USE OF INTERLABORATORY COMPARISON DATA BY LABORATORIES: Part A

Relevance of interlaboratory comparisons

Interlaboratory comparisons (ILCs) are performed for various reasons [1], e.g.

- to validate test procedures,
- to certify reference materials,
- to assess the competence of laboratories (proficiency testing),
- or more general, to investigate the degree of comparability among laboratories.

Irrespective of the specific aim(s) of an ILC, the results can be used by a participating laboratory

- to check the performance of its test procedures and / or its staff,
- to demonstrate its competence towards clients and accreditation bodies,
- to gain useful information for the evaluation of its measurement uncertainty.

The scoring of ILC data

In proficiency tests (PTs) the PT providers often evaluate scores as a quantitative measure of the laboratory performance. There is a number of different scores, three of which are used very commonly [2].

z-score:

$$z = \frac{|x_{lab} - x_{ass}|}{s}$$

z‘-score:

$$z' = \frac{|x_{lab} - x_{ass}|}{\sqrt{s^2 + u_{ass}^2}}$$

zeta score:

$$zeta = \frac{|x_{lab} - x_{ass}|}{\sqrt{u_{ass}^2 + u_{lab}^2}}$$

(with x_{lab}: laboratory result, x_{ass}: assigned value, s: standard deviation of performance assessment, u_{ass}: standard uncertainty of the assigned value, u_{lab}: standard uncertainty of the laboratory result)

The numerator of both scores gives the (absolute) difference between the laboratory result and the assigned value which can either be established by one or more reference laboratories or be derived as consensus value from the group of participating laboratories. The standard deviation s in the denominator of the z-score is a measure of the actual or accepted variability of the results. When the uncertainty of the assigned value is larger than 0,3s, this uncertainty is added to standard deviation at the denominator of z’-score. The denominator of the zeta score represents the combined standard uncertainty of the difference in the numerator. Thus, both scores have in common that the actual difference between the laboratory result and the assigned value is assessed against an estimate of the (expected or acceptable) spread of results. The difference between the two scores is that the z-score assesses all laboratories against the same numerical value whereas the zeta score allows the individually claimed accuracy of a laboratory to be taken into account.[1] It is recommended that

[1] Another score often used in the calibration field and quite similar to the zeta-score is the so called E_n number, defined as

proficiency testing providers use robust statistics to reject outliers. In the last revision of ISO 13528,several algoritims are described with efficiency and breakdown points of procedures.

PT providers often use the following classification of the result of a participating laboratory:

z, z’, zeta ≤ 2:	satisfactory result
2 < z, z’, zeta < 3:	questionable result,
z, z’, zeta > 3:	unsatisfactory result.

The analysis of the ILC data by the laboratory

In order to use the result of an ILC for the purposes mentioned in clause 1, the laboratory, after participation in an ILC, should carefully analyse its result taking into account existing information, e.g.

- statements on the uncertainty of the used test procedure in standards, literature etc.,
- their own evaluation of this uncertainty,
- standard deviation of the results of all laboratories participating in this ILC,
- uncertainty acceptable for the laboratory and its clients.

Even if the organiser of an ILC provides a classification of the results as satisfactory or unsatisfactory the laboratory should not simply rely on this judgement. If e.g. for the z-score a standard deviation s is used by the organiser which is not fit-for-purpose (ffp) for the laboratory, it might calculate a modified z-score using an sffp according to its own or its clients’ needs [3].

In the case of an unsatisfactory result the laboratory should perform a root cause analysis and, based on its result, should take corrective actions. Sometimes the organiser of the ILC might offer advice. After implementation of the corrective actions the laboratory should prove their effectiveness, e.g. by

- use of a suitable reference material,
- participation in another ILC.

Furthermore the results of ILCs are an important tool to verify the evaluation of the uncertainty of the used test procedures [4, 5, 6]. If the laboratory estimates of the uncertainty turn out to be too conservative or too optimistic, the laboratory should adapt them accordingly.

Conclusions

Irrespective of an eventual classification of the ILC results as satisfactory or unsatisfactory by the organiser of an ILC, the participating laboratory should carefully analyse its results on the basis of its own criteria. If a result then turns out to be unsatisfactory, the laboratory should take appropriate corrective actions and should satisfy itself that these actions have been effective.

Additionally the results of an ILC should be used to verify or improve the estimates of measurement uncertainty of the used test procedures.

$$E_n = \frac{|x_{lab} - x_{ass}|}{\sqrt{U_{ass}^2 + U_{lab}^2}}$$

In this case the denominator contains the expanded uncertainties U (k=2) instead of the standard uncertainties u (k=1) and consequently E_n<1 is the criterium for a satisfactory result while E_n>1 is considered as unsatisfactory.

References

[1] ISO/IEC 17043, Conformity assessment – General requirements for proficiency testing, 2010

[2] ISO 13528, Statistical methods for use in proficiency testing by interlaboratory comparisons, 2015

[3] IUPAC, The international harmonized protocol for the proficiency testing of analytical chemistry laboratories, Pure Appl. Chem., 78 (2006), 145 - 196

[4] NORDTEST Technical Report 537, Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories, 2017, www.nordicinnovation.net/nordtest.cfm

[5] EUROLAB Technical Report 1/2006, Guide to the Evaluation of Measurement Uncertainty for Quantitative Results, www.eurolab.org

[6] EUROLAB Technical Report, 1/2007, Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation, to be published