codex alimentarius commission

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

WORLD HEALTH ORGANIZATION

JOINT OFFICE: Viale delle Terme di Caracalla 00153 ROME Tel: 39 06 57051 www.codexalimentarius.net Email: codex@fao.org Facsimile: 39 06 5705 4593

Agenda Item 7 (a) **CX/PR 08/40/5**
March 2008

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON PESTICIDE RESIDUES
Fortieth Session
Hangzhou, China, 14 - 19 April 2008

DISCUSSION PAPER ON THE ESTIMATION OF UNCERTAINTY OF RESULTS FOR THE DETERMINATION OF PESTICIDE RESIDUES

(Prepared by IAEA)

INTRODUCTION

1. The estimation of measurement uncertainty (MU) in pesticide residue analysis particularly from first principles is extremely laborious, as residue determinations at trace levels (0.001 - 10 mg/kg) are often subject to considerable analytical variability. In order to minimise international trade disputes resulting from questionable exceedences of regulatory limits, the estimation and reporting of MU is essential for demonstrating the equivalence of analytical results generated by exporting and importing countries.

2. Laboratories are often limited in financial and personnel resources as well as in terms of time for the testing of samples. It is often impractical in routine laboratory work to estimate individual values for countless commodity/pesticide combinations[1], especially when using the more rigorous bottom-up approach. Consequently, it was proposed in ALINORM 07/30/24, paras. 156-160 to develop a simplified guidance document for the estimation of MU, that is, one based on proficiency testing (PT) results, method validation and quality control data.

3. At the 39th CCPR Meeting, it was agreed that a discussion paper which would form the basis of a guidance document should be prepared and tabled at the 40th CCPR Meeting. The Committee then would decide whether to undertake new work at its next session. An electronic working group (EWG) platform was set up by IAEA to facilitate progress of a draft document and for which a number of interested country delegates have subsequently enrolled (http://elearning.iaea.org/ATutor/login.php).

4. It is envisioned that a guidance document based on the discussion paper would be developed taking into consideration relevant publications and input from EWG participants. Codex Member States are therefore encouraged to involve their official laboratories to enable broad participation by experts and interested parties. The output should be a straightforward draft guideline based on empirical top-down concepts such as formulated by Horwitz.

[1] More than 1000 pesticides are known worldwide; more than 220 pesticides have a Codex Reference Number

5. The guidance document is intended to assist the understanding and adoption of the uncertainty concept in pesticide residue laboratories. The key objectives of the document should:

(a) be in compliance with ISO/IEC 17025;

(b) consider the high complexity of pesticide residue analysis (e.g. several working steps, limited resources, high number of possible commodity/pesticide combinations);

(c) be practically oriented and straightforward guidance (e.g. based on empirical top-down concepts such as formulated by Horwitz);

(d) not consider uncertainties related to sampling.

BACKGROUND

6. The need to control analytical procedures – and consequently the necessity of quantitative expressions of MU – is widely recognized. The technical part of ISO Standard 17025 requires MU as an essential parameter for which laboratories must establish estimates[2]. However, the routine estimation of MU appears to be problematic in many food control laboratories.

7. A number of guidance documents are available which describe different approaches towards estimating MU. With regard to pesticide residue analysis, bottom-up calculations in particular are perceived as overly complicated and extremely laborious. This may in part be due to the fact that the MU concept had originally been developed for physical measurements, the influencing factors and analytical parameters of which are limited and rather straightforward to define and calculate. The concept cannot easily be transposed to complicated and multi-factorial chemical residue analysis procedures.

8. Pesticide residue methods involve several independent processes: (a) sample preparation, processing and storage, (b) extraction of analyte(s), (c) clean-up, (d) quantitation of analyte(s). Each sub-procedure can involve several steps including sample comminution, weighing, pipetting, calibration and so on. Each procedural and/or working step may influence MU values, possibly differing from analyte to analyte, from commodity to commodity and is likely to be concentration dependent.

9. Accordingly, there is still insufficient understanding and incomplete adoption of the uncertainty concept. Common procedures for the particular purpose of pesticide residue analysis in food, with its diversity of influencing factors, are still missing. Therefore, a specific guidance applicable to pesticide residue analysis of food is deemed useful towards simplification and more likely acceptance of the uncertainty concept.

10. Taking into consideration relevant guidelines and reports (see below), particularly in terms of top-down approaches on MU, this discussion paper intends to outline specific pathways applicable to pesticide residues in food, including some practical examples.

MU CONCEPTS IN PESTICIDE RESIDUE ANALYSIS

11. Difficulties related to MU were also discussed by CCMAS in 2007 (see ALINORM 07/30/23, paras 6-10). Although pesticide residue analysis in its complexity was not of particular concern, the matter is perceived similarly in CX/MAS 07/28/2-Add.2. The United Kingdom had prepared a guidance document on MU explaining the situation and drawing together various developments in the area. In parts A to L the main relevant approaches as outlined in different publications are summarized and discussed. However, it does not specifically develop guidance for particular analytical procedures as to which approach would be applicable for which purpose.

---

[2] See ISO/IEC Standard 17025, Para. 5.10.3.1: In addition … test reports shall … include the following: "c) where applicable, a statement on the **estimated uncertainty** of measurement; information on uncertainty is needed in test reports when it is relevant to the validity of application of the test results, when a client's instruction so requires, or when the uncertainty affects compliance to a specification limit."

12. ISO/TS 21748:2004[3] provides additional mathematical concepts especially for estimating zones of acceptance and rejection around analytical values; however, straightforward top-down approaches are also discussed. One important statement relevant in this context is that the reproducibility standard deviation obtained from a collaborative study is considered a valid basis for MU evaluation. If accuracy (or trueness) data can be utilized, e.g. with respect to an established reference value based on (certified) reference material, then uncertainty associated with the estimated bias should be included in the MU budget. The process of evaluating uncertainty according to ISO/TS 21748 comprises the following elements:

- Repeatability, reproducibility and bias estimates from collaborative study.
- Laboratory bias and precision within that expected on the basis of the collaborative study.
- Laboratory bias and precision under control; effects appropriately combined to form a combined uncertainty estimate.

13. In guideline EA-4/16[4] it is recognized that "laboratories cannot in general be expected to initiate scientific research to assess the uncertainties associated with their measurements and tests". The guideline, among others, describes the use of validation and method performance data for uncertainty evaluation. Data accumulated during validation and verification of test methods, interlaboratory studies according to ISO 5725, accumulated quality control data, and proficiency testing schemes typically characterize test method performance.

14. SANCO[5] document (ACQ Guidelines) supports this line of action towards evaluating MU associated with proficiency test results. Eurolab Technical Report[6] and NORDTEST Report[7] TR 537 outline in greater detail, among others, the use of method validation and PT data for estimating MU.

ALTERNATIVE APPROACHES

15. A comprehensive and feasible MU concept is not provided by existing guidelines in terms of the practical application to pesticide residue analysis in foodstuffs. Calculating uncertainty budgets for thousands of relevant pesticide/crop combinations and dozens of analytical methods used in pesticide residue analysis is not practical in routine laboratory operations. Empirical approaches proposed recently show alternatives also for pesticide residue analysis of foodstuffs.

16. Practically oriented and straightforward guidance for application in the determination of pesticide residues in foodstuffs could be made available through top-down MU concepts. Validation data, repeatability, reproducibility, outcomes of PT schemes can be utilized for simplified MU estimation applicable in food control laboratories.

17. Based on a series of PT schemes, the ACQ Guidelines of the EC indicate that actual and target values according to different performance and quality criteria were well within the same order of magnitude. For instance, values derived from Fitness-for-Purpose (FFP), the Horwitz equation (see annex) and standard deviation calculated from EC PT schemes, after rejection of outliers (Qn), expressed in (%), were very similar. Accordingly, the evaluation of the recent EC PT schemes demonstrates that a FFP variability of 25% can be accepted as a sound representation of performance under these circumstances. As a consequence, taking the 25% variability as a standard deviation would lead to a generalized assumption of ±50% MU.

---

[3] Technical Specification ISO/TS 21748:2004: Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, First edition 2004-03-15

[4] EA-4/16 EA guidelines on the expression of uncertainty in quantitative testing, December 2003 rev00

[5] Document N° SANCO/2007/3131 - METHOD VALIDATION AND QUALITY CONTROL FOR PESTICIDE RESIDUE ANALYSIS IN FOOD AND FEED (www.crl-pesticides.eu)

[6] Eurolab Technical Report No. 1/2007, March 2007, Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation (www.eurolab.org)

[7] NORDTEST Report TR 537, HANDBOOK FOR CALCULATION OF MEASUREMENT UNCERTAINTY IN ENVIRONMENTAL LABORATORIES, EDITION 2

18. Accepting such a generalized approximation for pesticide multi-residue analysis, a generalized top-down approach might result in larger MU values than such derived for each individual pesticide/commodity combination by systematic bottom-up calculations. However, the application of generic MU is considerably more practical and easier to obtain. Generalized values, like ±50% MU, might enlarge safety margins around MRLs for a number of pesticides. This would make a difference especially when getting close to MRLs/ trigger values. On the other hand, for laboratories it would mean a considerable rationalization in terms of time, resources and workload that otherwise has to be devoted to systematic bottom-up MU evaluation.

Note: Analytical recovery values deviating from 100% are not corrected for in most cases. This missing correction alone could cause differences of up to 30%. Uncertainties related to sampling[8] are not widely considered as yet, but can be even higher. Also MRLs are mostly representing residue concentrations of $^1/_{100}$ to $^1/_{1000}$ in relation to toxicologically relevant levels, unless an exceedence concerns an Acute Reference Dose at the same time. In this context it may therefore be questioned whether a sharp definition of trigger values and associated safety areas around MRLs should be handled with extreme stringency.

DEVELOPMENT OF A SPECIFIC GUIDANCE ON MU EVALUATION FOR PESTICIDE RESIDUE ANALYSIS BASED ON EMPIRICAL DATA

19. Data derived from systematic method validation for verifying recovery values and associated standard deviations characterizing the use of analytical methods can be utilized. A step by step practical guidance should incorporate representative examples of commonly used analytical methods.

20. In practical terms a guidance document would incorporate empirical data and outcomes of PT schemes. In particular the following information and data could be utilized:

- Concentration dependent RSDs according to Horwitz could be utilized for estimating MU, e.g., for fatty matrices, whereas constant RSDs of 25% might become applicable for non-fatty matrices over the range of relevant trace levels.
- Sufficient method validation data, including recovery, repeatability and reproducibility.
- Control Charts data derived from the routine application of methods documented.
- Participation in PT schemes.

21. Implementing a PT-based simplified ±50% MU approach should only be used by individual laboratories if the following analytical performance and quality criteria can be demonstrated:

- Within-laboratory SD smaller than the between-laboratories SD.
- Successful participation in PT schemes (z-score $\leq|2|$ for 95%, z-score $\leq|3|$ for not more than 5% of the values).
- Little method and/or laboratory bias for recovery tests.
- Verification of analytical performance by regularly analysing suitable reference material, if available.

RECOMMENDATION / PROPOSAL

22. As is an emerging practice in the EC and elsewhere already, empirical top-down estimation of ±50% MU could complement a mathematically stringent bottom-up calculation model if the respective empirical quality criteria are met. Alternatively the Horwitz formula approach of estimating concentration dependant MU based on the evaluation of results of a multitude of interlaboratory collaborative tests could be applied as well.

[8] M H Ramsey and S L R Ellison (eds.) Eurachem/EUROLAB/ CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches Eurachem (2007). ISBN 978 0 948926 26 6. Available from the Eurachem secretariat.

23. It is proposed to further develop a specific guidance for the application of empirical MU concepts particularly applicable in the field of pesticide residue analysis of foodstuffs.

Note: Laboratories uncomfortable with these empirical approaches or where such is not deemed applicable may wish to apply the bottom-up step-by-step calculation to specifically generate distinct individual uncertainty estimates as given in references[9] providing guidance on bottom-up estimation MU, including treatment of result levels eventually conflicting with trigger values[10].

[9] primarily: EURACHEM/CITAC Guide CG 4, Quantifying Uncertainty in Analytical Measurement, Second Edition, QUAM 2000.1

[10] EURACHEM/CITAC Guide, Use of uncertainty information in compliance assessment, First Edition 2007

ANNEX: PRACTICAL MU ESTIMATION BASED ON TOP-DOWN APPROACHES

Limitations

In general, proficiency tests are not carried out frequently enough to provide good estimates of the performance of an individual laboratory's implementation of a test method. However, in the special case where:

- the types of test items used in the scheme are appropriate to the types tested routinely,
- the assigned values in each round are traceable to appropriate reference values, and,
- the uncertainty associated with the assigned value is small compared with the observed spread of results,

the dispersion of the differences between the reported values and the assigned values obtained in repeated rounds provides a basis for an evaluation of the uncertainty (see Eurolab and NORDTEST references).

A PT-based top-down approach is therefore applicable where PT data support this. Referring to EC-PT schemes this approach would be different for various matrices and pesticide/ matrix combinations.

Other matrix/pesticide combinations would need separate MU evaluation following the guidelines and approaches given elsewhere.

Underlying calculation formulas and statistics for PT based estimation of MU

Within-laboratory reproducibility standard deviation is combined with estimates of the method and laboratory bias using PT data:

$$U = k * u = \sqrt{u(R_W)^2 + u(bias)^2}$$

where:

$$u(bias) = \sqrt{RMS_{bias}{}^2 + u(C_{ref})^2}$$

and:

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}}$$

and:

$$u(C_{ref}) = \frac{S_R}{\sqrt{n}}$$

with:

U	=	expanded uncertainty
k	=	coverage factor
u	=	combined standard uncertainty
$u(R_W)$	=	within-laboratory reproducibility standard deviation
u(bias)	=	uncertainty component from method and laboratory bias, estimated from PT data
RMS_{bias}	=	root mean square of bias values
$bias_i$	=	bias of compound i
$u(C_{ref})$	=	average uncertainty of assigned values
S_R	=	interlaboratory standard deviation of PT

n	=	mean number of PT participants

Practical application

(1) Prerequisites:

- The laboratory has demonstrated its technical capability to generate reliable results at the required level of quality, i.e. by:
  - sound validation data for the respective analytical method;
  - acceptable quality control data, e.g. control charts for respective methods and compounds;
  - successful participation in PT schemes which fulfil PT quality criteria according to the Harmonized Protocol[11], ISO Guide 43-1 etc.;
  - the laboratory has been rated as well-performing (e.g. Category A, Sufficient Scope at 90%, e.g. according to PT evaluation within the EC).

(2) Uncertainty evaluation using laboratory evaluation data:

- identification of the main sources of uncertainty (weighing, calibration, purity, temperature, volumetric glassware, …);
- evaluation of the order of magnitude of the variability of basic laboratory operations in relation to the overall variability of the procedure.
- expected result:
  - variability of basic laboratory operations almost negligible;
  - random run-to-run variability as the principal source of MU.
- estimation of overall bias and recoveries from in-house validation experiments (fortification, spiking, reference materials, …):
  - the mean of the resulting relative standard deviation taken as relative uncertainty is associated with random variation;
  - no significant bias.

(3) Comparison with PT results:

- series of PT rounds with slightly varying concentrations and matrices,
- standard deviation for the relative differences of valid data is comparable to the expected relative standard deviation (comparing PT results with real laboratory data).

(4) Verification of uncertainty estimates:

- checks using observed within-laboratory precision,
- checks using certified reference materials or suitable test materials,
- checks using reference methods,
- checks based on the results of PT (including external QA data or measurement audits),
- checks based on comparison of results with other laboratories,
- comparison with other uncertainty estimates based on different approaches or different data.

(5) Conclusion:

- PT data can provide strong support for the laboratory estimate of MU based on validation data,

[11] M Thompson, S L R Ellison, R Wood; The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report); Pure Appl. Chem. 78(1) 145-196 (2006)

– PT data can form the basis for estimating MU, using the dispersion of relative differences.

Evaluation of uncertainty estimates against PT results

Checking the quality of uncertainty estimates may apply the zeta (ζ) score formula:

$$\zeta = \frac{x - x_a}{u(x)^2 + u(x_a)^2}$$

with:

x = laboratory result

x_a = assigned value

u(x) = standard uncertainty of laboratory results

$u(x_a)$ = standard uncertainty of assigned values

Uncertainties are considered overestimated if |ζ| is significantly less than 2; correct if |ζ| is in the range 0 to 2; underestimated if |ζ| is frequently over 2. Equivalent to the zeta score, the E_n number ($E_n = \zeta / 2$) can be calculated by replacing the expanded uncertainties U(x) and $U(x_a)$ by u(x) and $u(x_a)$ in the above formula.

MU estimation based on Horwitz formulas

Similarly to the PT based approach MU may be estimated using empirical Horwitz formulas (different equivalent expressions exist). These generalized expressions are used based on countless empirical interlaboratory comparison data. This approach takes into account that expected MU values are dependent on the residue level, i.e., the higher the residue concentration, the lower the anticipated MU. The Horwitz approach is expressed by the following equation:

$$RSD_R = 2^{1-0.5\log c} = 2 * c^{-0.1505}$$

with:

RSD_R = expected relative standard deviation (%)

c = concentration of the analyte (expressed as kg/kg, i.e., 0.01 mg/kg = 0.00000001 kg/kg)

Accordingly putting real figures into the above formulas concentration dependent RSD_R values are obtained, i.e.:

0.01 mg/kg ⇒ 32.0 %

0.1 mg/kg ⇒ 22.6 %

1 mg/kg ⇒ 16.0 %

These RSD_R values depending on the respective concentration levels which can be transformed into MU by multiplying with an appropriate coverage factor, normally k = 2. Advantages of this concept[12] include the incorporation of laboratory bias because laboratory variability is also randomized; deviations generated by different laboratories have been included; the Horwitz equation was found to be applicable to all concentration, methods and analytes.

[12] L. Alder et al.: Estimation of Measurement Uncertainty in Pesticide Residue Analysis. JAOAC International. Vol. 84, No 5, 2001, 1569-1577.

Drawbacks associated with the approach are that appropriate and sufficient data are needed as the basis for the estimation of a valid relation between concentration and uncertainty since the Horwitz data came from a highly diverse range of collaborative trials with concentrations from 0.05 µg/kg and 60%, involving not only pesticides. Prescribed methods were used, and PT data were not included. The resulting estimates of uncertainty accordingly are based on the distribution of between-laboratory standard deviations.

Summary

With the assumptions and prerequisites outlined for PT schemes and laboratory performance, based on top-down approaches, an estimate MU of ±50% as a generalized value could represent an acceptable and practical approximation to daily laboratory reality in pesticide residue analysis of foodstuffs.