THE ASSESSMENT OF THE TRUENESS OF A MEASUREMENT PROCEDURE BY USE OF A REFERENCE MATERIAL (RM)

Background

The terms accuracy, trueness and precision are explained e.g. in [1]. Accuracy as an umbrella term generally means the agreement of a measurement result with the (conventional) true value. For a series of repeated measurements accuracy can be split up into trueness and precision. The term precision characterises the dispersion between the single results, while trueness characterises the difference between the mean value of the series and the (conventional) true value.

Precision strongly depends on the conditions under which the series of measurement results are obtained. If the measurements are performed within the same laboratory by the same operator, using the same measurement procedure and equipment and within a short period of time, the precision under so-called repeatability conditions is relatively high, i.e. the standard deviation of the results is relatively low. Under reproducibility conditions, i.e. results obtained by different laboratories and different operators, using the same measurement procedure but different equipment, the precision is lower or the standard deviation of the results higher, respectively. The so-called intermediate precision conditions (called within-laboratory reproducibility conditions in [1]) are an intermediate case as the results are obtained within the same laboratory using the same measurement procedure maybe by different operators over a longer period of time.

While the evaluation of the precision of a measurement procedure (under repeatability or intermediate precision conditions) is rather straightforward for a laboratory, the trueness of the procedure is more difficult to assess. The use of a suitable reference material is one method which will be described below.

Use of a (certified) reference material

If a (certified) reference material is available whose reference quantity can be measured with the measurement procedure in question, a comparison of the result obtained and the reference value can be used to assess the trueness of the procedure.

The reference quantity of the RM is measured n times by the laboratory providing the single measured quantity values $x_{m,i}$, the mean value $\bar{x}_m$ and the standard deviation s_m. The absolute value of the difference Δ between the certified reference value x_{ref} and the mean measured value

$$|\Delta| = |\bar{x}_m - x_{ref}| \qquad \text{(eq. 1)}$$

is compared with the uncertainty of this difference caused by the uncertainty of the reference value u_{ref} taken from the certificate and the uncertainty of the measured mean value u_m

$$u_\Delta = \sqrt{u_{ref}^2 + u_m^2} \qquad \text{(eq. 2)}$$

where the standard uncertainty u_m can be estimated in a first approximation from the standard deviation of the measurement series:

$$u_m = \frac{s_m}{\sqrt{n}} \qquad \text{(eq. 3)}$$

The measured mean value is compatible with the reference value (i.e. there is no experimental evidence for a bias), if the following criterion holds:

$$|\Delta| \le k \cdot u_\Delta = k \cdot \sqrt{u_{ref}^2 + \frac{s_m^2}{n}} \qquad \text{(eq. 4)}$$

The coverage factor k is usually chosen as k = 2, which corresponds with a confidence interval of the uncertainty of approximately 95%[1].

Example:

Ochratoxin A (OTA) is a mycotoxin with inter alia carcinogenic, nephrotoxic and teratogenic properties. It can be present as a natural contaminant in some crops, e.g. in cereals, wine and coffee. A maximum allowed limit is defined in the EU [3]. The analysis can be performed by HPLC. A CRM for e.g. coffee is available [4].

A laboratory obtained from a measurement series (n=4) on this CRM the following results w_1 = 6.29 μg/kg; w_2 = 4.63 μg/kg; w_3 = 5.34 μg/kg; w_4 = 5.46 μg/kg. From these results, a mean value w_m=5.43 μg/kg and a standard deviation s=0.68 μg/kg are calculated. The OTA content in the CRM is certified as w_{ref} = 6.1±0.6 μg/kg where the given uncertainty is an expanded uncertainty U_{ref} with k=2. Thus the standard uncertainty of the CRM is $u_{ref} = U_{ref} / k$ = 0.3 μg/kg. Plugging all these values in eq. 4 results in:

$$\Delta = |5.43 - 6.1| = 0.67 < 0.91 = 2 \cdot \sqrt{0.3^2 + \frac{0.68^2}{4}} \text{ [μg/kg]}$$

As the criterion (eq. 4) is fulfilled the laboratory's results are compatible with the certified value.

Conclusions if the criterion is not fulfilled

In practice, quite often this criterion might not be fulfilled as the assumption in eq. 3 that the uncertainty of the measurement procedure could be evaluated from the standard deviation of an individual measurement series alone will often significantly underestimate the uncertainty. In particular this will be the case if the measurements were performed under repeatability conditions. This can be seen as an example from the results of the interlaboratory comparison which was organised to characterise the OTA reference material [4]. Fig. 1 shows the certified value and its expanded uncertainty together with the results of the participating competent laboratories. The latter are plotted as mean values (±) one laboratory standard deviation. Although not all of the results fulfilled the criterion (eq. 4) they could be used to determine the certified value.

If the criterion is not fulfilled there are two options to deal with this result [1]:

1) correction:

If there is reason to assume that the incompatibility of the measurement result is caused by a constant bias of the measurement procedure, the difference Δ (eq. 1) can be used to correct all future results obtained with this procedure:

$$x_{m,\text{corrected}} = x_m - \Delta \qquad \text{(eq. 5)}$$

In the uncertainty budget the standard uncertainty of the correction u_Δ should be added.

2) expansion of the measurement uncertainty

If there is doubt that the difference Δ is reflecting a constant bias of the method, one should make allowance for Δ when evaluating the measurement uncertainty $u(x)$ connected with the procedure.

[1] This statement is only valid if u_Δ is a reliable estimate of the standard uncertainty of the difference. For small measurement series (small n), i.e. low degree of freedom ν, a more accurate approach would replace k=2 by the corresponding value t(ν) from the Student's distribution (see e.g. annex G in [2]).

$$u(x) = \sqrt{\frac{s_m^2}{n} + u_{ref}^2 + \Delta^2} \qquad \text{(eq. 6)}$$

The result of eq. 6 is a rather conservative estimation of the measurement uncertainty which should be confirmed from time to time by repeated measurements of the RM and adapted, if necessary.

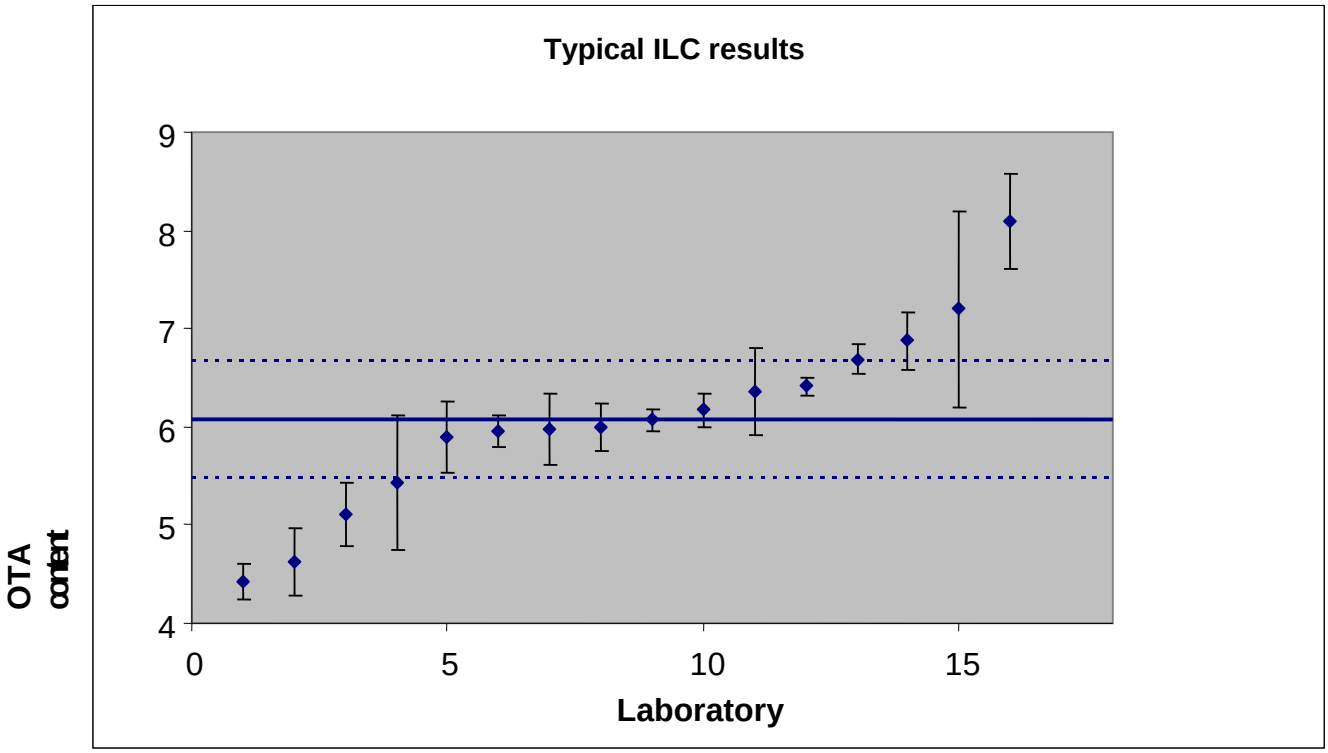

Figure 1: Results of the interlaboratory comparison performed to characterise the OTA-RM [4]. The certified value (full line) is shown together with the interval (broken lines) composed of the expanded uncertainty (k=2). The error bars of the individual laboratory mean values characterise the laboratories’ precision expressed as one standard deviation

References

[1] Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results, EUROLAB Technical Report 1/2006, www.eurolab.org
[2] JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM), http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf
[3] Commission Regulation (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs
[4] ERM®-BD475 Ochratoxin A (OTA) in ground roasted coffee, http://www.rm-certificates.bam.de/de/rm-certificates_media/rm_cert_food/erm_bd475e.pdf