European Directorate for the Quality of Medicines & HealthCare | Direction européenne de la qualité du médicament & soins de santé

Funded by the European Union and the Council of Europe

EUROPEAN UNION

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Implemented by the Council of Europe

General European OMCL Network (GEON)
QUALITY MANAGEMENT DOCUMENT

PA/PH/OMCL (18) 153 R1 CORR

EVALUATION OF MEASUREMENT UNCERTAINTY
ANNEX 2.5

Full document title and reference	Evaluation of Measurement Uncertainty ANNEX 2: ESTIMATION OF MEASUREMENT UNCERTAINTY USING TOP-DOWN APPROACH Annex 2.5: Use of Data from Proficiency Testing Studies for the Estimation of Measurement Uncertainty *PA/PH/OMCL (18) 153 R1 CORR*
Document type	Guideline
Legislative basis	Council Directive 2001/83/EC and 2001/82/EC, as amended
Date of first adoption	19 July 2019
Date of original entry into force	1 February 2020
Date of entry into force of revised document	
Previous titles/other references / last valid version	
Custodian Organisation	The present document was elaborated by the OMCL Network / EDQM of the Council of Europe
Concerned Network	GEON

Annex 2 to Guideline "Evaluation of Measurement Uncertainty"
PA/PH/OMCL (18) 145 (in its current version)

Estimation of measurement uncertainty using Top-down approach

Annex 2.5 Use of data from Proficiency Testing Studies for the estimation of measurement uncertainty

Proficiency testing (PT), as defined in ISO/IEC 17025, is the evaluation of participant performance against pre-established criteria by means of inter-laboratory comparisons (by definition, organization, performance and evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions, ISO/IEC 17043).

Therefore, as defined by Eurachem, a PT scheme (PTS) is a system for objectively evaluating a laboratory's performance, helping the participant to assess the accuracy of its measurements [1].

This annex provides examples for use of PTS data for estimation of uncertainty of measurement in routine testing, assuming a normal distribution of the results of a PTS, and therefore, standard deviations (*S*) are applied. Standard deviations (*S*) are used to reflect the expected variability between laboratory results, as Target Standard Deviation (*TS*, or TSD in the EDQM PTS reports) as well as the observed variability between measurements (*S* of repeatability).

By principle, a laboratory participating in a PTS has two main options:

1) *it can rely on its own data (Example 1)*

The uncertainty of measurement can be evaluated using the repeatability of the test results (precision component) and the z-score obtained (bias component). These two components can be obtained from the last participation on a PT round* or from the combination of data from several participations in PT rounds, as long as the performance of the laboratory remains consistent.

2) *it will rely on the data from all the participating laboratories (Example 2)*

The uncertainty of measurement can be evaluated using the standard deviation of results within and between participating laboratories, which means that reproducibility (precision and bias related components, respectively) is assessed. These two components can be obtained from the last participation on a PT round* or from the combination of data from several participations in PT rounds, as long as the performance of the laboratory remains consistent. This approach requires a history of more than 6 participations [2,3].

*Note: The laboratory should evaluate whether the measurements performed for the PT cover major uncertainty contributor and are sufficient to estimate the precision component. In case the major uncertainty contributor are not covered, the in-house method validation data should be used for the estimation of precision.

It should be kept in mind that the uncertainty of measurement is a combination of random and systematic errors. Random errors affect the precision and systematic errors, or bias, affect the trueness [4]. Relying exclusively on internal precision data is not recommended *(Example 1)*.

In a PTS, random errors are reflected by within-lab variability ($S_{within-lab}$) i.e. precision, which means that either the standard deviation (S) of the laboratory (repeatability) or the pooled standard deviation (S_{pool}) of all participant laboratories can be used, if homogeneity is demonstrated (e.g. using Cochran's test). Systematic errors are reflected by z-scores (bias), which means that the z-score obtained by the laboratory or z-scores pooled across participations from the same organization can be used (especially to assess the performance of a method across one organization). Therefore, the bias can be expressed as z-scores or as a variance component (between-lab variance component). The between-lab variance component and within-lab variance component (S_{pool}) are then combined to estimate the reproducibility of the method. There is a preference for the use of S obtained for all the participating laboratories for a measurement of systematic errors, as the mean z-score is 0 or close to 0 (assuming the assigned value is the overall robust mean of participant's results). Therefore, systematic errors can be represented by a difference of means or a standard deviation.

Regarding systematic errors, the significance of the bias is usually assessed (*e.g.* by a t-test) to know if it should be added in the calculation. Note that the t-test is used in place of the z-test when the S is estimated from the data (such as the examples in other annexes). In PTS rounds, the Target S, or TS, is known, and so a z-value (z-score) is calculated in place of the t-value. The critical value is 2 (precisely 1.96 for P = 95% level of confidence). It simply means that if the performance is 'satisfactory', i.e. a z-score ≤ 2, then the bias can be neglected from the combined standard uncertainty. Nevertheless, it is always preferable to take into account the bias component in the estimation of uncertainty, regardless of the z-score achieved by the lab.

The relative disadvantage of using PT samples is the lack of traceable reference values similar to those for certified reference materials (CRM). Consensus values in particular are prone to occasional error. This certainly demands due care in their use for uncertainty estimation. When the uncertainty around the consensus value is significant, then it should be taken into account in the uncertainty budget [1].

Before proceeding to the actual examples, it is important to clarify how to use correctly the data from PTSs.

Random and Systematic errors: within-lab variability, $S_{within-lab}$ (within-laboratory precision) and between-lab variability, $S_{between-lab}$ (bias)

Five different laboratories participated in a PTS and, therefore, there are five estimates of the within-lab variation for the test. Each laboratory performed three independent determinations (number of replicates, n = 3), which means that the number of degrees of freedom for each laboratory was 2 (Table 1).

Table 1. Data reported for a PTS with five participant laboratories

Lab ID	Replicate 1 (mg)	Replicate 2 (mg)	Replicate 3 (mg)	$S_{within\text{-}lab}$ (mg)	Degrees of freedom, DF
1	986	941	975	23.5	2
2	765	791	780	13.1	2
3	958	987	970	14.6	2
4	913	917	945	17.4	2
5	883	857	894	19.0	2

$S_{within\text{-}lab}$ and $S_{between\text{-}lab}$ can be calculated using simple statistical formulae if the degrees of freedom are the same for all laboratories. Otherwise, the use of a statistical software is recommended.

Precision: $S_{within\text{-}lab}$ and S_{pool}

If the results are homogeneous (*e.g.* using the Cochran test), they can be combined, obtaining S_{pool} which is always a weighted mean of the individual $S_{within\text{-}lab}$. The weights are the degrees of freedom (DF = n - 1) used to calculate the various $S_{within\text{-}lab}$.

When n is the same per lab, the weighted mean is equal to the simple mean (arithmetic mean). Regardless of n, it is considered to be a good practice to determine the weighed mean of the *S*, S_{pool}:

$$S_{pool} = \sqrt{\frac{\sum_{i=1}^{j} DF_i \times S^2_{within-lab(i)}}{\sum DF_i}}$$

$$S_{pool} = \sqrt{\frac{2 \times 23.5^2 + 2 \times 13.1^2 + 2 \times 14.6^2 + 2 \times 17.4^2 + 2 \times 19.0^2}{(2+2+2+2+2)}} = 17.88 \text{ mg}$$

As a reminder, to combine variabilities you must always sum variances (S^2), never standard deviations (*S*).

Bias: S_{bias} ($S_{between\text{-}lab}$) and S_R

In the case where the number of replicates is the same per each lab, S_{bias} ($S_{between\text{-}lab}$) can be deduced from the variation between means of the participants using simple statistical formulae. Otherwise, the use of a statistical software is recommended.

First, the overall mean is calculated as a weighted mean of the various laboratory means, where the weights are the number of replicates (in the present example, n = 3). When n is the same per lab, the weighted mean is equal to the simple mean (arithmetic mean).

$$Overall\ Mean = \frac{\sum_{i=1}^{j} n_i \times Mean_i}{\sum n_i}$$

$$Overall\ Mean = \frac{3\times967+3\times779+3\times972+3\times925+3\times878}{3+3+3+3+3}$$

In addition, the standard deviation between the means of the participants is an estimate of standard deviation of reproducibility (S_R) for the given number of replicates (n) (Table 2).

Table 2. Data reported for a PTS with five participant laboratories

Lab ID	n	Mean (mg)
1	3	967
2	3	779
3	3	972
4	3	925
5	3	878
Overall Mean		904
S_R		79.71

The standard deviation of reproducibility (S_R) is a combination of $S_{within\text{-}lab}$ or S_{pool} (repeatability) and $S_{between\text{-}lab}$ (bias).

$$S_R = \sqrt{S_{between-lab}^2 + \frac{S_{pool}^2}{n}}$$

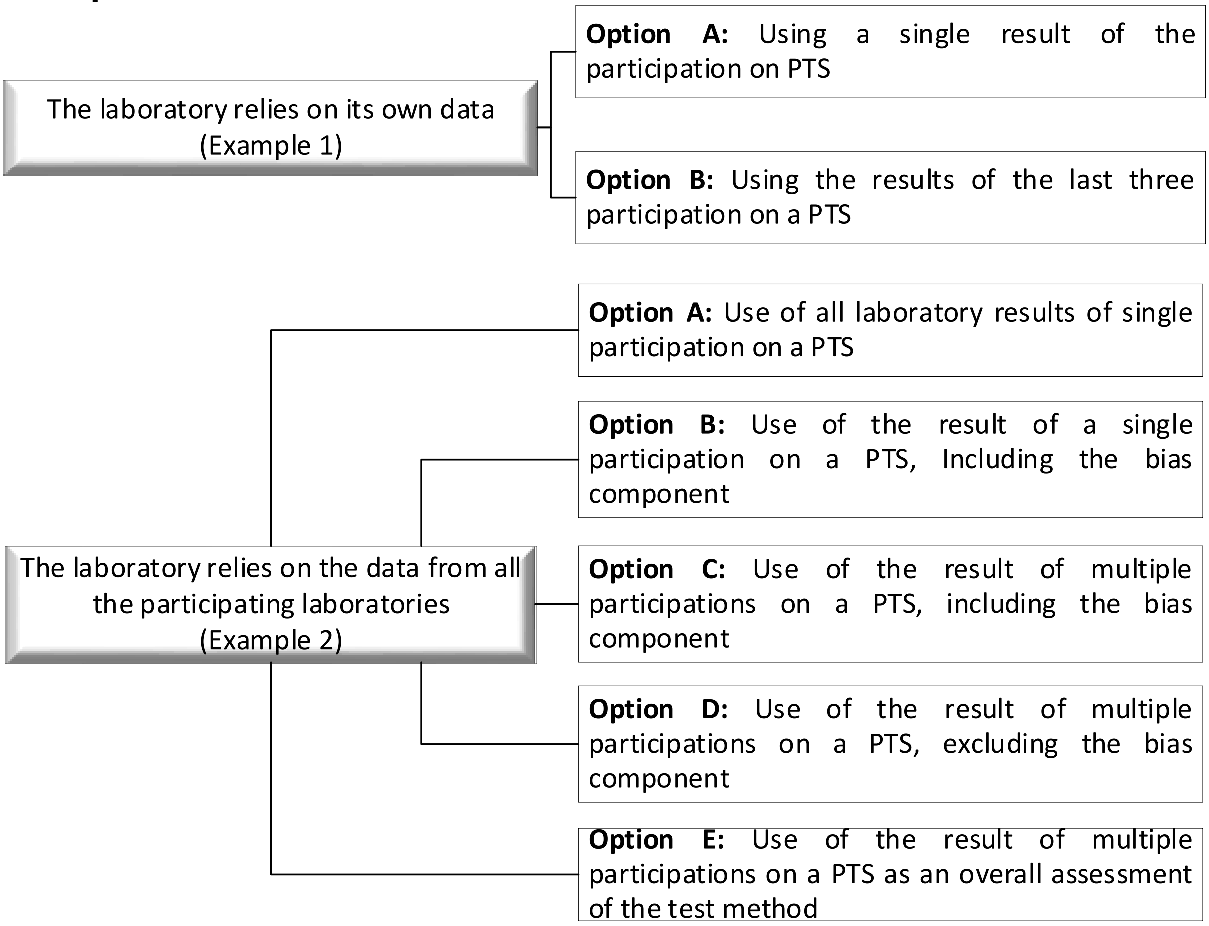
The between-lab variation is then obtained:

$$S_{between-lab} = \sqrt{S_R^2 - \frac{S_{pool}^2}{n}}$$

$$S_{between-lab} = \sqrt{79.71^2 - \frac{17.88^2}{3}} = 79.04 \text{ mg}$$

The S_R combines both precision (random errors) and bias (systematic errors in terms of uncertainty of measurement, or aggregation of z-scores in PTS terms).

Examples


The laboratory relies on its own data (Example 1)
Option A: Using a single result of the participation on PTS
Option B: Using the results of the last three participation on a PTS
The laboratory relies on the data from all the participating laboratories (Example 2)
Option A: Use of all laboratory results of single participation on a PTS
Option B: Use of the result of a single participation on a PTS, Including the bias component
Option C: Use of the result of multiple participations on a PTS, including the bias component
Option D: Use of the result of multiple participations on a PTS, excluding the bias component
Option E: Use of the result of multiple participations on a PTS as an overall assessment of the test method

Example 1: Use of data from PTS for the estimation of measurement uncertainty for quantitative determination of content of a test sample relying on its own data

1. Description of the analytical procedure

The melting point is determined according to Ph. Eur. 2.2.14.
The laboratory participated in three round of PTS for melting point (ºC), according to Ph. Eur. 2.2.14, and the observed data are provided in Table 3.

Table 3. Data obtained for several rounds of the same PTS on melting point.

Round ID	Lab. Mean Result	$S_{within\text{-}lab}$	Assigned Value (°C)	$U_{assigned\ value}$ (k = 2*)	S_R	Number of participant labs	Z-score (bias)
1 (n = 3)	115.5	0.26	115.1	0.1	1.09	45	0.4
2 (n = 3)	160.0	0.31	160.0	0.2	1.34	51	0.0
3 (n = 3)	229.0	0.15	228.8	0.2	1.36	44	0.2

n = number of replicates
* for approximately 95% level of confidence

2. Estimation of measurement uncertainty

2.1 Specification of a measurand

The measurand is melting point expressed as ºC. The target standard deviation (TS) is 1ºC.

2.2 Quantification of the uncertainty of measurement

Option A: Using a single result of the participation on PTS

The laboratory chooses to rely on the result of its participation on Round 1, for which the z-score obtained was 0.4.

By definition, the uncertainty will be estimated combining precision and bias:

$$u_c = \sqrt{u^2_{precision} + u^2_{bias}}$$

However, due to the satisfactory result obtained in Round 1 (z-score ≤ 2), the bias is not significant and, therefore, could be disregarded (not recommended). The estimation of uncertainty will rely solely on the within-lab precision, $S_{within\text{-}lab}$ = 0.26 °C. In this case, the uncertainty around the bias has not been taken into consideration.

Therefore: $$u_c = \sqrt{s^2_{within-lab}} = s_{within-lab} = 0.26\ ℃$$

The expanded uncertainty U is:

$$U = k \times u_c$$

For Round 1 (Table 1), and assuming k = 2 and approximately 95% level of confidence:

$$U = 2 \times 0.26 = 0.52℃.$$

or, expressed as a relative expanded uncertainty:

$$U_{rel} = \left(100 \times \frac{0.26}{115.5}\right) = 0.46\%.$$

Alternatively, the laboratory can use S_{pool} calculated from all the participants (see above *Precision:* $S_{within\text{-}lab}$ *and* S_{pool}).

Option B: Using the results of the last three participations on a PTS

The laboratory chooses to rely on the results of the last three participations described in Table 3.

By definition, the uncertainty will be estimated combining precision and bias:

$$u_c = \sqrt{u^2{}_{precision} + u^2{}_{bias}}$$

However, due to the satisfactory results obtained for the three rounds (z-score ≤ 2), the bias is not significant and, therefore, could be disregarded (not recommended). The estimation of uncertainty will rely solely on the within-lab precision. If the within-lab precision is similar on the several rounds, the pooled S can be used (see above *Precision:* $S_{within\text{-}lab}$ *and* S_{pool}).

In the present case, the results were obtained using triplicates (Table 3).

Table 4. Estimation of S_{pool} from the data described on Table 3

Round ID	Lab. Mean Result (ºC)	$S_{within\text{-}lab}$	n	DF	S_{pool}
1	115.5	0.26	3	2	$\sqrt{\frac{2\times0.26^2+2\times0.31^2+2\times0.15^2}{(2+2+2)}}$ = 0.25 °C
2	160.0	0.31	3	2	
3	229.0	0.15	3	2	

Therefore, for k = 2 and approximately 95% level of confidence, using S_{pool}:

$$u_c = S_{pool}$$
$$U = k \times u_c$$

$$U = 2 \times 0.25 = 0.50\ ℃.$$

2.3 Reporting of results

For a melting point of 115.1 ºC, the result should be reported as:

Option A: 115.1 ± 0.5 ºC, (k = 2, for approximately 95% level of confidence).

Option B: 115.1 ± 0.5 ºC, (k = 2, for approximately 95% level of confidence).

For convenience, the results were reported with one decimal place, and U with one significant figure.

Note: when treating data from different rounds of a PTS, for which the consensus value and TS are also different (TSs vary with different consensus values), it is advised, instead of using S_{pool}, to use RSD_{pool}. In this sense, the U must be converted from percentage to ºC, which will vary with the target value. Using S_{pool} directly, the U (in ºC) obtained applies to the entire working range tested.

Example 2. The laboratory relies on the data from all the participating laboratories

In this group of examples, it is important to refer *Bias:* S_{bias} ($S_{between\text{-}lab}$) *and_*S_R: an estimate of reproducibility is obtained by S_R, which is a combination of the $S_{within\text{-}lab}$ (repeatability) and $S_{between\text{-}lab}$ (bias).

1. Description of the analytical procedure

The density of liquid sample is determined according to Ph. Eur. 2.2.25.
The laboratory participates in a PTS for density, expressed in g/cm^3. For informative purposes, Table 5 reports on the results obtained in 10 runs of the PTS for density. The condition to use the last result (round n. 10) is to have a history of minimum of 6 participations. The values reported for S_R include $S_{within\text{-}lab}$ and $S_{between\text{-}lab}$. According to the Eurachem Citac Guide CG4 and Eurolab Technical Report 1/2007, it is possible for the participant laboratory to consider that the uncertainty of measurement is obtained directly from the within-laboratory data [2,3].

Table 5. Results obtained for the participation in 10 rounds of a PTS on density

Round ID	Lab. Mean Result (g/cm^3)	$S_{within\text{-}lab}$	Assigned Value (g/cm^3)	$U_{assigned\ value}$ (k = 2*)	S_R	Number of participant labs	z-score (bias)
1 (n=3)	1.119	0.0015	1.119	0.003	0.0024	61	0.0
2 (n=3)	0.845	0.0021	0.846	0.003	0.0019	57	-0.5
3 (n=3)	0.862	0.0019	0.865	0.003	0.0022	47	-1.5
4 (n=3)	1.261	0.0028	1.261	0.003	0.0031	44	0.0
5 (n=3)	0.958	0.0022	0.96	0.004	0.0020	42	-1.0
6 (n=3)	0.911	0.0016	0.912	0.004	0.0027	57	-0.5
7 (n=3)	0.842	0.0018	0.845	0.004	0.0021	44	-1.5
8 (n=3)	0.914	0.0025	0.918	0.004	0.0031	50	-1.8
9 (n=3)	0.911	0.0015	0.911	0.004	0.0018	43	0.0
10 (n=3)	0.863	0.0020	0.864	0.003	0.0023	57	-0.5

n = number of replicates
* for approximately 95% level of confidence

2. Estimation of measurement uncertainty

2.1 Specification of a measurand

The measurand is density of a liquid expressed as g/cm^3. The target standard deviation (TS) is 0.002 g/cm^3.

According to Ph. Eur 2.2.25, Density, ρ_{20}, is defined as the mass of a unit volume of the substance at 20 °C, expressed in kilograms per cubic metre or grams per cubic centimetre (1 kg/m^3 = 10^{-3} g/cm^3):

$$\rho_{20} = \frac{mass_{substance}}{unit\ volume_{substance}}$$

2.2 Quantification of the uncertainty of measurement

Option A: Use of all laboratory results of single participation on a PTS

The reproducibility standard deviation, S_R, can be a suitable estimate for the measurement uncertainty, as it already includes the contribution from repeatability and from bias. Since this already comprises systematic effects due to different ways of operation in the laboratories involved in the PTS, an additional uncertainty contribution accounting for systematic effects is normally not necessary [6].

Therefore, the combined standard uncertainty $u(y)$, is given by the between-laboratory reproducibility standard deviation, S_R, where questionable and unsatisfactory results are excluded (i.e. |z scores|> 2).

$$u_c = s_R$$

$$U = k \times u_c$$

From Table 5, the participation of the laboratory on the more recent round (Round 10) of a PTS for density (one round), with n = 57 participants, all with satisfactory results.

This is applicable when the number of replicates is 3. In case of different number of replicates (*n*), the following formula has to be applied:

$$S_R = \sqrt{S^2_{between-lab} + \frac{S^2_{pool}}{n}}$$

This means that, for a higher number of replicates, such as n = 3, the influence of S^2_{pool} is smaller, so S_R obtained will be smaller, hence U will also be lower than the one obtained for n = 1. Therefore, if the laboratory is using in routine testing one replicate (n = 1), and the participation in the PTS was performed with three replicates (n = 3), U obtained from the PTS data is underestimated for the foreseen routine use.

Calculated results for uncertainty of measurement based on inter-laboratory precision (reproducibility) are given in Table 6 (for n = 3).

Table 6. Calculated uncertainty of measurement based on reproducibility

Assigned Value	0.864 g/cm^3
Result of the participating laboratory	0.863 g/cm^3
Number of replicates for the PTS	3
Within-laboratory precision – repeatability, $S_{\text{within-laboratory}}$ (S_r)	0.0020 g/cm^3
Reproducibility (all participants), S_R	0.0023 g/cm^3
Between-laboratory precision $S_{between-laboratory}(S_g)$	$S_{between-laboratory}(S_g) = \sqrt{S_R^2 - \frac{S^2_{within-laboratory}}{n_{replicatesPTS}}} =$

	$\sqrt{0.0023^2 - \frac{0.0020^2}{3}} = 0.0020\ g/cm^3$	
Number of independent sample preparations (n) in routine testing	**3**	**1**
Combined standard uncertainty, u_c	$u_c = S_R$, *which explained by* $u_c = \sqrt{S^2_{between-laboratory} + \frac{S^2_{within-laboratory}}{n_{routine\ replicates}}} = \sqrt{0.0020^2 + \frac{0.0020^2}{3}} = 0.0023$	$u_c = \sqrt{S^2_{between-laboratory} + \frac{S^2_{within-laboratory}}{n_{routine\ replicates}}} = \sqrt{0.0020^2 + \frac{0.0020^2}{1}} = 0.0028$
Expanded uncertainty (U) , *U = 2 x* u_c	0.0046 ≅ 0.005 g/cm^3	0.0056 ≅ 0.006 g/cm^3

Coverage factor, k = 2, level of confidence: approximately 95%, U reported with one significant figure.

Option B: Use of the result of single participation on a PTS, including the bias component

From Table 5, it is selected the participation of the laboratory on Round 10 of a PTS for density (one round), with n = 57 participants, all with satisfactory results.

If the PTS report does not state the uncertainty of the assigned value, $u_{C_{ref}}$, it can be determined by:

$$u_{C_{ref}} = \frac{S_R}{\sqrt{n_{total}}}$$

where n_{total} is the number of participating laboratories [7]. However, this equation may require additional factors accounted for if the assigned value is determined by median, which must be checked in advance [3].

It is important to mention that $u_{C_{ref}}$ is deemed negligible if it is lower than 0.3 times the target standard deviation of the PTS round.

Therefore, the standard uncertainty is the combination of precision, $u_{precision}$, and bias, u_{bias}, data:

$$u_c = \sqrt{u^2_{precision} + u^2_{bias}}$$

with S_R being used as $u_{precision}$, and $u_{bias} = \sqrt{bias^2 + \left(\frac{s_{bias}}{\sqrt{n}}\right)^2 + u_{Cref}{}^2}$, taking into account that the laboratory decided to use the results from the last PTS participation and bias has to be included. These assumptions are applicable to the specific case where n = 3 replicates are used for the PTS, as well as in routine testing. As in Option A (table 6), if the laboratory performs one replicate (n = 1) in routine testing, it should be taken into account that:

$$S_R = \sqrt{S^2_{between-laboratory} + \frac{S^2_{within-laboratory}}{n_{routine\ replicates}}}$$

Assuming that n = 3 (for the PTS replicates and for the routine replicates), the bias is calculated as:

$$\Delta = |\text{Value obtained} - \text{True value}|.$$

S_{bias} is determined from independent replicates if only one PTS is used; in this case, three independent sample preparations were used and the bias can be determined using the mean value of these preparations (Table 7):

$$\text{Bias: } \Delta = |\ 0.863\text{-}0.864| = 0.001$$

Note: as the bias will be used as squared value ($Bias^2$), it is not critical to force the determination of the bias as an absolute value.

As alternative to the use of the mean value of the replicates, the results obtained for the calculation of bias are provided in Table 7.

Table 7. Calculation of the bias for Round 10

	Independent sample preparations (N=3) (g/cm^3)	**Bias** **Δ = Value obtained – True value**
	0.863	-0.001
	0.861	-0.003
	0.865	0.001
Average	0.8633 (reported value: 0.863)	-0.001
S	0.002	0.002

For the determination of the bias, a constant value (true value) is subtracted to each replicate, which means that S_{bias} is equal to $S_{repeatability}$. This is supported by the fact that when variances are applied to the equation of the bias, the true value has a variance of zero because a constant has no variation.

Therefore, in this case, and using standard deviations:

$$u_{bias} = \sqrt{0.001^2 + \left(\frac{0.002}{\sqrt{3}}\right)^2 + \left(\frac{0.003}{2}\right)^2} = 0.00214\ g/cm^3$$

which is converted into 0.248% when compared with the reported value, 0.863 g/cm^3.

Combined standard uncertainty u_c, is determined using *S* as:

$$u_c = \sqrt{0.002^2 + 0.00214^2} = 0.00293\ g/cm^3$$

The expanded uncertainty (U), is:

$$U = k \times u(y)$$

for k = 2, for approximately 95% level of confidence, and one significant figure:

$U = 2 \times 0.00293\ g/cm^3 = 0.006\ g/cm^3$, or U_{rel} = 0.7% (when comparing with the reported value, 0.863 g/cm^3).

Option C: Use of the result of multiple participations on a PTS, including the bias component

All the results from the participation of the laboratory on 10 rounds of a PTS for density are selected, all with satisfactory results (Table 5).

Therefore, the standard uncertainty is the combination of precision, $u_{precision}$, and bias, u_{bias}, data:

$$u_c = \sqrt{u^2{}_{precision} + u^2{}_{bias}}$$

with S_R being used as $u_{precision}$, and $u_{bias} = \sqrt{RMS_{bias}^2 + u_{Cref}{}^2}$,

where:

- S_R is the pooled standard deviation of all the individual S_R for 10 rounds described in Table 5,
- RMS is the root mean square of all the bias values from 10 rounds, $RMS_{bias} = \sqrt{\frac{\sum(bias_i)^2}{n_{PTS}}}$
- u_{Cref} is the highest value obtained for the reported uncertainty of the assigned value for all the 10 rounds (Table 5)

Note: these assumptions are applicable to the specific case where n = 3 replicates are used for the PTS, as well as in routine testing. As in Option A (Table 6), if the laboratory performs one replicate (n = 1) in routine testing, it should be taken into account that:

$$S_R = \sqrt{S_{between-laboratory}^2 + \frac{S_{within-laboratory}^2}{n_{routine\ replicates}}}$$

Although all the z-scores were satisfactory, the laboratory chooses to include the bias in the estimation of the combined uncertainty.
The pooled S_R is calculated as follows:

$$S_{pool} = \sqrt{\frac{\sum_{i=1}^{j} DF_i \times S^2{}_{R(i)}}{\sum DF_i}}$$

$$\Leftrightarrow$$

$$S_{pool} = \sqrt{\frac{60 \times 0.0024^2 + 56 \times 0.0019^2 + 46 \times 0.0022^2 + 43 \times 0.0031^2 + 41 \times 0.0020^2 + 56 \times 0.0027^2 + 43 \times 0.0021^2 + 49 \times 0.0031^2 + 42 \times 0.0018^2 + 56 \times 0.0023^2}{(60 + 56 + 46 + 43 + 41 + 56 + 43 + 49 + 42 + 56)}} = 0.01193$$

or

$$10^{-4}\sqrt{\frac{60 \times 24^2 + 56 \times 19^2 + 46 \times 22^2 + 43 \times 31^2 + 41 \times 20^2 + 56 \times 27^2 + 43 \times 21^2 + 49 \times 31^2 + 42 \times 18^2 + 56 \times 23^2}{(60 + 56 + 46 + 43 + 41 + 56 + 43 + 49 + 42 + 56)}}$$

The bias is calculated as: Δ = Value obtained – True value or Δ = |Value obtained – True value|, and S_{bias} is determined from the individual participations in each round of the PTS. However, as the bias can be calculated without obtaining the absolute values, enabling to check tendencies (positive or negative), this is how it was calculated in Table 8.

The obtained results from calculation of bias are provided in Table 8.

Table 8. Calculation of the bias for all rounds

Round ID	Lab. Mean Result (g/cm^3)	Assigned Value (g/cm^3)	Bias Δ=Value obtained – True value
1 (n=3)	1.119	1.119	0.000
2 (n=3)	0.845	0.846	-0.001
3 (n=3)	0.862	0.865	-0.003
4 (n=3)	1.261	1.261	0.000
5 (n=3)	0.958	0.96	-0.002
6 (n=3)	0.911	0.912	-0.001
7 (n=3)	0.842	0.845	-0.003
8 (n=3)	0.914	0.918	-0.004
9 (n=3)	0.911	0.911	0.000
10(n=3)	0.863	0.864	-0.001
		Average (g/cm^3)	-0.002

Therefore,

$$RMS_{bias} = \sqrt{\frac{\sum(bias_i)^2}{n_{PTS}}} = \sqrt{\frac{0.000^2 + 0.001^2 + 0.003^2 + 0.000^2 + 0.002^2 + 0.001^2 + 0.003^2 + 0.004^2 + 0.000^2 + 0.001^2}{10}}$$

$$= 0.00000410\ g/\ cm^3$$

and

$$u_{bias} = \sqrt{RMS_{bias}^2 + u_{Cref}{}^2} = \sqrt{0.00000410^2 + 0.004^2} = 0.004\ g/cm^3$$

Combined standard uncertainty u_c, is determined using S as:

$$u_c = \sqrt{0.01193^2 + \left(\frac{0.004}{2}\right)^2} = 0.012096\ g/cm^3$$

The expanded uncertainty (U), is:

$$U = k \times u_c$$

for k = 2, for approximately 95% level of confidence and two significant figures:

$$U = 2 \times 0.012096\ g/cm^3 = 0.024\ g/cm^3\ .$$

Option D: Use of the result of multiple participations on a PTS, excluding the bias component

All the results from the participation of the laboratory on 10 rounds of a PTS for density are selected, all with satisfactory results (table 5).

Therefore, the standard uncertainty is the combination of precision, $u_{precision}$, and bias, u_{bias}, data but taking into account that all the z-scores were satisfactory, and using the same assumptions as Option A (Table 5),

$$u_c = S_R$$

where S_R is the pooled standard deviation of all the individual S_R for 10 rounds described on Table 5.

The pooled S_R is calculated as follows:

$$S_{pool} = \sqrt{\frac{\sum_{i=1}^{j} DF_i \times S^2{}_{R(i)}}{\sum DF_i}}$$

$$\Leftrightarrow$$

$$S_{pool} = \sqrt{\frac{60 \times 0.0024^2 + 56 \times 0.0019^2 + 46 \times 0.0022^2 + 43 \times 0.0031^2 + 41 \times 0.0020^2 + 56 \times 0.0027^2 + 43 \times 0.0021^2 + 49 \times 0.0031^2 + 42 \times 0.0018^2 + 56 \times 0.0023^2}{(60+56+46+43+41+56+43+49+42+56)}} = 0.01193$$

Hence,

$$u_c = S_R = 0.01193\ g/cm^3$$

The expanded uncertainty (U), is: $U = k \times u_c$

For k = 2, for approximately 95% level of confidence and two significant figures:

$$U = 2 \times 0.01193\ g/cm^3 = 0.024\ g/cm^3.$$

2.3 Reporting of results

For a density of 0.863 g/cm^3, the result should be reported as:

Option A: 0.863 ± 0.008 g/cm^3, (k = 2, for approximately 95% level of confidence).

Option B: 0.863 ± 0.006 g/cm^3, (k = 2, for approximately 95% level of confidence).

Option C: 0.863 ± 0.024 g/cm^3, (k = 2, for approximately 95% level of confidence).

Option D: 0.863 ± 0.024 g/cm^3, (k = 2, for approximately 95% level of confidence).

For convenience, the results were reported with three decimal places.

Option E: Use of the result of multiple participations on a PTS as an overall assessment of the test method

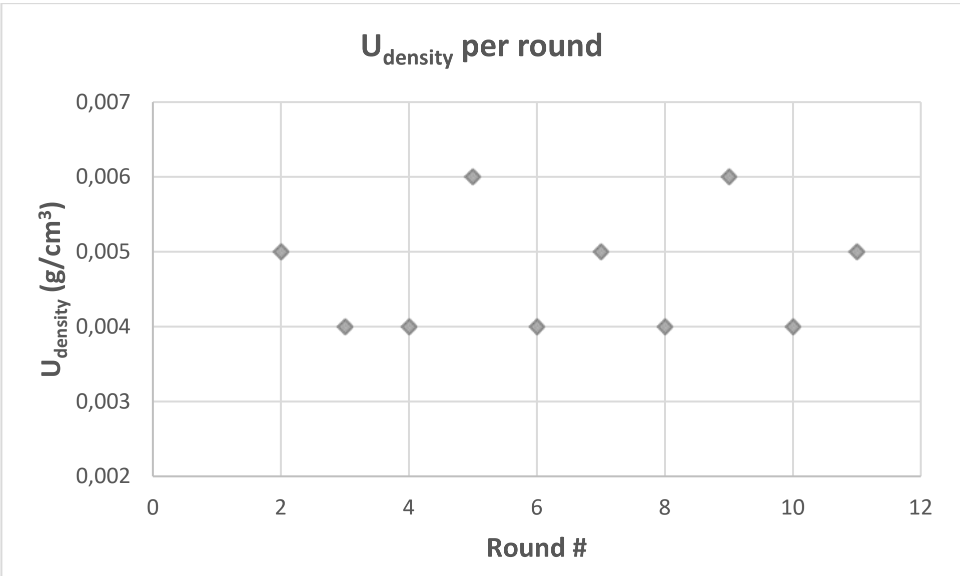
In this case, U for each round should be determined individually, according to Option A, and a plot of U *vs* Round ID should be performed, so that the laboratory is able to monitor its performance within a specific time-frame. If the data from Table 5 is used, the raw data for the plot obtained are provided in Table 9. The values reported for S_R include $S_{within\text{-}lab}$ and $S_{between\text{-}lab}$.

Table 9. Calculation of U using original data from Table 5

Round ID	Lab. Mean Result (g/cm³)	$S_{within\text{-}lab}$	Assigned Value (g/cm³)	$U_{assigned\ value}$ (k = 2*)	S_R	U (k = 2*)
1 (n=3)	1.119	0.0015	1.119	0.003	0.0024	0.005
2 (n=3)	0.845	0.0021	0.846	0.003	0.0019	0.004
3 (n=3)	0.862	0.0019	0.865	0.003	0.0022	0.004
4 (n=3)	1.261	0.0028	1.261	0.003	0.0031	0.006
5 (n=3)	0.958	0.0022	0.96	0.004	0.0020	0.004
6 (n=3)	0.911	0.0016	0.912	0.004	0.0027	0.005
7 (n=3)	0.842	0.0018	0.845	0.004	0.0021	0.004
8 (n=3)	0.914	0.0025	0.918	0.004	0.0031	0.006
9 (n=3)	0.911	0.0015	0.911	0.004	0.0018	0.004
10(n=3)	0.863	0.0020	0.864	0.003	0.0023	0.005

n = number of replicates
* for approximately 95% level of confidence

The main purpose of plotting U in each round is to check its evolution. Moreover, for a well established test method, where acceptance criteria for precision have been established, a line can be added, in order for the laboratory to control if the acceptance criteria is suitable taking into account the uncertainty estimated (Plot 1).



Plot 1. Expanded uncertainty (U) estimated from the results of PTS, per round.

References

1. Selection, Interpretation and Use of Proficiency Testing Schemes (2011), 2nd Edition, Eurachem.
2. Eurachem/CITAG Guide CG4, Quantifying Uncertainty in Analytical Measurement, 3rd Edition (2012).
3. Eurolab Technical Report 1/2007 – Measurement uncertainty revisited.
4. Miller JC & Miller JN, Statistics for Analytical Chemistry, 3rd Edition (UK, 1993)
5. Eurolab Technical Report 1/2006 - Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results – Measurement uncertainty revisited.
6. Eurolab Technical Report 1/2002 - Measurement Uncertainty in Testing.
7. Nordtest Technical Report 537, Handbook For Calculation of Measurement Uncertainty In Environmental Laboratories, Edition 3.1, 11/2012.
8. BARWICK, V.J., ELLISON, S.L.R.: VAM Project 3.2.1 Development and Harmonisation of Measurement Uncertainty Principles. Part (d): Protocol for uncertainty evaluation from validation data (Jan 2000).