EURAMET TC Project
Final Report

Project Title

Evaluation of repeatability measurement procedures in gravimetric volume calibrations of glassware

Coordinator, Institute, Country

Elsa Batista, IPQ (Portugal

EURAMET Registration No.

1525

Subject Field

Flow

KCDB Identifier

--

Date

2024-03

Instituto Português da Qualidade

PORTUGUESE INSTITUTE FOR QUALITY
Rua António Gião, 2
2829-513 CAPARICA
PORTUGAL
Tel (+ 351) 21 294 81 00 *E-mail: ipq@ipq*
Fax (+ 351) 21 294 81 01 *URL: www.ipq.p*

Final Report V2

Evaluation of repeatability measurement procedures in gravimetric volume calibrations of glassware

EURAMET Project no. 1525
Cooperation in research

Coordination: Elsa Batista, IPQ, Portugal
Participants: Andrea Malengo, INRIM, Italy
Urška Turnšek, MIRS, Slovenia
Olaf Schnelle-Werner, ZMK, Germany
George Rodrigues, Artel, USA
Isabel Chavez Baucom, OWM, USA
Georgia Harris (Retired NIST employee; work on this project was done while at NIST)

March 2024

Contents

1. Introduction

Following the cooperation between EURAMET TC F and ASTM Committee E41 in the revision of the standard E542 - Standard Practice for Calibration of Laboratory Volumetric Apparatus it was decided to investigate the different approaches regarding repeatability measurements of glassware calibration performed by different National Metrology Laboratories, Accredited laboratories and manufactures around the world and its influence to the determined volume and uncertainty.

Suitable repeatability estimates are needed for evaluating measurement results and for determining the measurement uncertainties. Lack of repeatability agreement within a set of measurement results can lead to significant problems with the operating characteristics of the volumetric instrument.

This protocol describes the different approaches in repeatability measurements from each participant and its impact on the final uncertainty calculation and volume determination. At least two common approaches for obtaining repeatability statistics will be investigated: The first approach uses standard deviation control charts to monitor the measurement process at the time of calibration by ensuring that observed and accepted standard deviations agree. The second approach uses a larger number of repeated measurements for each calibration (5 or 10 replicates).

Two sets of flasks of the following 3 nominal volumes were calibrated:100 mL, 500 mL, 1000 mL. Two flasks were broken during the comparison and replaced by others of the same capacity.

The comparison schedule and participants are described in table 1.

Table 1 - schedule and participants

Country	Laboratory	Responsible	Date of measurements
Portugal	IPQ	Elsa Batista	July 2021
Italy	INRIM	Andrea Malengo	September 2021
Slovenia	MIRS	Urška Turnšek	October 2021
Germany	ZMK	Olaf Schnelle-Werner	November 2021
USA	Artel	George Rodrigues	March 2022
USA	NIST OWM*	Georgia Harris	June 2022

*NIST Office of Weights and Measures has responsibility to support legal metrology in the USA. Measurements selected in this analysis for NIST OWM were submitted by Isabel Chavez Baucom.

2. The instruments

The chosen instruments are a) one - mark volumetric flasks (see Fig. 1), nominal capacities: 100 mL (9572, 9573), 500 mL(9576, 9577), 1000 mL (9578, 9579), class A, made out of borosilicate glass, narrow-necked, pear-shaped, manufactured according to ISO 1042:1998.

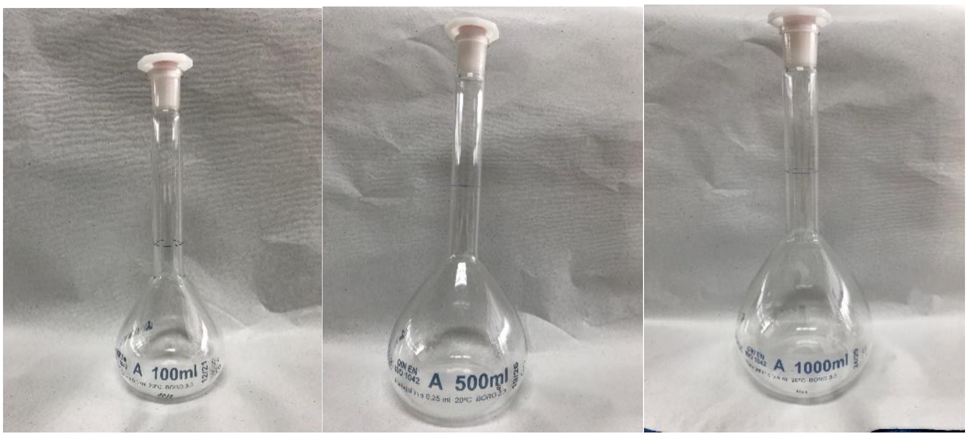

Figure 1 – Flask of 100 mL, 500 mL, and 1000 mL

During the comparison the 500 mL flask 9576 and the 1000 mL flask 9579 were broken and replaced by two other flaks.

3. Calibration method

3.1 Method description

Calibration of the flasks consisted of the determination of the amount of water contained in the flask at reference temperature of 20 ºC, using the gravimetric method. The following equation described in ISO standard 4787 [1] and in ASTM E 542 –01:2002 [2] were used (NIST SOP 14 is considered equivalent, NISTIR 7383 (2019):

$$V_{20} = (I_I - I_E) \times \frac{1}{\rho_W - \rho_A} \times \left(1 - \frac{\rho_A}{\rho_B}\right) \times \left[1 - \gamma(t - 20)\right] \qquad (1)$$

Where:

V_0 volume, at the 20 ºC , in mL

I_I weighing result of the flask full of liquid, in g

I_E weighing result of the empty and dry flask, in g

ρ_W water density, in mg/μL, at the calibration temperature t, in ºC, is advisable to use the Tanaka density formula [3]

ρ_A air density, in g/cm^3

ρ_B density of masses used during measurement (substitution weighing scheme) or during calibration of the balance, in g/cm^3

γ cubic thermal expansion coefficient of the material of the flask, in °C^{-1}

t water temperature used in the calibration, in °C

The participating laboratories used their own test procedure of calibration. The simplified calibration procedure is:

- Weigh the empty dry standard and record the mass, I_E

- Fill the flask up to the reference line, adjust the meniscus and wipe out (drying) any water drops above the reference line.
- Weigh the filled standard recording the mass I_{L}.

3.2 Meniscus reading

- The meniscus shall be set so that the plane of the upper edge of the graduation line is horizontally tangential to the lowest point of the meniscus.
- The shape of the meniscus is set such that the surface of the liquid forms a curve that meets the glass tangentially.
- NIST OWM meniscus readings were performed following NISTIR 7383, GMP 3, Option A (or ASTM E542 7.2.2.2) and introduce volumetric biases. Measurements and calculations performed at NIST during calibration indicate a bias/offset of approximately the following values: 0.028 mL at 100 mL; 0.053 mL at 500 mL; and 0.094 mL at 1000 mL. This bias was corrected according to ASTM E542 in order to obtained comparable results with the other laboratories that use ISO 4787 procedure.

4. Evaluation of the measurement results

4.1 Reference value

To determine the reference value the formula of the weighted mean is used, by means of the inverses of the squares of the associated standard uncertainty are the weighting factors [4]:

$$y = \frac{x_1/u^2(x_1)+\ldots+x_n/u^2(x_n)}{1/u^2(x_1)+\ldots.+1/u^2(x_n)} \qquad (2)$$

To determine the standard uncertainty u(y) associated with y is used the following expression:

$$u(y) = \sqrt{\frac{1}{1/u^2(x_1)+\ldots+1/u^2(x_n)}} \qquad (3)$$

4.2 Consistency determination

To identify an overall consistency of the results a chi-square test can be applied to all *n* calibration results.

$$\chi^2_{obs} = \frac{(x_1-y)^2}{u^2(x_1)} + \ldots + \frac{(x_n-y)^2}{u^2(x_n)} \qquad (4)$$

where the degrees of freedom are: $\nu = n - 1$

The consistency check is regarded as failed if: $\Pr\left\{\chi^2(\nu) > \chi^2_{obs}\right\} < 0{,}05$. The function *CHIINV(0,05; n-1)* in MS Excel was used. The consistency check was failing if *CHIINV(0,05; n-1) <* $\chi^2_{obs.}$

If the consistency check did not fail then *y* was accepted as the KCRV x_{ref} and $U(x_{ref})$ was accepted as the expanded uncertainty of the KCRV.

If the consistency check failed then the laboratory with the highest value of $\frac{(x_i-y)^2}{u^2(x_i)}$ is excluded from the next round of evaluation and the new reference value, reference standard uncertainty and chi-squared value is calculated again without the excluded laboratory.

The En value was also calculated. This value is defined as [5]:

$$E_{n_{lab-1}} = \frac{E_{lab-i} - E_{RV}}{\sqrt{U^2(E_{lab-i}) - U^2(E_{RV})}} \tag{5}$$

where $\varepsilon_{\mathrm{lab-i}}$ is the error of lab-i for a certain point, $\varepsilon_{\mathrm{RV}}$ is the comparison reference value (RV) for the error and $\mathrm{U}(\varepsilon_{\mathrm{lab-i}})$ and $\mathrm{U}(\varepsilon_{\mathrm{RV}})$ and the expanded uncertainties (*k*=2) of those values.

With the value of $\mathrm{E_n}$ one can conclude that:

- The results of the laboratory for a certain point are consistent (passed) if $\mathrm{E_n} \leq 1$

- The results of the laboratory for a certain point are inconsistent (failed) if $\mathrm{E_n} > 1$

5. Measurement results

5.1 Flask stability

The volume measurements obtained by IPQ in the beginning of the comparison (IPQ-1) and in the end of the comparison (IPQ-2) are presented in the following table. For two of the flasks this stability could not be assessed because they were broken.

Table 2 – Volume variation measurement results

		Nominal volume	Volume (μl)	*U_{exp}* (μl)	ΔV (μl)
Flask 9572	IPQ – 1	100	99,955	0,013	0,002
	IPQ - 2	100	99,953	0,012	
Flask 9573	IPQ – 1	100	99,962	0,015	0,002
	IPQ - 2	100	99,960	0,012	
Flask 9577	IPQ – 1	500	499,863	0,053	0,025
	IPQ - 2	500	499,838	0,050	
Flask 9578	IPQ – 1	1000	999,71	0,10	0,05
	IPQ - 2	1000	999,66	0,10	

From the obtained results its can be verify that the flasks were stable, the volume variation is smaller than the uncertainty. Only the first results from IPQ were used to determine the reference value.

5.2. Volume results, 100 mL flask

The obtained results for the 100 mL flasks are the following:

Table 3 - Volume results 100 mL

	Flask 9572		**Flask 9573**	
Participant	**Volume (mL)**	***U* (mL)**	**Volume (mL)**	***U* (mL)**
IPQ	99,955	0,013	99,962	0,015
INRIM	99,942	0,013	99,963	0,013
MIRS	99,955	0,031	99,987	0,023
ZMK	99,928	0,050	99,964	0,050
Artel	99,937	0,031	99,969	0,031
NIST OWM	99,939	0,046	99,948	0,046

5.3. Volume results, 500 mL flask

The obtained results for the 500 mL flasks are the following:

Table 4 - Volume results 500 mL

	Flask 9577		Flask 9576		Flask 1926	
Participant	**Volume (mL)**	***U* (mL)**	**Volume (mL)**	***U* (mL)**	**Volume (mL)**	***U* (mL)**
IPQ	499,863	0,053	499,818	0,057	499,951	0,050
INRIM	499,879	0,024	499,847	0,024	-	-
MIRS	499,918	0,052	499,897	0,051	500,032	0,042
ZMK	499,85	0,23			499,95	0,23
Artel	499,990	0,077			499,95	0,08
NIST OWM	499,867	0,070			499,82	0,07

One of the flasks were broken at MIRS (9576) and replaced by another (1926), IPQ measured this flask at the end of the comparison.

5.4. Volume results,1000 mL flask

The obtained results for the 1000 mL flasks are the following:

Table 5 - Volume results 1000 mL

	Flask 9578		Flask 9579		Flask 1	
Participant	**Volume (mL)**	***U* (mL)**	**Volume (mL)**	***U* (mL)**	**Volume (mL)**	***U* (mL)**
IPQ	999,71	0,10	1000,12	0,11	1000,16	0,10
INRIM	999,77	0,04	1000,165	0,036	-	-
MIRS	999,71	0,12	1000,14	0,14	-	-
ZMK	999,78	0,45	999,91	0,45	-	-
NIST OWM	999,61	0,17	-	-	1000,01	0,17

One of the flasks were broken at OWM (9579) and replaced by another (1), IPQ measured this flask at the end of the comparison. Artel did not perform the measurements at this capacity since it's out of their range.

5.5. Determination of the reference value, flasks 100 mL

All the measurement results, the reference value and its uncertainty are presented in the following figures.

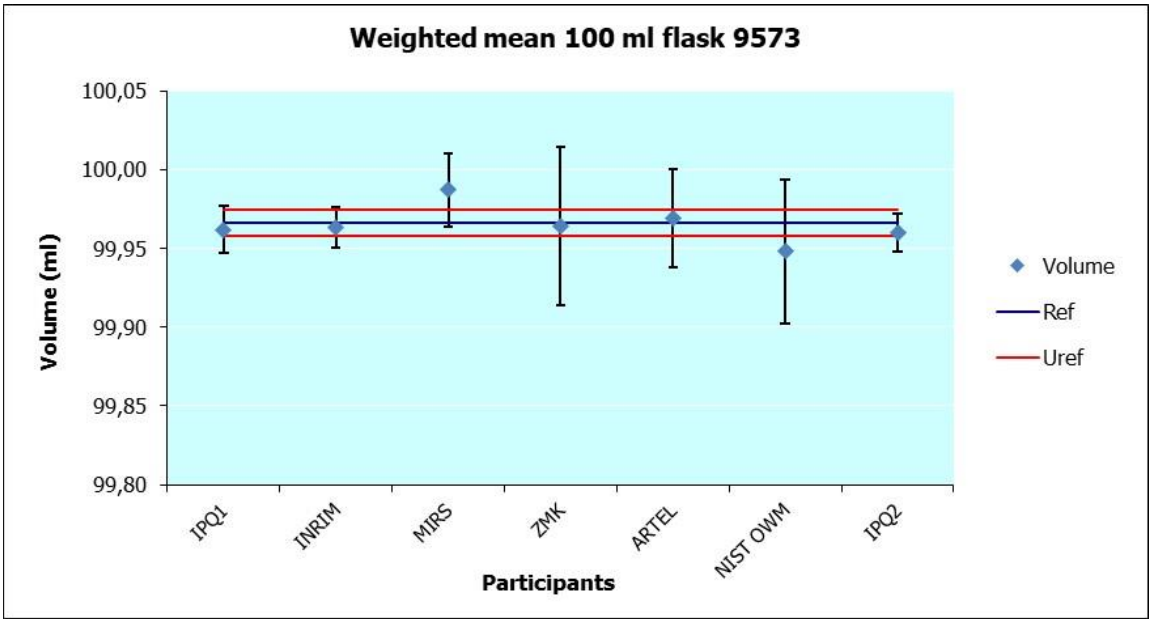

Figure 2 – Volume results with reference value – 100 mL flask 9573

From this figure it can be observed that all the volume results are consistent with the reference value.

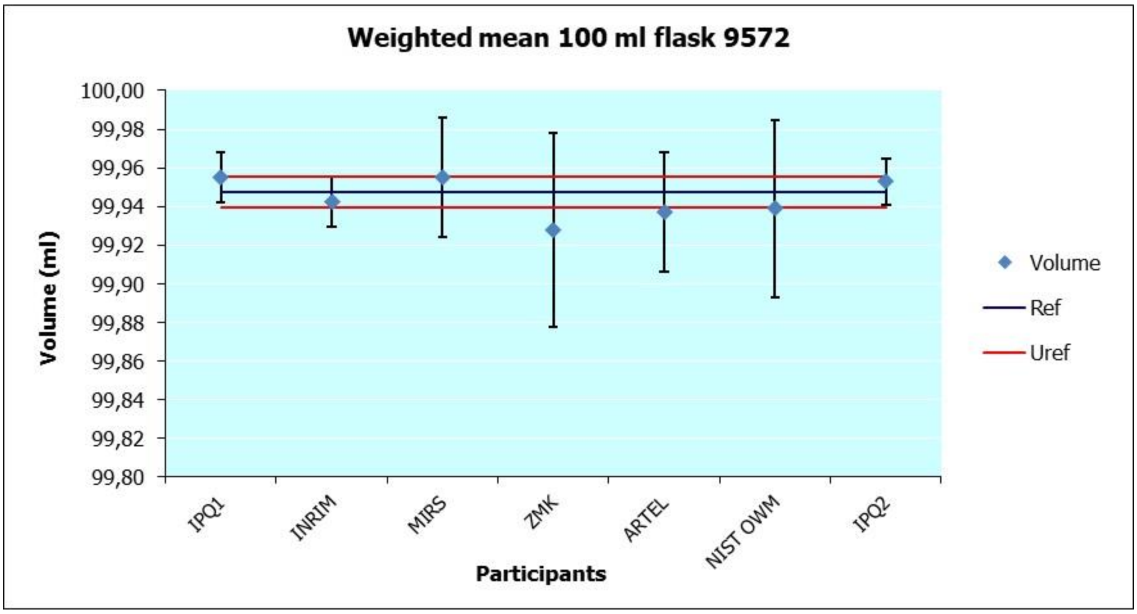

Figure 3 – Volume results with reference value – 100 mL flask 9572

From this figure it can be observed that all the volume results are consistent with the reference value.

The En values were also determined for both flasks and the results are presented in figure 4.

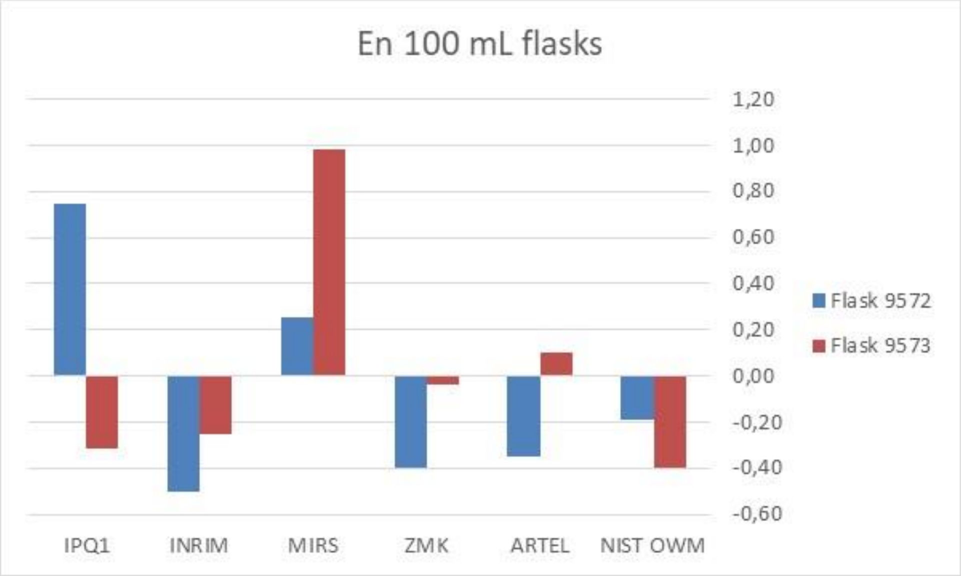

Figure 4 – En value for the 100 mL flasks

From this figure it can be observed that no inconsistent result was found.

5.6. Determination of the reference value, flasks 500 mL

All the measurement results, the reference value and its uncertainty are presented in the following figures.

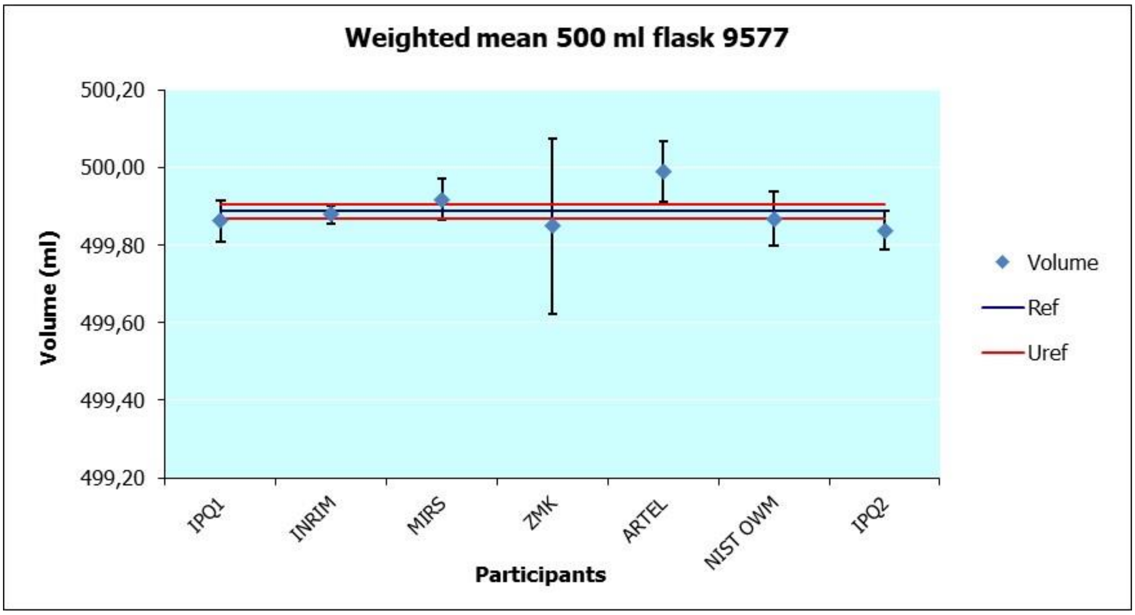

Figure 5 – Volume results with reference value – 500 mL flask 9577

From this figure it can be observed that one results, Artel is inconsistent with the reference value.

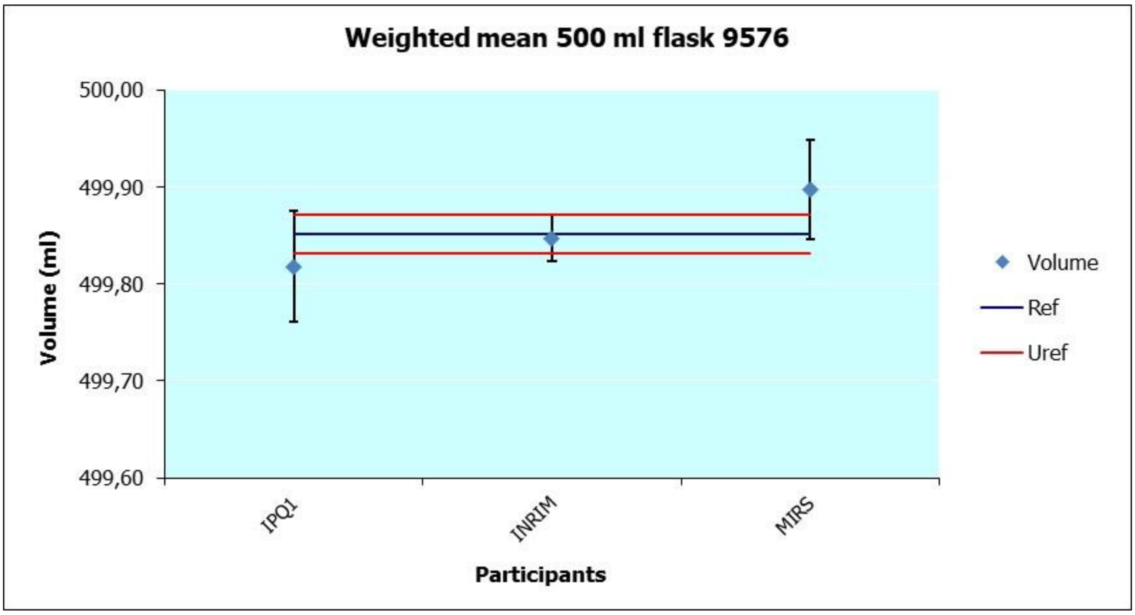

Figure 6 – Volume results with reference value – 500 mL flask 9576

From this figure it can be observed that all the volume results are consistent with the reference value.

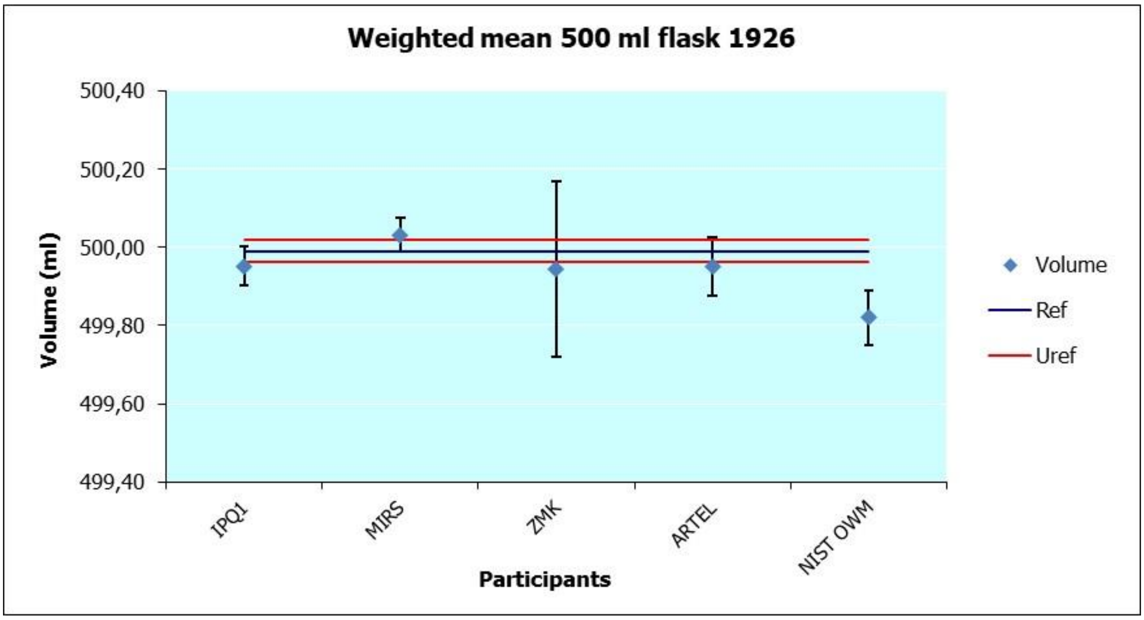

Figure 7 – Volume results with reference value – 500 mL flask 1926

From this figure it can be observed that one results, NIST OWM is inconsistent with the reference value.

The En values were also determined for both flasks and the results are presented in figure 8.

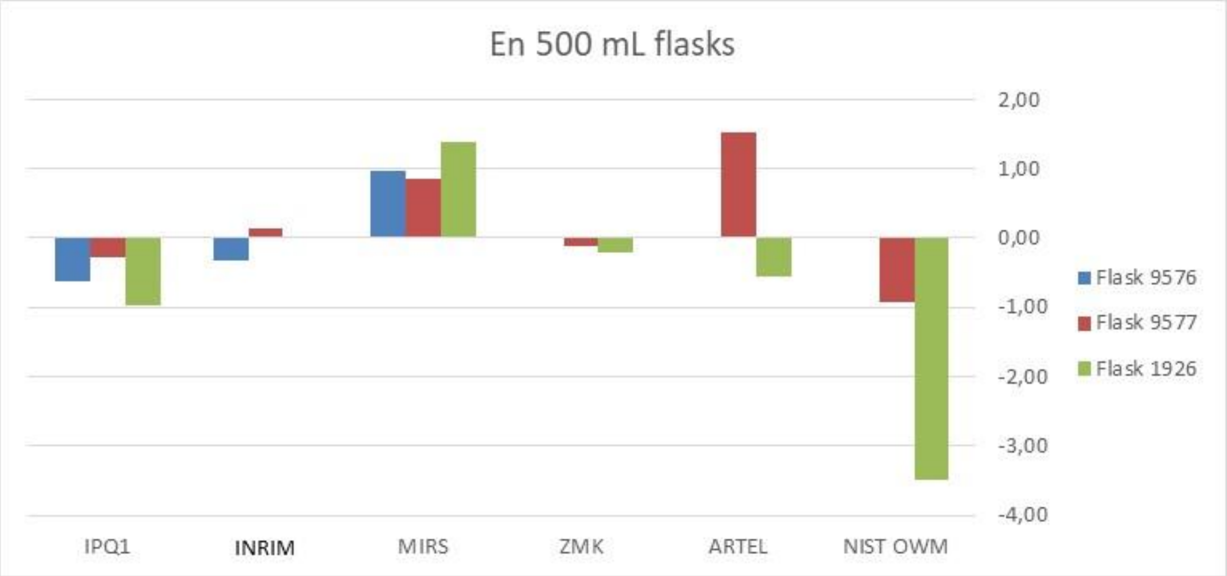

Figure 8 – En value for the 500 mL flasks

For the 3 flasks there were two inconsistent values.

5.7. Determination of the reference value, flasks 1000 mL

All the measurement results, the reference value and its uncertainty are presented in the following figures.

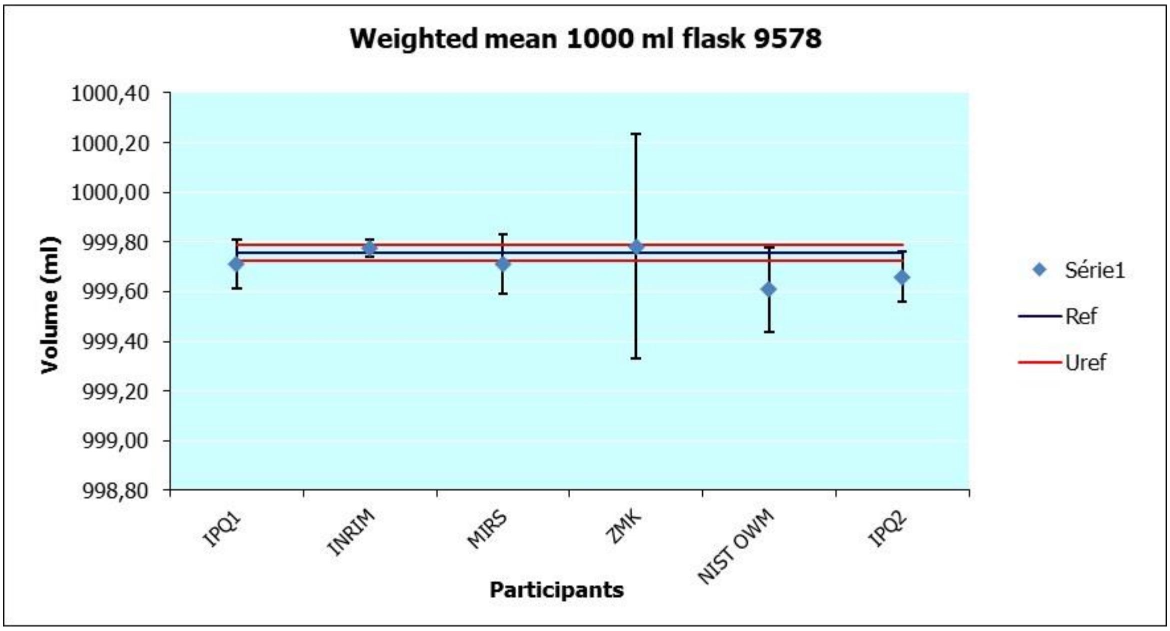

Figure 9 – Volume results with reference value – 1000 mL flask 9578

From this figure it can be observed that there is no inconsistent result.

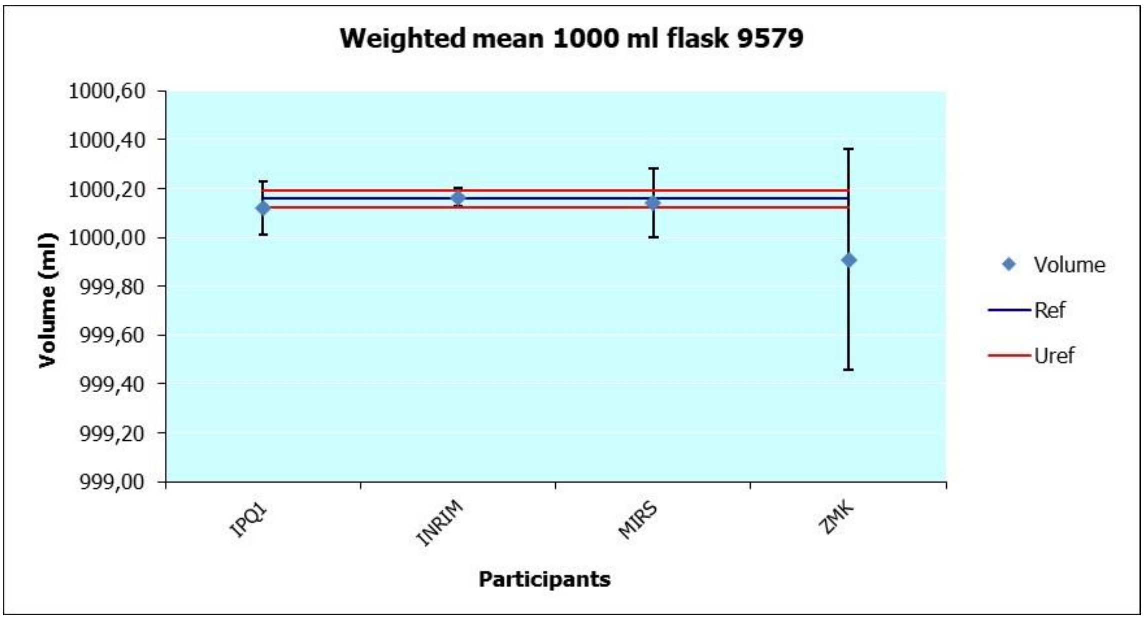

Figure 10 – Volume results with reference value – 1000 mL flask 9579

From this figure it can be observed that all the volume results are consistent with the reference value.

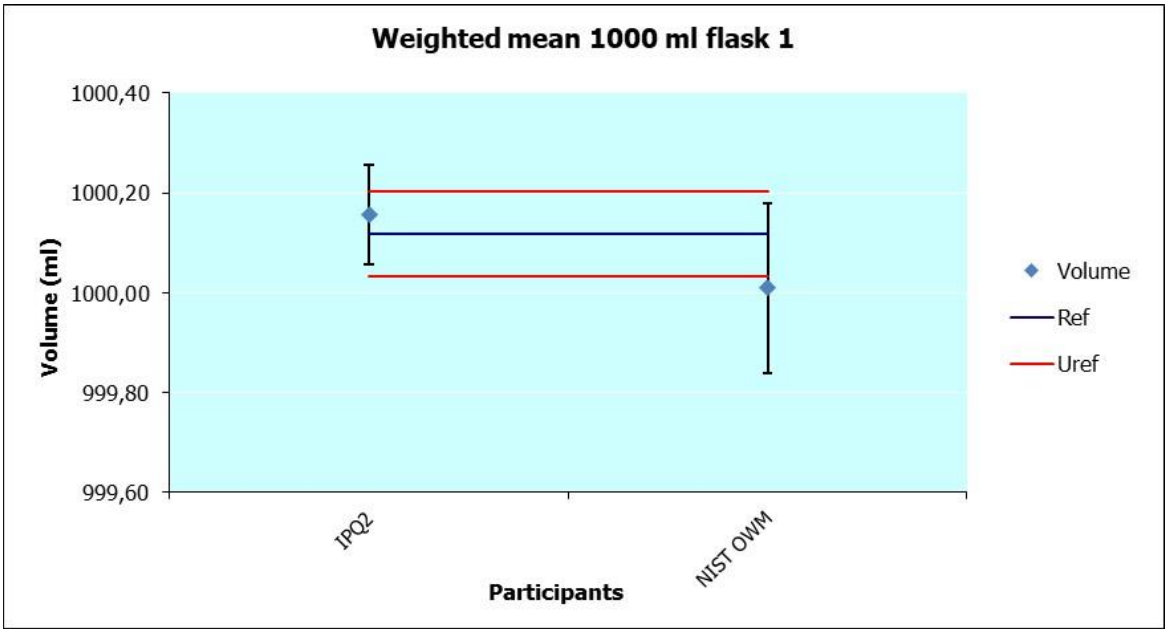

Figure 11 – Volume results with reference value – 1000 mL flask 1

From this figure it can be observed that the results are consistent with the reference value.

The En values were also determined for both flasks and the results are presented in figure 12.

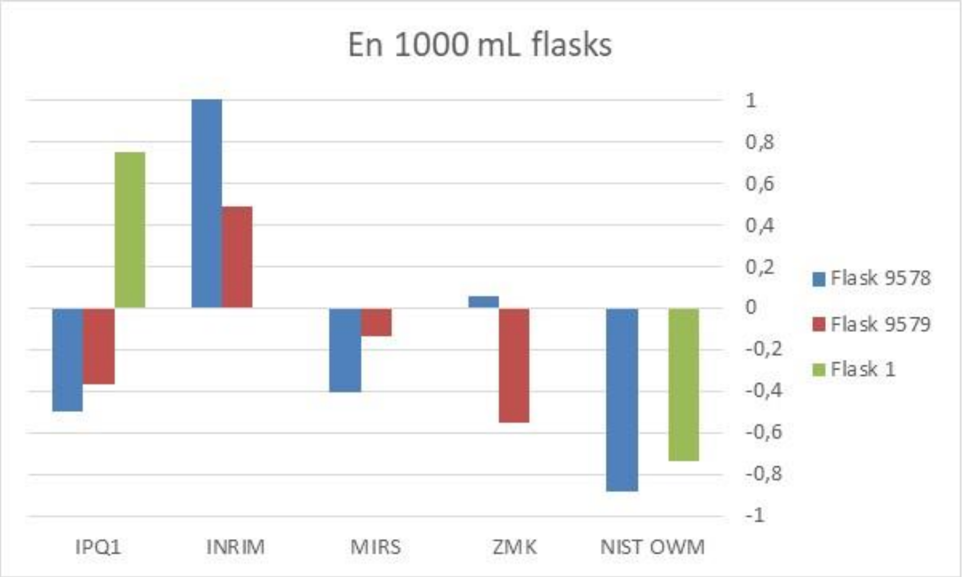

Figure 12 – En value for the 1000 mL flasks

For the 3 1000 mL flasks measured it can be observed that there were no inconsistent values.

6. Reproducibility tests

There were two sets of each flask that were circulated by the participants, this allows to verify if there was any variation within the laboratories.
In relative terms, the errors of the measurements *Di/V* obtained by the laboratories are shown in figure 15, where the error bars correspond to the relative expanded uncertainties.

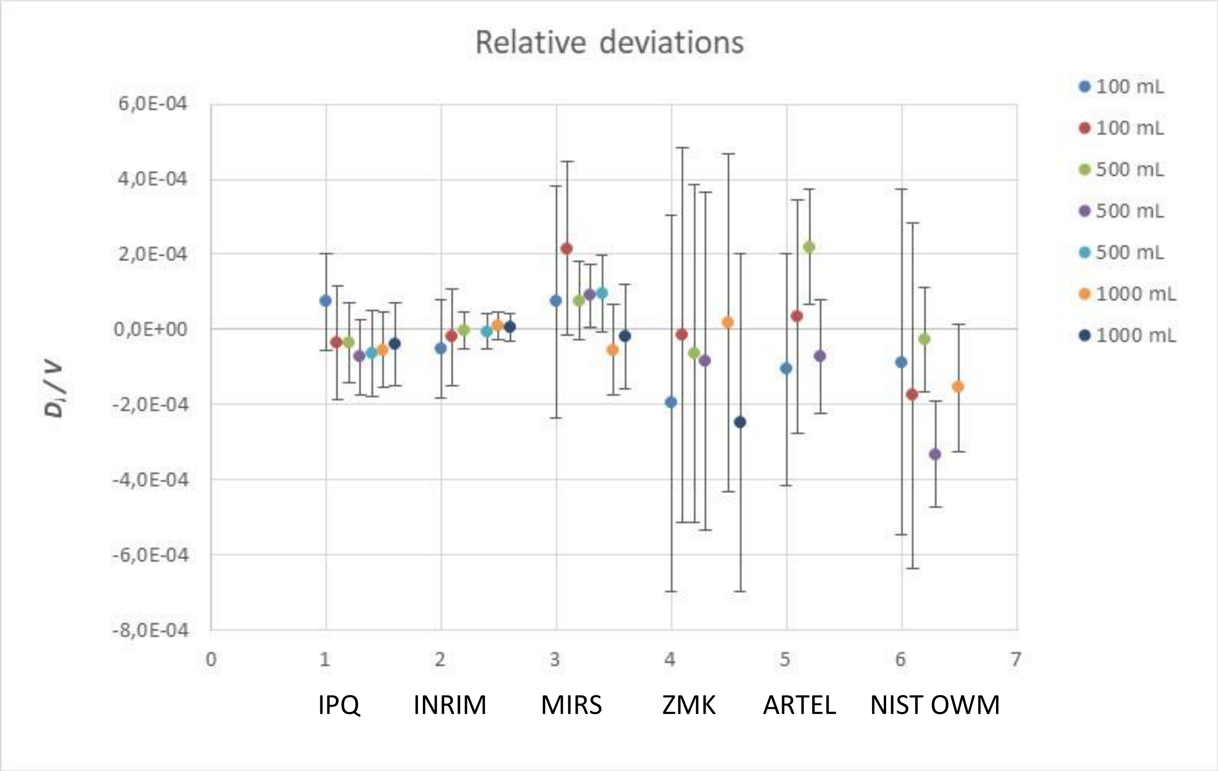

Figure 13 – Relative deviation of the volume flasks

From the figure above it can be seen that the results of each laboratory have similar relative uncertainty values and same relative error value when compared with the reference, with a strong correlation. Assessing the standard deviation of the relative errors, the reproducibility of the measurements for each laboratory was evaluated, which, as expected, is proportional to the declared uncertainties, always remaining below about 0,8 times the expanded uncertainty. The analysis is shown in table 6 and figure 14, where the relative standard deviation is plotted against the mean value of the expanded relative uncertainties. This shows that the stated uncertainties are adequately evaluated, it is noted that the ZMK laboratory could improve the stated uncertainty.

Table 6 – Uncertainty analysis of the volume variations

Participant	**Mean *U* (%)**	**dev std (%)**
IPQ	1,0E-04	5,1E-05
INRIM	6,1E-05	2,0E-05
MIRS	1,4E-04	8,9E-05
ZMK	4,0E-04	1,0E-04
ARTEL	1,9E-04	1,5E-04
NIST OWM	1E-04	2,3E-04

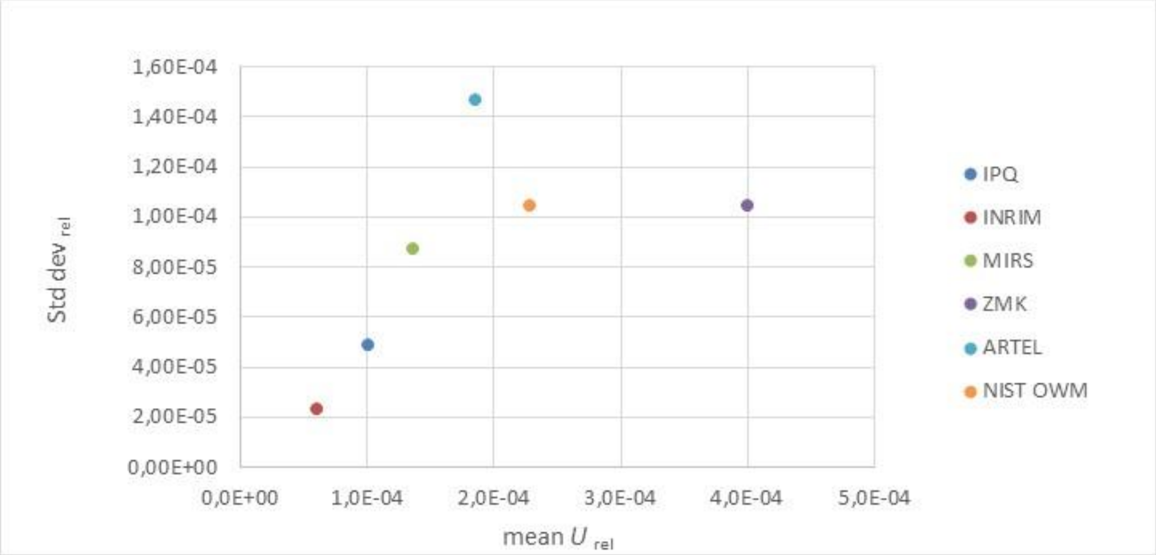

Figure 14 – Uncertainty analysis of the volume variations

7. Uncertainty calculation

Each laboratory described the uncertainty components. Both values, i.e. standard uncertainty and expanded uncertainty were stated, along with the relevant coverage factor *k*. The repeatability measurement procedure was described in detail; documented procedures were referenced (EURAMET cg-19 and NIST SOP 14 provide details).
For the evaluation of the measurement uncertainty, reference was made to the Guide to the Expression of Uncertainty in Measurement at approximately 95 % confidence interval. [6].
In general, the laboratories that used EURAMET cg 19 [7] had lower uncertainty that the ones using the NIST SOP 14.
The largest source of uncertainty was the repeatability and meniscus for the European countries and the meniscus reading and control limits for the USA laboratories.
For the repeatability measurements all laboratories used 10 replicated except for OWM NIST that used 5 repetitions.

8. Conclusions

This project had the purpose to investigate the different approaches regarding repeatability measurements of glassware calibration performed by different National Metrology Laboratories, Accredited laboratories and manufactures around the world and its influence on the determined volume and uncertainty. Six laboratories have participated and two sets of flasks of the following 3 nominal volumes were calibrated: 100 mL, 500 mL, 1000 mL. From 38 results only 2 results were inconsistent.

OWM NIST has used the procedure described in NIST GMP 3 (Option A: setting bottom of the meniscus between upper and lower portion of graduation line ellipse), which leads to lower volume values than the laboratories that use the ISO 4787 meniscus adjustment procedure but his bias was corrected according to ASTM E542 in order to obtained comparable results with the other laboratories.

OWM NIST used the repeatability five runs and compared to historical process standard deviations where available with control charts and standard deviation charts with F-test used to assess repeatability for each flask compared to an accepted value.

Also, the use of less repetitions leads to higher uncertainty values therefore it is advisable to use at least 10 repetitions, EURAMET cg 19 and ISO 4787 annex E, a).

Regarding reproducibility results it was verified that the laboratories had good agreement within instruments and with similar uncertainties.

9. References

[1] ASTM E 542:2021 - Standard practice for calibration of laboratory volumetric apparatus;

[2] ISO 4787:2020; Laboratory glassware – Volumetric glassware – Methods for use and testing of capacity;

[3] Tanaka, M., et. al; Recommended table for the density of water between 0 °C and 40 °C based on recent experimental reports, Metrologia, 2001, Vol.38, 301-309.

[4] M.G. Cox, The evaluation of key comparison data, Metrologia, 2002, Vol. 39, 589-595.

[5] ISO 13528 - Statistical methods used in proficiency testing by interlaboratory comparisons

[6] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML; Guide to the expression of uncertainty in measurement (GUM), Geneva, 1995.

[7] EURAMET calibration guide-19 Version 3.0, Guidelines on the determination of uncertainty in gravimetric volume calibration