, Slovenia
- Research Centre Řež, Czech Republic
- NRG, Netherlands
- SCK-CEN, Belgium

2 Preparation and Irradiation of Samples

The niobium samples comprised 1 mg strip, and 8 mg disk geometries with tantalum impurity concentrations of 0.3 ppm, 4.0 ppm or 19.6 ppm for strips and 0.3 ppm and 4.0 ppm for disks. They were supplied by Laboratoire de Dosimétrie, de Capteurs et Instrumentation (LDCI), Cadarache, France. Each sample was given a name in the form Nb-n, where n is in integer in the range 1 to 61, not necessarily consecutive.

Tantalum impurities are known to affect x-ray fluorescence corrections in measurements of ^{93m}Nb activity. By providing samples with three different concentrations, the intention was to provide evidence for the impact, if any, of tantalum on the measurement results. Of the concentrations used: 0.3 ppm may be considered as pure; 4 ppm as a “standard” or typical concentration and 19.6 ppm considered as an impure sample.

The irradiations took place from 13th to 23rd October 2017 in the MARIA reactor operated by National Centre for Nuclear Research (NCBJ), Poland. In total 43 samples were provided for measurement, of which 38 were used in the inter-comparison.

3 Protocol

Because a significant proportion of the participating organisations rely on techniques requiring sample dissolution, it was not possible to carry out an inter-comparison along traditional lines in which each participant measures the same set of samples. Instead, distinct batches of near identical samples, each of which was subjected to the same irradiation conditions.

Using a non-destructive technique, one of the participating organisations was responsible for measuring each sample with high fidelity to provide a control for any variation between samples of the same batch. Once complete, subsets of samples were sent to each participant for measurement, with some samples retained as a precaution against loss.

Results were sent back to a non-participating member of the EWGRD who acted as

referee by collating and analysing the data. The data were treated anonymously, and the participants identified as Org. A to Org. L, inclusive.

Participants were requested to record sample mass measurements and specific activities, together with corresponding uncertainties within a pre-prepared template, prescribing the reference time of the irradiation (against which specific activity data were to be decay-corrected), the units of measurement and number formats required. Uncertainties (at one standard deviation, 1σ) were requested as the combination Type A, i.e. those that can be treated using statistical methods, and Type B, i.e. those that cannot.

Participants were also asked to provide:

- Sample mass measurements and uncertainties;
- Sample activities, and specific activities (activity divided by mass);
- A brief description of measurement technique(s) and equipment, including any relevant software;
- Identify any corrections applied or measures taken (e.g. for dead time, self-absorption, fluorescence correction);
- Calibration, including any standard sources;
- Nuclear data used: i.e. what libraries were used or, where allocated separately, the values of half-lives and photon emission yields assumed in the analysis;
- A short description of the treatment of uncertainties.

4 Sample Mass Measurements

Table 1 shows the identification (ID) number, properties, mass measurements and associated uncertainties for each of the samples included in the inter-comparison, as well as the organisations measuring each sample. Sample properties are defined in terms of geometry, mass and approximate tantalum (Ta) concentration, which constituted six batches of samples (B1 to B6), each of which subject to the same irradiation conditions.

Organisation L was the one responsible for measuring all samples. Although the samples retained by Org. L were also measured, they are not included in Table 1 and have not played any role in the inter-comparison, other than as control for sample-to-sample variation.

On the whole, the mass measurements given in the table show excellent agreement between each pair of organisations. However, for the reasons given below, there are limitations on what can be inferred from the data:

- With only two values per sample, it is difficult to make statistical inferences on the accuracy of any individual's mass measurements as done in the accompanying EWGRD inter-comparisons on gamma spectrometry measurements [1] and [2].
- Org. K used dissolution techniques and only weighed a fraction of specimens Nb-41 and Nb-42. The reported mass values are substantially lower than the corresponding values reported by Org. L.
- Several organisations reported that samples had not arrived intact (Nb-4, Nb-31, Nb-45, Nb-46, Nb-51 and Nb-52), some in multiple pieces. Consequently, there is uncertainty that the recipient organisations measured all material weighed by Org. L and for these samples, the comparison of mass measurements may not be relevant, as shown by the systematically lower value of the second measurement in some circumstances.

Of the 38 pairs of data, 30 make no reference to sample fragmentation (third bullet) nor division of the sample for dissolution (second bullet). Within this subset, analysis of the data presented in Table 1 showed that 26 agreed within ± 3 standard deviations (σ_m) of each other and 21 pairs show agreement within $\pm 1\%$. Of the four outside $\pm 3 \sigma_m$ three were measured by Org. E (Nb-7, Nb-18 and Nb-35) suggesting issues with mass measurements and/or

Table 1 Details of Samples and Mass Measurements

Specimen ⁽¹⁾	Description	Meas. by	Value 1 (mg)	Value 2 (mg) ⁽⁴⁾	Specimen ⁽¹⁾	Description	Meas. by	Value 1 (mg)	Value 2 (mg) ⁽⁴⁾	
Nb-1	Batch 1 (B.1) 1 mg strip, 0.3 ppm	L, G	0.882±0.008	0.887±0.007	Nb-28	Batch 4 (B.4) 8 mg disk, 0.3 ppm	L, G	9.103±0.008	9.102±0.009	
Nb-2		L, J	0.912±0.008	0.915±0.008	Nb-30		L, D	8.615±0.009	8.610±0.020	
Nb-3		L, B	0.957±0.009	1.000±0.100	Nb-31 ⁽²⁾		L, E	9.062±0.009	(8.170±0.002)	
Nb-4 ⁽²⁾		Batch 2 (B.2) 1 mg strip, 4 ppm	L, H	0.977±0.008	(0.623±0.022)	Nb-41 ⁽³⁾	Batch 5 (B.5) 8 mg disk, 0.3 ppm	L, K	9.048±0.009	(1.485±0.045)
Nb-5			L, I	0.824±0.008	0.820±0.020	Nb-42 ⁽³⁾		L, K	8.805±0.009	(1.290±0.039)
Nb-6			L, F	0.883±0.008	0.884±0.001	Nb-45 ⁽²⁾		L, C	8.812±0.009	(7.636±0.003)
Nb-7			L, E	0.925±0.008	0.880±0.002	Nb-46 ⁽²⁾		L, C	9.053±0.008	(8.952±0.004)
Nb-10	L, G		1.087±0.008	1.084±0.007	Nb-51 ⁽²⁾	L, A		9.144±0.010	(2.515±0.003)	
Nb-11	L, J		1.165±0.008	1.165±0.008	Nb-52 ⁽²⁾	L, A		9.057±0.008	(2.866±0.003)	
Nb-12	L, B		0.980±0.008	1.000±0.100	Nb-33	L, G		8.783±0.008	8.781±0.009	
Nb-13	Batch 3 (B.3) 1 mg strip, 19.6 ppm	L, H	1.202±0.009	1.209±0.025	Nb-34	Batch 6 (B.6) 8 mg disk, 4 ppm	L, B	9.320±0.009	9.300±0.100	
Nb-14		L, I	1.261±0.008	1.260±0.020	Nb-35		L, E	8.819±0.009	9.130±0.002	
Nb-15		L, F	1.087±0.009	1.087±0.002	Nb-36		L, C	9.503±0.008	9.501±0.004	
Nb-18		L, E	1.380±0.008	1.170±0.002	Nb-37		L, C	9.430±0.008	9.427±0.004	
Nb-19		L, G	1.256±0.008	1.255±0.007	Nb-40		L, D	9.065±0.008	9.020±0.020	
Nb-20		L, J	0.871±0.008	0.858±0.008	Nb-61		L, A	8.977±0.008	8.975±0.003	
Nb-21		L, B	1.025±0.008	1.000±0.100	Nb-62		L, A	9.011±0.009	9.010±0.003	
Nb-22	Batch 3 (B.3) 1 mg strip, 19.6 ppm	L, H	1.193±0.009	1.188±0.025	Notes: (1) Only samples measured by two organisations are shown. (2) Found to be broken into multiple fragments on receipt. (3) Only fraction of sample measured. (4) Mass measurements performed on incomplete samples are shown in parenthesis because their comparison with Org. L is not valid.					
Nb-23		L, I	1.068±0.008	0.980±0.020						
Nb-24		L, F	0.738±0.009	0.738±0.001						
Nb-25		L, E	1.042±0.008	1.020±0.002						

Table 2 Details of Specific Activity Measurements Provided by Each Participant

		Org L ⁽²⁾	Org K ⁽³⁾	Org J	Org I	Org H	Org G	Org F	Org E	Org D	Org C ⁽³⁾	Org B ⁽¹⁾	Org A ⁽³⁾	Mean ⁽¹⁾ σ_{pop}
Technique: solid (S) or dissolution (D)		S	D	D	D	S	S	D	D	S	D	S	D	n/a
Batch 1 1mg strip 0.3ppm	Meas. Act.	2.57E+07		2.55E+07	2.75E+07	2.79E+07	2.60E+07	2.81E+07	3.10E+07			<i>5.00E+07</i>		2.74E+07
	$\sigma_{meas.}$	2.6%		6.7%	3.3%	4.7%	3.5%	3.9%	0.9%			<i>20%</i>		7.0%
	Δ to mean	-6.2%		-6.9%	0.3%	1.9%	-4.9%	2.7%	13.2%			<i>82.6%</i>		
Batch 2 1mg strip 4 ppm	Meas. Act.	2.59E+07		2.59E+07	2.74E+07	2.81E+07	2.56E+07	2.76E+07	3.50E+07			<i>5.00E+07</i>		2.68E+07
	$\sigma_{meas.}$	2.6%		7.3%	2.7%	3.9%	3.5%	3.9%	1.1%			<i>20%</i>		4.0%
	Δ to mean	-3.0%		-3.2%	2.3%	5.0%	-4.4%	3.3%	30.8%			<i>86.9%</i>		
Batch 3 1mg strip 19.6 ppm	Meas. Act.	1.82E+07		1.70E+07	1.77E+07	2.07E+07	1.76E+07	1.83E+07	2.01E+07			<i>3.30E+07</i>		1.85E+07
	$\sigma_{meas.}$	2.6%		7.1%	3.1%	3.9%	3.6%	4.0%	1.1%			<i>24%</i>		7.4%
	Δ to mean	-1.5%		-8.2%	-4.6%	11.8%	-5.0%	-1.1%	8.6%			<i>78.3%</i>		
Batch 4 8mg disk 0.3ppm	Meas. Act.	2.57E+07					2.60E+07		2.95E+07	2.99E+07				2.78E+07
	$\sigma_{meas.}$	2.6%					3.4%		1.0%	11.5%				7.9%
	Δ to mean	-7.3%					-6.3%		6.2%	7.4%				
Batch 5 8mg disk 0.3 ppm	Meas. Act.	1.74E+07	1.78E+07								1.73E+07		1.78E+07	1.76E+07
	$\sigma_{meas.}$	2.6%	5.1%								4.0%		2.8%	2.2%
	Δ to mean	-1.1%	1.0%								-1.7%		1.3%	
Batch 6 8mg disk 4 ppm	Meas. Act.	8.08E+06					7.97E+06		8.83E+06	9.35E+06	7.92E+06		8.15E+06	8.30E+06
	$\sigma_{meas.}$	2.6%					3.4%		1.2%	11.4%	4.0%		2.8%	6.2%
	Δ to mean	-2.6%					-4.0%		6.4%	12.7%	-4.5%		-1.8%	
Notes:	(1)	Org. B (values in italic) not included in calculation of mean and σ_{pop} .												
	(2)	For Org. L, values given are mean of measurements on all samples in batch.												
	(3)	Each of the values shown for Org. K, Org C, and Org. A are the average of measurements performed on two samples in each group.												

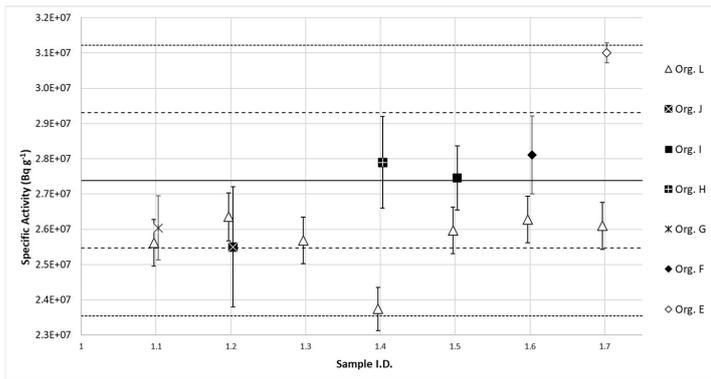


Fig. 1 ^{93m}Nb Specific activity measurements obtained for B.1, 1mg strip with 0.3ppm Ta

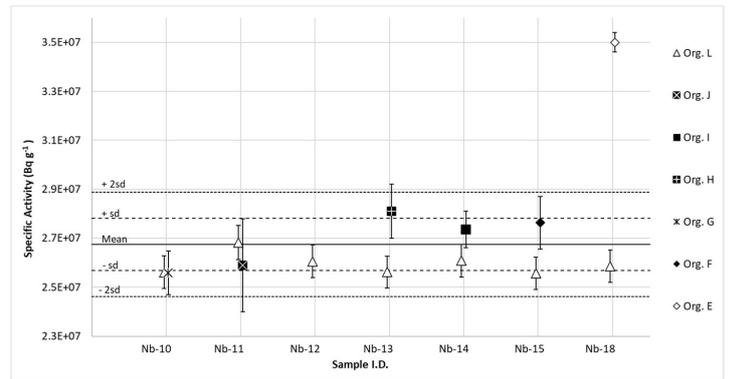


Fig. 2 ^{93m}Nb Specific activity measurements obtained for B.2, 1mg strip with 4 ppm Ta

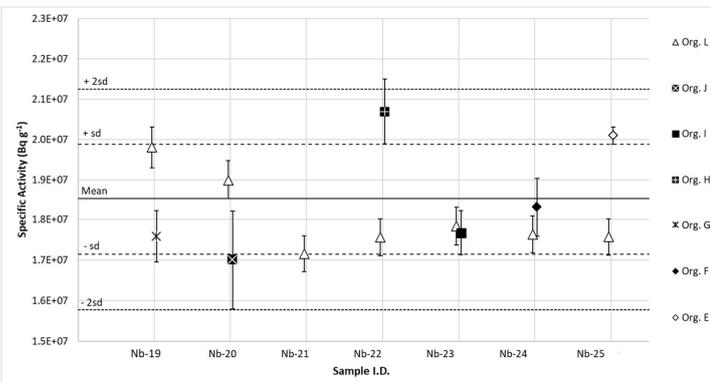


Fig. 3 ^{93m}Nb Specific activity measurements obtained for B.3, 1mg strip with 19.6 ppm Ta

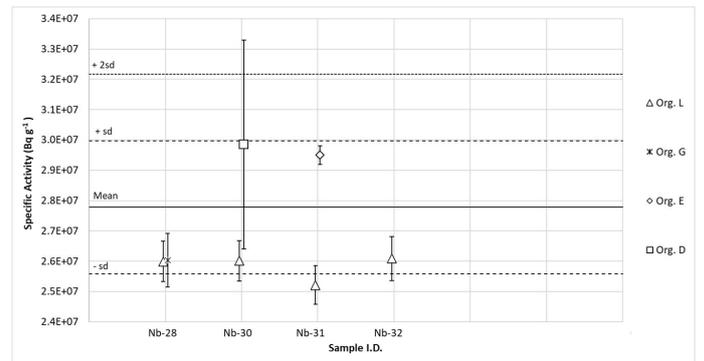


Fig. 4 ^{93m}Nb Specific activity measurements obtained for B.4, 8mg disk with 0.3 ppm Ta

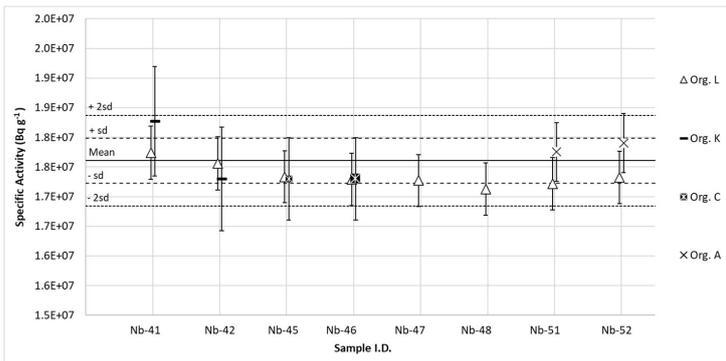


Fig. 5 ^{93m}Nb Specific activity measurements obtained for B.5, 8mg disk with 0.3 ppm Ta

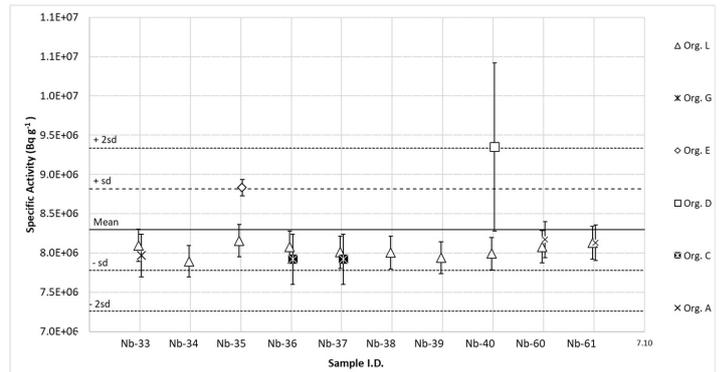


Fig. 6 ^{93m}Nb Specific activity measurements obtained for B.6, 8mg disk with 4 ppm Ta. Note, Nb-33 was measured by Org. G

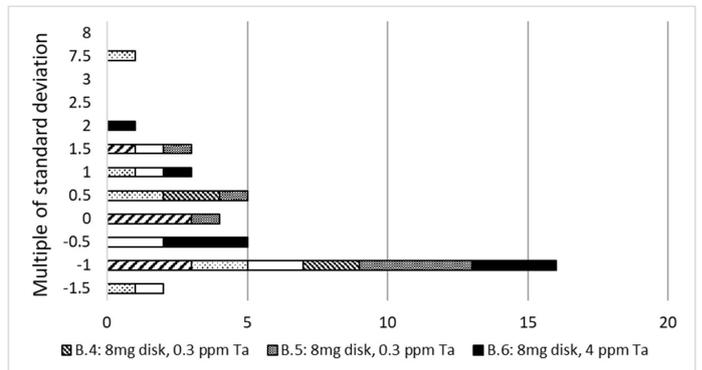
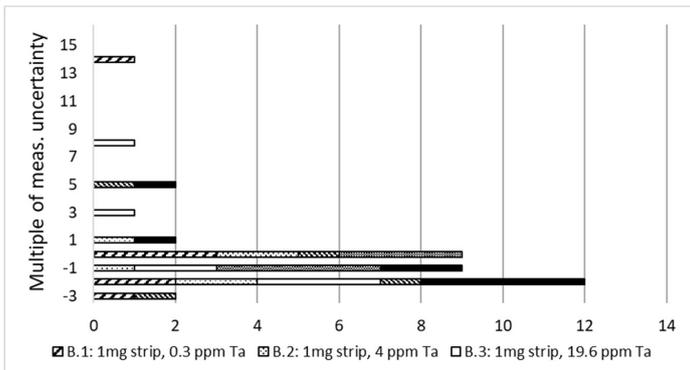


Fig. 7 Distributions of discrepancies from mean relative to measurement uncertainties, σ_{meas} (on left) and relative to standard deviation, σ_{pop} (on right). The legend applies to both graphs.

inadequate allocation of uncertainties. However, only Nb-18 and Nb-23 are more than $\pm 5\%$ different: Nb-18 (Org.L and Org E) with a difference of $+17.9\%$ and Nb-23 (Org. L and Org. I) a difference of $+9.0\%$. If the data for these two samples represent a genuine difference in the measured masses, they are large enough to significantly affect the corresponding comparison of specific activity measurements (see Sections 6 and 6.1).

5 Specific Activity Measurements

Individual specific activity measurements of samples from each of the six batches are compared in Figures 1 to 6, which provide a graphical comparison of the specific activities obtained for each batch. Each figure shows the specific activity measured by each participant against the sample number. Error bars show the individual measurement uncertainty provided by the participant at one standard deviation. The horizontal lines plotted show the mean and multiples of the population standard deviation (σ_{pop}) for the batch, given in Table 2 as described below.

Table 2 compares the $^{93\text{m}}\text{Nb}$ specific activity results for each batch from each of the participating organisations. Note: except for Organisation L, participants measured a subset of batches, hence the empty cells in the table. For each organisation, the table provides the measurement technique used; the specific activity obtained and its associated measurement uncertainty (σ_{meas}); and the deviation (Δ) of the measured activity from the mean for the batch, expressed as a percentage. In order to provide a single representative value for each organisation's batch measurements averages were taken as follows:

- For Org. L, the data presented represent the mean of all samples within the batch (excluding retained samples);
- For Orgs. A, C and K, because each received pairs of samples in the same batches, the data shown in Table 2 for these participants represent the mean of a pair of samples.

In its righthand column, Table 2 shows the mean specific activity obtained for each batch, together with the standard deviation of the data for that batch, σ_{pop} . In calculating the mean and standard deviation, the Org. L data were treated as a single point represented by its mean (excluding retained samples). This was done to prevent Org. L having a dominant influence on the analysis.

From the data provided in Table 2, it is evident that the specific activity data provided by Org. B are significantly greater than their respective batch means, with substantially greater uncertainties for all samples measured. This was discussed with the correspondent who confirmed that $^{93\text{m}}\text{Nb}$ measurement is a technique in development at that organisation, for which the inter-comparison provided a useful exercise. With the participant's consent, it was deemed that Org. B's data were not representative of the data as a whole and would not be included in the global analysis (mean and standard deviations) nor the intercomparison of the data. These data are not shown in Figures 1 to 6.

The available measurements from all individual samples are represented in Figures 1 to 6, including those performed by Org. L on the retained samples, and the pairs of data from Orgs. A, C and K which were averaged in Table 2.

6 Discussion and Analysis

For the first batch of samples (1 mg strip with 0.3 ppm Ta), Figure 1 shows a generally good level of agreement, with all measurements lying within $\pm 2\sigma_{\text{pop}}$ of the mean value for this batch, and only two values outside the standard deviation, σ_{pop} , of $\pm 7\%$. Compared to the population of data plotted in Figure 1, the position of the mean appears higher than might be

expected from the disposition of data. However, it is important to recognise that seven of the data points lying below it are from Org. L, of which only the average value contributes to the mean and σ_{pop} .

With the exception of Nb-4, measurements performed by Org. L on all seven samples show consistency within their associated measurement uncertainties of 2.6%, i.e. no evidence of sample-to-sample variation. Whilst they are both within $\pm 2 \sigma_{\text{pop}}$, Org. E's measurement of Nb-7 and Org. L's value for Nb-4 are clearly the most outlying. The presence of a more discrepant value in the Org. L data does not adversely affect the mean value and σ_{pop} for this batch because it only represents one value out of the six used to calculate the average measurement of Org. L. By contrast, Org. E's outlying measurement of Nb-7 does affect the mean and standard deviation for this batch. Excluding this outlier would reduce the mean from $2.74 \times 10^7 \text{ Bq cm}^{-2}$ to $2.68 \times 10^7 \text{ Bq cm}^{-2}$, a reduction of 2.2% and a reduction in from 7% to 4%. In part, at least, Org. E's measurement of Nb-7 is likely to have been affected by the mass measurement which was 4.9% lower than that obtained by Org. L.

Figure 2 shows a similar pattern for the measurements performed on the second batch, 1 mg strip with 4 ppm Ta. The measurements performed by Org. L on each of these samples are mutually consistent within measurement uncertainties (σ_{meas}) of $\pm 2.6\%$, indicating no sample-to-sample variation. Although the agreement between organisations is generally good, there is a potentially significant outlier on the measurement of Nb-18 by Org. E. Being some 30% greater than all the other measurements performed on this batch, Org. E's measurement of Nb-18 is considered anomalous and is associated with the 17% difference in mass measurements reported in Table 1. As a result, it is excluded from the calculated mean and σ_{pop} for this batch given in Table 2 and represented in Figure 2. With this point rejected, all but five data lie within $\pm \sigma_{\text{pop}}$ (4%) and those that don't lie comfortably within ± 2 Further consideration of Nb-18 is given in Section 6.1.

In Figure 3, Org. L's measurements of the third batch, 1 mg strip with 19.6 ppm Ta, show their greatest variability, with their own standard deviation of 5.2% and differences amongst them exceeding $2\sigma_{\text{meas}}$, albeit with all data less than 7.3% from the mean. As for the other participants, there is good agreement with all measurements within $\pm 2\sigma_{\text{pop}}$ of the mean and only three more than σ_{pop} from the mean. It is noted that despite the potentially discrepant mass measurement by Org. I for Nb-23 (Table 1), the specific activity obtained is in good agreement with the rest of the batch.

The fourth batch, shown in Figure 4, is the first of two sets of 8 mg disk 0.3 ppm Ta specimens. Being the smallest batch, it contains four specimens, of which only three were measured by two participants, and it has the largest σ_{pop} value of 7.9%. Compared to a measurement uncertainty of $\pm 2.6\%$, data from Org. L for each of these samples are mutually consistent suggesting no sample-to-sample variability. For the other participants, this batch appears to provide the most polarised inter-comparison with data split between two high values from Orgs. D and E (Nb-30 and Nb-31), of which the former has a significantly larger σ_{meas} of $\pm 11.5\%$, and a markedly lower value from Org. G (Nb-28) with modest σ_{meas} noted, however, that with its relatively large uncertainty, Org D's measurement is nonetheless consistent with the lower values. Finally, Org. E's mass measurement for Nb-31 is 10% lower than the corresponding Org. L value suggesting that had Org. E obtained the same mass measurement, the agreement would have been much better although still high.

For the fifth batch (8mg disk 0.3 ppm Ta), Figure 5, the absence of sample-to-sample variation is demonstrated by Org. L's measurements, which show mutual consistency within $\pm\sigma_{\text{pop}}$ of 2.2% compared to measurement uncertainties of $\pm 2.6\%$. Overall, the batch shows excellent agreement over eight samples, fourteen individual measurements and four organisations (Orgs. A, C, K and L). The level of agreement is represented by the lowest σ_{pop} (2.2%) and the fact that all but three of the individual measurements lie between the mean and $\pm\sigma_{\text{pop}}$. Unlike the first four batches, this batch contained three pairs of samples measured by the same organisations, namely: Nb-41 and Nb-42 (Org. K), Nb-45 and Nb-46 (Org. C) and Nb-51 and Nb-52 (Org. A). It is therefore interesting to note that in each pair, variations seen in the corresponding Org. L data appear to be reproduced, including the modest shape apparent at either end of the data.

Figure 6 shows the specific activity measurements for the sixth batch (8mg disk 4 ppm Ta samples). The largest of the six batches with ten samples plotted, the data from Org. L are all consistent within their measurement uncertainties of $\pm 2.6\%$, confirming the absence of sample-to-sample variability. There is excellent agreement between the fifteen measurements provided by Orgs. A, C, G and L, all of which agree within $\pm 1.4\%$ of $8.08 \times 10^6 \text{ Bq g}^{-1}$. However, results for Nb-35 and Nb-40 (Orgs. E and D, respectively) are more discrepant, contributing to higher mean for the batch as a whole ($8.30 \times 10^6 \text{ Bq g}^{-1}$) and a significantly larger σ_{pop} of $\pm 6.2\%$.

6.1 Potential Outliers

Discussion of the specific activity results in Section 6 has identified potentially outlying values, in particular Nb-18. To determine whether any of the data should be considered as outliers, statistical tests have been performed on the results of each of the batches using both Inter-Quartile Range and Grubbs test [3]. The results of both tests confirmed that only Org. E's result for Nb-18 should be treated as an outlier, justifying its exclusion from the mean and σ_{pop} calculated for Batch 2.

Further investigation into Org. E's result for Nb-18 showed that had its specific activity been calculated with the mass value measured by Org. L, the specific activity measurement would reduce from $3.50 \times 10^7 \text{ Bq g}^{-1}$ to $2.96 \times 10^7 \text{ Bq g}^{-1}$. At this value, it would be $2.7 \sigma_{\text{pop}}$ greater than the mean, but would not be indicated as an outlier in either of the tests applied.

6.2 Distribution of Discrepancies

Figure 7 presents histograms showing the distributions of measurement discrepancies relative to the mean for each batch expressed as multiples of σ_{meas} (to the left) and σ_{pop} (to the right). The data are presented so that the distributions are evident both for individual batches and the comparison as a whole.

Compared to measurement uncertainties, it is clear that the majority of data (34 out of 40) are clustered between $-3\sigma_{\text{meas}}$ and $+1\sigma_{\text{meas}}$, suggesting that for these measurements, at least, the uncertainties allocated are representative of the measurement precision. However, the remainder of the data (including one not shown because it is out of range of the graph) appear significantly high creating a somewhat skewed distribution.

Unsurprisingly, the comparison of discrepancies relative to standard deviation, σ_{pop} , provides a narrower distribution for the obvious reason that σ_{pop} represents the distribution of all of the data. However, the comparison relative to σ_{pop} provides additional insight showing the distribution to be very peaked at $-1\sigma_{\text{pop}}$ with 23 out of the 40 values concentrated in the three intervals centred on this, with the rest of the population lying above. This suggests that:

- The asymmetric tendency to measure high is present in a substantial proportion of the data, and not just the more obviously discrepant values identified in the comparison of relative to σ_{meas} ;
- The asymmetry observed in the discrepancies has pulled the mean upwards, with the consequence that the modal values (represented by the peak at $-1\sigma_{\text{pop}}$) appear as under-measurement when, in fact, they are most probably accurate.

6.3 Normalised Results

Table 3 compares the relative performance of each organisation by normalising the value it obtained for each batch, S_i , relative to the mean of all measurements for that batch and expressing it as a deviation (in %), i.e. $[S_i - \bar{S}_i] / \bar{S}_i$. The mean batch values used in the denominator here are those in Table 2 and indicated by the solid horizontal lines plotted in Figures 1 to 6. In doing this, the averages obtained over all of the batches measured represent the overall performance of each organisation, effectively aggregating the observations made in the section above on Figures 1 to 6. In addition to this, σ_{org} , the standard deviation of each organisation's normalised results is given, to provide an indication of the variability of each organisation's performance from batch to batch. The average value of each organisation's measurement uncertainty, σ_{meas} , is also included to help gauge the significance of these results. The key observations are as follows:

- The majority (7 out of 11) of organisations' mean discrepancies are less than their σ_{meas} , supporting the uncertainties allocated. Of those that aren't, two are comparable and the other is Org. E the most discrepant organisation.
- Excluding the value of 10.3% for Org. E, values of σ_{org} are between 0.9% and 5.1%, lower than the range of σ_{pop} values, which lie between 2.2% and 7.4% (Table 2). This suggests that each individual organisations' results tend to be systematically removed from the consensus (to some extent) with some variation around that represented by σ_{org} . In contrast, the batch-to-batch variability is larger because systematic differences embedded within each organisation's data are not correlated from one to the other. This was also observed in [1].
- Org. E appears to systematically over-predict all of the samples it measured, by on average 13%. and significantly greater than the declared measurement uncertainties. This is evident as a clear outlier in Figure 2, as well as values markedly above the distribution in Figures 1, 3, 4 and 6.
- Org. D over-predicts both of the samples it measured by on average 10%, lying clearly above the distribution of data in Figures 4 and 6. However, this is mitigated by substantially larger measurement uncertainties.
- Mean normalised results for all other organisations are mutually reconcilable within $\pm 2\sigma_{\text{meas}}$.

To assess the impact of the two most discrepant organisations on the inter-comparison as a whole, the final row in Table 3 shows how the mean normalised deviation for each organisation is changed if Org. D and Org. E are excluded from the calculation of the mean value for each batch. The result is a tighter distribution amongst the remaining organisations with noticeably less indication that they have under-measured relative to the mean. Doing this does raise the possibility that Org. H has also over-measured, being the only one of the remaining participants to be more than σ_{meas} from the mean.

6.4 Sensitivity to Measurement Methods

Table 3 also records the method used by each participant showing that a subset of seven used sample dissolution techniques, "D" and a subset of four used solid sample techniques, "S". For each subset, it provides the average of the normalised specific activity deviations (i.e. $[S_i - \langle S_i \rangle_{\text{batch}}] / \langle S_i \rangle_{\text{batch}}$) for each batch as well as the data as a whole.

t-Test results for each batch, and the data as a whole, indicate the likelihood of the subsets of data from the two methods having the same mean. Although Batches 1 and 2 appear to show differences between mean values of the the two subsets greater than 5%, the t-Test results suggest little to no difference evident between any of the individual batch results, and none for the data overall. It is important to note that the distribution of participants between batches is not consistent (e.g. Batches 4, 5 and 6 were measured by fewer participants), which this has the potential to influence the outcome of this comparison from batch to batch.

Whilst limitations in the dataset are acknowledged, the inference reached here is that there is no evidence to suggest systematic differences between organisations on the basis of the techniques deployed.

6.5 Tantalum Concentration

One of the objectives of the inter-comparison was to provide batches of specimens with a range of tantalum concentrations to determine whether this had any bearing on the ability of participants to provide consistent results, whatever the counting technique used. Despite the limitation on the distribution of batches between the participants identified above, it would be expected that any adverse effects of increasing tantalum would be evident as increased variability in the results and/or higher numbers of outliers. On inspection, the data presented in Table 2 show no evidence of any increase in σ_{pop} associated with increasing tantalum concentration, and a qualitative review of Figures 1 to 6 shows no greater tendency for there to be more outlying values in the batches with highest concentrations. These observations were supported by simple regression of σ_{pop} and batch range (i.e. highest minus lowest) vs. tantalum concentration, which indicated no significant correlation.

Therefore, whilst recognising its limitations, this inter-comparison exercise has produced no evidence that tantalum concentrations up to c. 20 ppm have affected the accuracy with which participants have measured the specific activity of $^{93\text{m}}\text{Nb}$.

Table 3 Normalised Specific Activity Measurements

	Org L	Org K	Org J	Org I	Org H	Org G	Org F
Technique	S	D	D	D	S	S	D
Batch 1	-6.2%	-	-6.9%	0.3%	1.9%	-4.9%	2.7%
Batch 2	-3.0%	-	-3.2%	2.3%	5.0%	-4.4%	3.3%
Batch 3	-1.5%	-	-8.2%	-4.6%	11.8%	-5.0%	-1.1%
Batch 4	-7.3%	-	-	-	-	-6.3%	-
Batch 5	-1.1%	1.0%	-	-	-	-	-
Batch 6	-2.6%	-	-	-	-	-4.0%	-
Mean.	-3.6%	1.0%	-6.1%	-0.7%	6.2%	-4.9%	1.6%
$\sigma_{org}^{(2)}$	$\pm 2.6\%$	n/a	$\pm 2.6\%$	$\pm 3.5\%$	$\pm 5.1\%$	$\pm 0.9\%$	$\pm 2.4\%$
Av. σ_{meas}	$\pm 2.6\%$	$\pm 5.1\%$	$\pm 7.0\%$	$\pm 3.0\%$	$\pm 4.1\%$	$\pm 3.5\%$	$\pm 3.9\%$
Rev. Mean⁽¹⁾	-1.5%	1.0%	-4.9%	0.5%	7.6%	-2.3%	2.9%
	Org E	Org D	Org C	Org A	Mean D	Mean S	t-Test prob
Technique	D	S	D	D			
Batch 1	13.2%	-	-	-	2.3%	-3.1%	32%
Batch 2	30.8%	-	-	-	8.3%	-0.8%	33%
Batch 3	8.6%	-	-	-	-1.3%	1.8%	65%
Batch 4	6.2%	7.4%	-	-	6.2%	-2.1%	n/a
Batch 5	-	-	-1.7%	1.3%	0.2%	-1.1%	n/a
Batch 6	6.4%	12.7%	-4.5%	-1.8%	0.1%	2.0%	77%
Mean.	13.0%	10.1%	-3.1%	-0.2%	0.8%	1.9%	72%
$\sigma_{org}^{(2)}$	$\pm 10.3\%$	$\pm 3.7\%$	$\pm 2.0\%$	$\pm 2.1\%$			
Av. σ_{meas}	$\pm 1.1\%$	$\pm 11.5\%$	$\pm 4.0\%$	$\pm 2.8\%$	-	-	-
Rev. Mean⁽¹⁾	16.0%	15.5%	-1.6%	1.4%	2.2%	4.8%	64%
Note: (1) Results normalised vs. batch average values excluding Org. D and Org. E. (2) σ_{org} is the standard deviation of normalised specific activity measurements for each organisation. (3) Where t-Test results are not given, there were insufficient data in one of the subsets.							

7 Suggestions for Improvement

As the first international Round Robin inter-comparison of the Dosimetry Community’s facilities for the measurement of the activation product ^{93m}Nb , this exercise has provided valuable data and insights. Because of the relatively late adoption of this radionuclide by the Reactor Dosimetry Community, and because of the challenges associated with performing multi-participant measurements on samples that have traditionally required destructive techniques, it is unsurprising that such an activity has not been undertaken previously.

Being first in kind, it is also not surprising that, with the benefit of hindsight, there are ways that this exercise could be improved in any future inter-comparison. The points below provide a non-exhaustive list of the more important lessons learned:

- Although the technique of irradiating batches of near identical samples was successful, the batch technique would have been much more effective if each participant had received and measured an equivalent set of samples, each drawn from the same of batches. The inconsistent distribution of samples between the participants has weakened some of the conclusions of the exercise.
- Increasing the numbers of dosimeters received and measured by each participant would have improved the statistical significance of the results.
- Recognising that about a third of participants used non-destructive (solid) techniques, the sharing of samples between those organisations could have been possible before providing to an organisation using

dissolution techniques. Had this been realised, it would have been possible to make the limited number of samples go further between the participants and may have made the conclusions more compelling.

- Whilst the provision of samples with different geometries has provided important variations between the measurements performed, a more systematic investigation of the impacts of geometry would require careful design. In particular, future inter-comparisons would benefit from a greater range of specific activation values, including some that would stretch the abilities of the facilities.
- A more conclusive investigation into the effects of tantalum concentrations would have been obtained if there had been a better distribution of tantalum values between the batches, with more at higher concentrations in particular.
- Damaged samples should not be forward to participants, if it can be avoided. To that end, an increase in the number of samples irradiated would provide some mitigation against the consequences of damaged specimens (dosimeters broken during their recovery from irradiation could have been substituted).
- A protocol for dealing with the most outlying measurements should be agreed in advance.

8 Conclusions

Following on from an initial Round Robin inter-comparison of gamma spectrometry measurements reported in 2014, the European Working Group on Reactor Dosimetry has undertaken a Round Robin exercise to compare measurements of the important activation product ^{93m}Nb reaction, which plays a key role in underwriting reactor dosimetry assessments.

The results presented anonymously in this paper show that:

- An inter-comparison of ^{93m}Nb specific activity measurements performed by 12 European Organisations has been successfully carried out, overcoming the challenge presented by the use of destructive (i.e. dissolution) techniques by the majority of participants.
- The technique of irradiating batches of near identical samples was successful. The use of a single organisation to provide a control against sample-to-sample variation was a key part of this.
- Comparisons performed on measurements from six batches of samples (each with its own geometry and tantalum concentration) show agreement to within standard deviations ranging from $\pm 2.2\%$ to $\pm 7.9\%$.
- The majority of measurements are clustered within around 2% of the modal value for their batch.
- The distribution of measurements relative to their respective batch mean values is skewed, with few measurements below the modal value, and a tail comprising a small but significant minority of participants tending to over-measure.
- Although mass measurements were shown to be consistent for the majority of specimens, some of the most discrepant mass data were shown to be associated with discrepant specific activity measurements.
- From the data provided, there is no evidence to suggest that either measurement technique or tantalum concentration have had any impact on the accuracy of the measurements.

Finally, the inter-comparison described in this paper has provided valuable data and insights into the measurement of this key radionuclide. However, being first in kind, there are a number of ways that this exercise could be improved, seven of which are described in the paper.

References

1. D.A. Thornton, J. Wagemans, A.C. Fernandes, J.M. Girard, D.P. Kis, V. Klupák, R. Mutnuru, H. Philibert, G. Rousseau, J.P. Santos, T Serén, E. Zsolnay, *The EWGRD Round Robin Measurement Exercise*, submitted to the 15th International Symposium on Reactor Dosimetry, EPJ Web of Conferences **106**, 06005 (2016)
2. D.A. Thornton, C. Domergue A. Fernandes P. Frajtag, P. Haroková, D.P. Kis, V. Klupák, V. Kochkin M. Lépy, M. Mitev, R. Prokopowicz, V. Radulović, B. Straathof, J. Wagemans, E. Zsolnay, *The Second EWGRD Round Robin: Inter-Comparison of Gamma Spectrometry Measurements on Activation Dosimeters*, submitted to the 17th international Symposium on Reactor Dosimetry, Lausanne, 2023
3. F.E Grubbs, *Procedures for Detecting Outlying Observations in Samples*, *Technometrics*, Vol. 11, No. 1, February 1969

Acknowledgements

The EWGRD would like to thank the participating authors and their organisations for their support in the completion of the exercise and the production of this paper. In particular, the authors would like to acknowledge the support of CEA, NCBJ and SCK-CEN whose commitment and initiative made the exercise possible.