The trueness criterion: an alternative concept for the evaluation of proficiency tests

proofacs

B.K. Schindler

PROOF-ACS GmbH, Tempowerkring 1, 21079 Hamburg, Germany; email: birgit.schindler@proof-acs.de

Introduction

Proficiency testing (PT) is a key element of external quality assurance in analytical laboratories according to ISO/IEC 17025. In practice, data from PTs are used to ensure and demonstrate the validity of analytical methods.

The food market stakeholders require analytical results, which are used for crucial decisions on whether or not...

... the quality of products is adequate.

- ... delivered goods are marketable (release-analysis).
- ... the applied pesticides are below the MRL.
- \ldots products are ready for consumption.

In day-to-day routine, analytical results are used as "one-off" results to answer the question

"What concentration levels are really present in the sample?"

Food safety thus requires

- Reliable results,
- **True** results, which represent the actual analyte concentration in the analysed sample,
- **Repeatable** results, with a possibly lowest intra- and inter-day variation.

Consequently, the accepted results of laboratories participating in PTs should not significantly deviate of the true level. Therefore, the target deviation applied in a PT has to be quite narrow.

One approach related to the assessment of ...



... the analytical performance of laboratories is the **COMPARABILITY CRITERION**

- The results of the labs are compared with the *average* of the results of all participants.
- Reference value: statistically derived mean, calculated of the results of the participants.
- Applied model: z-score.
- Fit for purpose criterion: 2-fold target standard

Trueness vs. z-score – Examples of Proficiency tests

Example 1: Perchlorate in courgette [2]

- Contaminant in fruits and vegetables
- Single-Residue Method
- LC-MS/MS



• Internal standard: ¹⁸0₄-labelled perchlorate

Comparison of the two models with respect to the accepted ranges – illustrated with results of an actual PT on perchlorate in courgette:

70-120 % model		z-score model
Piked Level 70% recovery 70% recovery 70% recovery 120% recovery 500 120% recovery 120% r	Accepted assurance assurance assurance assurance as a second assurance as a second as a se	SNED VALUE
70-120 %	Applied model	z-score
Spiked level: 0.065 mg/kg	Reference	Assigned value: 0.066 mg/kg
Results within 70–120 % recovery of the spiked level	Criterion for acceptance of results as satisfactory	Results with a z-score ≤ 2 , based on the target standard deviation according to Horwitz
70–120% of the spiked level; here: 0.046–0.078 mg/kg	Accepted range	58–144 % of the assigned value; here: 0.038–0.095 mg/kg
→ More sophisticated	Level of difficulty	→ More generous

Example 2: p,p-DDT in black tea [3]

Black tea – a demanding matrix
Pesticide multi-method analysis required
p,p-DDT: GC-MS/MS
Results are highly influenced by the appropriate condition of the liner



deviation according to Horwitz [1].

Another approach considers ...



... the true analyte level in the test sample – **TRUENESS CRITERION**

- The results of the labs are compared with the *actual analyte concentration in the sample*.
- Reference: spiked value.
- Applied model: 70-120 % model.
- fit for purpose criterion: 70 up to 120 % recovery of the spiked level.

The application of the trueness criterion requires:





Absence of analytes in the raw material

Stability of the analytes in the test material



Verification of the true analyte level in the test samples The application of the two models to the results of a PT on pesticides in tea results in completely different outcomes related to the evaluation of analytical performances of participating laboratories:

- Lab 5 and lab 15 would be categorised as non-satisfying (z-score > 2) according to the z-score model → not comparable.
- However, among the labs in this test, lab 5 and lab 15 are the ones that provided true results (within 70–120 % recovery of the spiked level).



Conclusion

The trueness criterion provides information related to the ability of a laboratory to reliably quantify the actual analyte concentration in the sample. The outcome of several PTs, even in complex matrixes, showed that a target range of 70 to 120 % of the spiked level is feasible. Taking into consideration that PT samples are analysed with special care, a recovery of 70–120 % is considered as an appropriate performance criterion as proposed by the SANCO Document 12571/2013 [4] for method validation.

The trueness criterion (70–120 % model) takes into consideration the requirements of laboratories' clients in a more appropriate way and answers the question "What concentration levels are really present in the sample?". The additional information on the trueness of the results contributes significantly to the identification of shortcomings in order to improve the laboratories' analytical performances.

REFERENCES: [1] Horwitz W., Evaluation of Analytical Methods Used for Regulation of Foods and Drugs. Anal Chem. 1982; 54(1): 67A-76A. [2] PROOF-ACS GmbH, Ring Test P1410-RT "Chlorate and perchlorate in basil and courgette, February 2014 (www.proof-acs.de). [3] Bundesverband Naturkost Naturwaren, Laboratory Performance Accessment, Analysis of Pesticides and Anthraquinone in Black Tea, May 2013 (www.n-bnn.de). [4] European Commission Health & Consumer Protection Directorate-General, Guidance document on analytical quality control and validation procedures for pesticide residue analysis in food and feed, SANCO/12571/2013, Implemented 2014 (http://www.eurl-pesticides.eu/library/docs/allcrl/AqcGuidance_Sanco_2013_12571.pdf).