GUIDELINES ON ESTIMATION OF UNCERTAINTY OF RESULTS

(CAC/GL 59-2006)

1. INTRODUCTION

It is a requirement under ISO/IEC 17025 that laboratories determine and make available the uncertainty associated with analytical results. To this end, food testing laboratories operating under Guidelines on Good Laboratory Practice in Pesticide Residue Analysis (CAC/GL 40-1993) should have available sufficient data derived from method validation/verification, inter-laboratory studies and in-house quality control activities, which can be applied to estimate the uncertainties particularly for the routine methods undertaken in the laboratory. These guidelines were prepared taking into account the general recommendations of the Codex Committee on Methods of Analysis and Sampling (CCMAS).

1.1 CONCEPT AND COMPONENTS OF UNCERTAINTY

Measurement uncertainty refers to the 'uncertainty' associated with data generated by a measurement process. In analytical chemistry, it generally defines the uncertainty associated with the laboratory process but may also include an uncertainty component associated with sampling.

The uncertainty 'estimate' therefore describes the range around a reported or experimental result within which the true value can be expected to lie within a defined level of probability. This is a different concept to measurement error which can be defined as the difference between an individual result and the true value. The reporting of uncertainty is intended to provide a higher level of confidence in the validity of the reported result.

Contributions to data uncertainty are manifold and described in detail in Tables 1and 2. The evaluation of uncertainty ideally requires an understanding and estimation of the contributions to the uncertainty of each of the activities involved in the measurement process.

2. IDENTIFICATION OF UNCERTAINTY SOURCES

In general, the uncertainty of measurements is comprised of many components, arising from activities involved with the sample. The uncertainty of an analytical result is influenced by three major phases of the determination:

- External operations: sampling (S_S), packing, shipping and storage of samples[1];
- Preparation of test portion: sub-sampling, sample preparation and sample processing (S_{Sp});
- Analysis (S_A): extraction, cleanup, evaporation, derivatisation, instrumental determination[2].

The combined standard (S_{Res}) and relative (CV_{Res}) uncertainty may be calculated according to the error propagation law:

$$S_{Res} = \sqrt{S_S^2 + (S_{Sp}^2 + S_A^2)} \;;\; S_{Res} = \sqrt{S_S^2 + S_L^2} \qquad (1)$$

If the whole sample is analysed, the mean residue remains the same and the equation can be written as:

$$CV_{Res} = \sqrt{CV_S^2 + CV_L^2} \text{ and } CV_L = \sqrt{CV_{Sp}^2 + CV_A^2} \qquad (2)$$

Where CV_L is the relative uncertainty of the laboratory phase of the determination which may derive from the sub-sampling, sample preparation, sample processing and analytical steps.

---

1 Packing, shipping, storage, and laboratory preparation of samples may have significant influence on the residues detected, but their contribution to the uncertainty can often not be quantified based on the current information. Examples of such errors are e.g, selection of sampling position, time of sampling, Incorrect labelling decomposition of analytes or contamination of the sample.

2 If the result has been corrected for the recovery, the uncertainty associated with this correction shall be incorporated.

Amendment 2011.

It should be noted that a laboratory is normally only required to estimate the uncertainty associated with those processes for which it has control, that is, only those processes that take place in the laboratory if sampling is not the responsibility of the laboratory staff.

2.1 ERRORS IN ANALYTICAL MEASUREMENTS

In most measurements we can distinguish between three types of errors: gross, random and systematic errors.

Gross errors refer to unintentional/unpredictable errors while generating the analytical result. Errors of this type invalidate the measurement. Laboratory quality assurance procedures should minimize gross errors. It is not possible or desirable to statistically evaluate and include the gross errors in the estimation of uncertainty. They need no further discussion in this document.

Random errors are present in all measurements, and cause replicate results to fall on either side of the mean value. The random error of a measurement cannot be compensated for, but increasing the number of observations and training of the analyst may reduce the effects.

Systematic errors occur in most experiments, but their effects are quite different. The sum of all the systematic errors in an experiment is referred to as the bias. Since they do not sum to zero over a large number of measurements, individual systematic errors cannot be detected directly by replicate analyses. The problem with systematic errors is that they may go undetected unless appropriate precautions are taken. In practice, systematic errors in an analysis can only be identified if the analytical technique is applied to a reference material, the sample is analysed by another analyst or preferably in another laboratory, or by re-analysing the sample by another analytical method. However, only if the reference material matches identically in terms of analyte, matrix, and concentration does it meet the ideal conditions for determining the bias of the method. The bias of a method may also be investigated by recovery studies. However, recovery studies assess only the effects of analysis (S_A) and do not necessarily apply to naturally incurred samples, or components of the bias that may be introduced prior to the analytical step. In pesticide analysis, results are not normally corrected for the recovery, but should be corrected if the average recovery is significantly different from 100%. If the result has been corrected for recovery, the uncertainty associated with recovery should be incorporated in the uncertainty estimation of the measurement.

Some examples of sources of errors are illustrated in Tables 1 and 2. It should be noted that not all sources mentioned have to be evaluated in the uncertainty estimation. Some sources are already incorporated in the overall uncertainty, while others are negligible and may be disregarded. However, it is important to recognise and assess all sources before elimination. Further information may be obtained from published documents[3,4].

Table 1: Sources of error in preparation of the test portion

	Sources of systematic error	Sources of random error
Sample preparation	The portion of sample to be analysed (analytical sample) may be incorrectly selected.	The analytical sample is in contact and contaminated by other portions of the sample.
		Rinsing, brushing is performed to various extent, stalks and stones may be differentially removed.

[3] EURACHEM Guide to Quantifying Uncertainty in Analytical Measurements, 2nd ed. 1999, http://www.measurementuncertainty.org.

[4] Ambrus A. Reliability of residue data, Accred. Qual. Assur. 9, pp. 288-304. 2004.

	Sources of systematic error	Sources of random error
Sample processing (S_{Sp})	Decomposition of analyte during sample processing, cross contamination of the samples.	Non homogeneity of the analyte in single units of the analytical sample.
		Non homogeneity of the analyte in the ground/chopped analytical sample.
		Variation of temperature during the homogenisation process.
		Texture (maturity) of plant materials affecting the efficiency of homogenisation process.

Table 2: Sources of error in analysis (S_A):

	Sources of systematic error	Sources of random error
Extraction / Cleanup	Incomplete recovery of analyte.	Variation in the composition (e.g. water, fat, and sugar content) of sample materials taken from a commodity.
	Interference of co-extracted materials (load of the adsorbent).	Temperature and composition of sample/solvent matrix.
Quantitative determination	Interference of co-extracted compounds.	Variation of nominal volume of devices within the permitted tolerance intervals.
	incorrect purity of analytical standard.	Precision and linearity of balances.
	Biased weight/volume measurements.	Incomplete and variable derivatisation reactions.
	Operator bias in reading analogue instruments, equipment.	Changing of laboratory-environmental conditions during analysis.
	Determination of substance which do not originate from the sample (e.g. contamination from the packing material).	Varying injection, chromatographic and detection conditions (matrix effect, system inertness, detector response, signal to noise variation etc.).
	Determination of substance differing from the residue definition.	Operator effects (lack of attention).
	Biased calibration.	Calibration.

3. PROCEDURES FOR ESTIMATING MEASUREMENT UNCERTAINTY

Whilst there are a number of options available to laboratories for the estimation of measurement uncertainty, there are two procedures described as the 'bottom up' approach and the 'top down' approach[1] that are the most commonly used.

The bottom-up method:

The bottom up or component-by-component approach incorporates an activity-based process whereby the analyst breaks down all the analytical operations into primary activities. These are then combined or grouped into common activities and an estimate made of the contribution of these activities to the combined uncertainty value of the measurement process. The bottom up approach can be very laborious and requires a detailed knowledge of the whole analytical process. The benefit to the analyst is that this approach provides a clear understanding of the analytical activities which contribute significantly to the measurement uncertainty and which therefore may be assigned as critical control points to reduce or manage measurement uncertainty in future applications of the method.

The top-down method:

The top down approach is based on method validation and long-term precision data derived from laboratory control samples, proficiency testing results, published literature data and/or inter-laboratory collaborative trials. Uncertainty estimates based on inter-laboratory studies may also take into account the between-laboratory variability of the data and provides a reliable estimate of the method performance and the uncertainty associated with its application. It is important to acknowledge however that collaborative studies are designed to evaluate the performance of a specific method and participating laboratories. They normally do not evaluate imprecision due to sample preparation or processing as the samples generally tend to be highly homogenized.

Pesticide residue analytical laboratories normally look for over 200 residues in numerous commodities that lead to practically infinite number of combinations. Therefore, it is suggested that, for estimating the uncertainty associated with multi residue procedures, laboratories use a properly selected range of analytes and sample matrices which represents the residues and commodities to be analysed in terms of physical chemical properties and composition according to the relevant parts of the Guidelines on Good Laboratory Practice in Pesticide Residue Analysis rather than establishing the uncertainty for each method/analyte/ matrix combination. The selection of a representative range of analytes and matrices to provide an uncertainty estimate should be supported by validation data and studies on the selected matrix/analyte combination.

In summary, laboratories should use either their own long-term precision data or the activity-based procedure (component by component calculation) to establish and refine the uncertainty data.

In certain situations it may also be appropriate to estimate the uncertainty contribution due to sample variability. This will require an understanding of the analyte variability within the sample lot and is not readily available to the laboratory or the analyst. The values obtained from the statistical analysis of over 8500 residue data (Table 4) provide currently the best estimate[5]. These estimates can be incorporated into the combined uncertainty value.

Likewise, it may be necessary to take into consideration the stability of analytes during sample storage and processing if these are likely to result in analyte variability between analysts and laboratories.

3.1 UNCERTAINTY ESTIMATES OF RESULTS INVOLVING ANALYSIS OF MULTI-COMPONENTS

The estimation of uncertainty of results for multi-component residues arising from the application of technical mixtures including structural and optical isomers, metabolites and other breakdown products may require a different approach particularly where the MRL has been established for the sum of all or some of the component residues. The assessment of the random and systematic errors of the results based on the measurements of multiple peaks is explained in detail in a recent publication[6].

---

5 Ambrus A and Soboleva E. Contribution of sampling to the variability of residue data, JAOAC. 87, 1368-1379, 2004.

6 Soboleva E., Ambrus A., Jarju O., Estimation of uncertainty of analytical results based on multiple peaks, J. Chromatogr. A. 1029. 2004, 161-166.

4. GUIDANCE VALUES FOR ACCEPTABLE UNCERTAINTIES

The establishment of the standard deviation of a series of tests ran by a single laboratory, as a measure of standard uncertainty, requires the results a large data-set that is not always available. However, for smaller amounts of data the true standard deviation can be estimated as follows:

Depending on the number of observations (n), the relation of the true (σ) standard deviations, calculated (S) standard deviations, and the expected range of the mean value ($\bar{x}$) at 95% probability are illustrated in Table 3. The multiplying factor, ***f***, provides the link between the estimated and true values as the function of the number of measurements.

Table 3
The values of *f* for calculation of expected ranges of standard deviation and mean values

N	$S_{min} = f_1\sigma$	$S_{max} = f_2\sigma$	$\bar{x} = \pm f_3 S$
	f_1	f_2	f_3
5	0.35	1.67	1.24
7	0.45	1.55	0.92
15	0.63	1.37	0.55
31	0.75	1.25	0.37
61	0.82	1.18	0.26
121	0.87	1.13	0.18

For instance: the repeatability of the laboratory operations, CV_L, was determined from 5 test portions drawn from a homogenised sample containing incurred residues. The average residue found was 0.75 mg/kg with a standard deviation of 0.2 mg/kg. The true residue of the processed sample can be expected between 0.75 ± 1.24*0.2 = 0.75 ± 0.248 mg/kg, while the true uncertainty of the measurement results is likely to be between 0.0696 (0.2*0.35) and 0.334 (0.2*1.67) mg/kg in 95% of the cases.

The guidance values for standard uncertainty, given in Table 4, are based on a large number of data and can be used to assess the reality of the estimated uncertainty in a laboratory in order to avoid an unreasonable high or low value.

Table 4. Typical expected uncertainties of major steps in the sampling and analysis of pesticide residues

Procedure	Relative uncertainty	Comments
Sampling of commodities of plant origin Reflects the variation of mean residues being in composite samples taken randomly from a lot. It does not incorporate the errors of follow-up procedures.	Medium and small commodities. (Sample size ≥10)[a]: 26-30%[b]. Large commodities. (Sample size ≥5)[a]: 36-40%[b].	For testing compliance with MRLs, the sampling uncertainty is defined as 0, as the MRLs refer to the average residues in bulk samples.
Sampling of animal products	The relation between the number of samples (n) to be taken for detection of a specified percentage of violation (β_p) with a given probability (β_t), is described by[a]: $1-\beta_t = (1-\beta_p)^n$.	The primary samples should be selected randomly from the whole lot.

Procedure	Relative uncertainty	Comments
Sample processing Includes the physical operation performed for homogenizing the analytical sample and sub-sampling, but excludes decomposition and evaporation of analytes.	Largely varying depending on sample matrix and equipment. No typical value can be given. The analysts should try to keep it[c] below 8-10%.	It may be influenced by the equipment used for chopping / homogenising the sample and the sample matrix, but it is independent from the analyte.
Analysis It includes all procedures performed from the point of spiking of test portions.	Within laboratory reproducibility: 16-53% for concentrations of 1 µg/kg to 1 mg/kg[c]. Average between laboratories reproducibility within 0.001-10 mg/kg: 25%[d].	The typical CV_A can be conveniently determined from the recovery studies performed with various pesticide-commodity combinations on different days and during the use of the method.

Notes:

(a) *Recommended Method of Sampling for the Determination of Pesticide Residues for Compliance with MRLs, (CAC/GL 33-1999)*;

(b) *Ambrus A. Soboleva E. Contribution of sampling to the variability of residue data, JAOAC, 87, 1368-1379, 2004*;

(c) *Guidelines on Good Laboratory Practice in Pesticide Residue Analysis (CAC/GL 40-1993)*;

(d) *Alder L., Korth W., Patey A., van der Schee and Schoeneweis S., Estimation of Measurement Uncertainty in Pesticide Residue Analysis, J. AOAC International, 84, 1569-1578, 2001*.

In addition to the estimated uncertainties made by the individual laboratories, regulatory authorities and other risk managers may decide on a default expanded uncertainty of measurements which can be used in judging compliance with MRLs (See section 5) based on between-laboratories reproducibility values. For instance, a 50% expanded uncertainty for CV_L is considered to be a reasonable default value.

5. USE OF UNCERTAINTY INFORMATION

If required, the result should be reported together with the expanded uncertainty, U, as follows:

$$\text{Result} = x \pm U \text{ (units)}$$

The expanded uncertainty, U, may be calculated from the standard combined uncertainty (S_{Res}) with a coverage factor of 2 as recommended by EURACHEM or with the Student ***t*** value for the level of confidence required (normally 95%) where the effective degree of freedom is less than 20. The respective calculations for the expanded uncertainty are as follows:

$$U = 2S_{Res} \quad \text{or} \quad U = t_{v,0.95}S_{Res} \qquad (3)$$

The numerical value of the reported results should follow the general rule that the last digits can be uncertain. Rounding the results should be done only when the final result is quoted since rounding at the initial stages of calculation may introduce unnecessary bias in the calculated values.

For the purpose of explication, it is assumed that the best estimate of the residue content is reported for a sample. How the results are interpreted depends upon the purpose of the testing. Typical reasons include testing compliance with the national MRL, certifying compliance with the Codex MRL of a commodity for export.

5.1 TESTING COMPLIANCE WITH AN MRL

Figure 1 shows how the testing results can be displayed in terms of the measured value of the residue, the corresponding uncertainty interval, and the MRL.

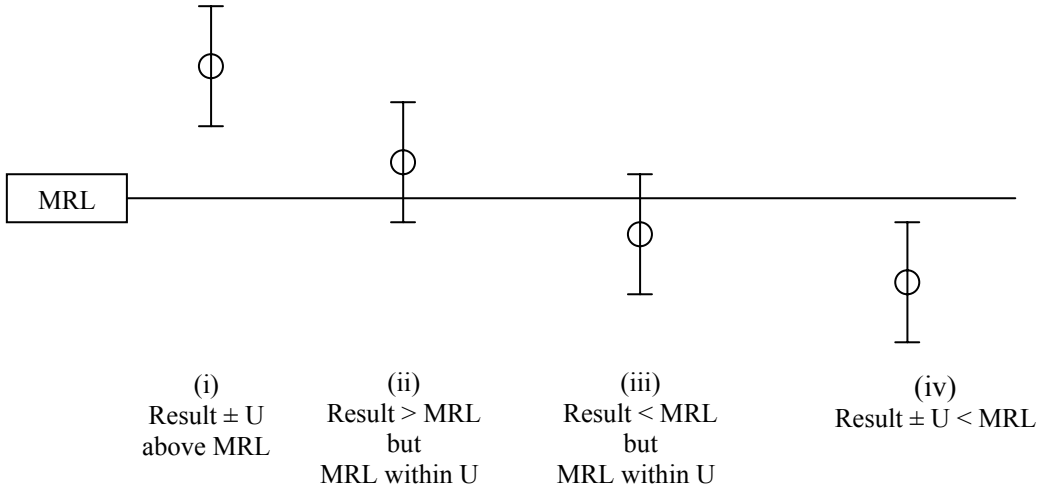

Figure 1. Illustration of the relationship of measured value expected uncertainty and MRL

Situation (i)

The analytical result bounded by the measurement uncertainty endpoints is greater than the MRL. The result indicates that the residue in the sampled lot is above the MRL.

Situation (ii)

The analytical result is greater than the MRL with the lower endpoint of the measurement uncertainty less than the MRL

Situation (iii)

The analytical result is less than the MRL with the upper endpoint of the measurement uncertainty being greater than the MRL.

Situation (iv)

The analytical result bounded by the expanded measurement uncertainty endpoints is less than the MRL.

5.2 DECISION ENVIRONMENT

The situations illustrated in Figure 1 are relevant for products of plant origin. The compliance of residues with MRLs for animal products should be decided following sampling plans based on distribution free statistics and examples given in the document on Recommended Methods of Sampling for the Determination of Pesticide Residues for Compliance with MRLs**.**

Since the residues in every sample that concurs with the minimum sample size and sample mass specified in the Codex Sampling Procedure should comply with the MRL, the expanded uncertainty should be calculated using SL from equation 1 as U = kSL.where SL = CVL* residue.

The decision-making in Situation (i) is clear. In order to avoid lengthy explanation of the uncertainty involving the performance of the analysis for testing compliance with the MRL at the national level in locally produced or imported commodities, the laboratory may report the results as the sample contains “not less than ‘x – U’ residues”. This satisfies the requirement that the MRL was exceeded beyond any reasonable doubt accounting for measurement uncertainty.

In situation (iv) the sample is clearly compliant with the MRL.

In situations (ii) and (iii) it cannot be concluded that the MRL is exceeded or compliant without reasonable doubt. Action by decision makers may need further consideration as discussed below.

The implications of situations (ii) and (iii) will depend on national practices and may have considerable impact on the acceptance of trade consignments. Caution should be exercised in distributing products in domestic markets or international trade with test results illustrated in situations (ii) and (iii). For example when certifying products for export it may not be advisable to export consignments with residue results as described in situations (ii) and (iii). For countries importing commodities with residue levels as described in situation (ii) it may be difficult to verify compliance with the MRL with an acceptable level of confidence. Situation (iii) generally may not lead to actions by the importing party.

Glossary of terms used in the text[a]

Blank (sample, reagent)	(i) Material (a sample, or a portion or extract of a sample) known not to contain detectable levels of the analyte(s) sought. Also known as a matrix blank. (ii) A complete analysis conducted using the solvents and reagents only in the absence of any sample (water may be substituted for the sample, to make the analysis realistic). Also known as a reagent blank or procedural blank.
Combined standard uncertainty	For a measurement result, y, the total uncertainty, $u_c(y)$ is an estimated standard deviation equal to the positive square root of the total variance obtained by combining all uncertainty components using the law of propagation of uncertainty (error propagation law).
Contamination	Unintended introduction of the analyte into a sample, extract, internal standard solution etc., by any route and at any stage during sampling or analysis.
Residue definition	The definition of a residue is that combination of the pesticide and its metabolites, derivatives and related compounds to which the MRL applies or which is used for dietary exposure assessment.
Determination system	Any system used to detect and determine the concentration or mass of the analyte. For example, GC-FPD, LC-MS/MS, LC with post-column derivatisation, ELISA, TLC with densitometry, or bioassay.
Level	In this document, refers to concentration (e.g. mg/kg, µg/ml) or quantity (e.g. ng, pg).
Lot	A quantity of a food material delivered at one time and known, or presumed, by the sampling officer to have uniform characteristics such as origin, producer, variety, packer, type of packing, markings, consignor, etc.
Matrix effect	An influence of one or more undetected components from the sample on the measurement of the analyte concentration or mass. The response of some determination systems (e.g. GC, LC-MS, ELISA) to certain analytes may be affected by the presence of co-extractives from the sample (matrix).
Procedural blank	See blank.
Reagent blank	See blank.
Response	The absolute or relative signal output from the detector when presented with the analyte.
Spike or spiking	Addition of analyte for the purposes of recovery determination or standard addition.
Standard uncertainty	Expressed as the standard deviation of an uncertainty component.
Unit (as part of sample)	A single fruit, vegetable, animal, cereal grain, can, etc. For example, an apple, a T-bone steak, a grain of wheat, a can of tomato soup.
Violative residue	A residue which exceeds the MRL or is unlawful for any other reason.

Note (a). The definitions given are based on the following references[7,8,9,10]. Additional definitions are given in the Guidelines on Good Laboratory Practice in Pesticide Residue Analysis.

---

7 EURACHEM (2000) EURACHEM/CITAC Guide Quantifying Uncertainty in Analytical Measurements 2nd ed. http://www.measurementuncertainty.org.

8 Recommended Method of Sampling for the Determination of Pesticide Residues for Compliance with MRLs.

9 Willetts P, Wood R (1998) Accred Qual Assur 3: 231-236.

ANNEX

Introductory notes

As noted in the Guideline document CAC/GL 59-2006, the estimation of measurement uncertainty (MU) associated with analytical data is a requirement for laboratories accredited under ISO/IEC 17025 and an expectation for all laboratories operating under Good Laboratory Practice (GLP) in pesticide residue analysis. Decisions in regard to compliance of food, whether for domestic or international standards for chemical residues and contaminants, need to take into consideration the uncertainty associated with the test results reported by laboratories for analysis of specific lots or consignments.

It is not uncommon for laboratories to report widely different estimates of MU in Proficiency Tests (PT) despite the fact that they employ very similar test methods for analysis. This evidence suggests that the estimation of MU still appears to be a developing science for a number of food laboratories. This Annex is intended to describe some of the options laboratories might employ in estimating measurement uncertainty, particularly the use of in-house method validation, quality control and long-term precision data for multi-residue pesticide methods. It is also anticipated that a more harmonised approach to the estimation of MU for pesticide residue results will minimise possible disputes in compliance decisions for residue levels near MRLs.

There are broadly two approaches commonly employed for the determination of MU; the so-called GUM (*Guide to the Expression of Uncertainty in Measurement*) or 'bottom-up' approach and the 'top-down' procedures based around application of analytical precision and bias.

The GUM approach is based on a rigorous analysis of all the individual components of an analytical process and the estimation of random and systematic errors assigned to these steps. This process, whilst initially very laborious, requires the analyst to have or develop a detailed understanding of the analytical steps on the process and identify the critical control points in the method. Unless all steps are considered in the process, it is possible to underestimate the MU. On the other hand, some operational errors may cancel out which, if ignored, could provide an overestimate of the uncertainty. It is generally acknowledged that the bottom-up approach is more suited to physical metrology than to analytical chemistry activities and, in particular, to the more complex multi-pesticide residue methods.

Proponents of the top-down approach note that laboratory data collected from in-house validation, long-term precision and analytical quality control (QC) is likely to provide more reliable information on MU. Where available, PT data can also be used to estimate MU, either as the sole basis for estimates or more often in combination with in-house data. The inter-laboratory reproducibility data from PT studies can also provide a useful 'benchmark' for single laboratory estimates.

All options should be considered in the estimation of MU. The initial aim should be to obtain the best possible estimate using the information available. Initial laboratory estimates should be verified by comparison with alternative methods, literature reports and comparisons from PT studies. Furthermore professional judgement has an important role when estimating and verifying measurement uncertainty. Estimates should be reviewed as more precision data becomes available, for example, within-batch QC data routinely generated during the course of an analytical programme.

This Annex focuses on the estimation of MU using the top-down approach, based on data obtained from different sources.

Applying a default value for MU for pesticide residues in foods

EU member states have adopted a MU 'default' value of ± 50% for pesticide residues in food consignments entering the EU. The default value is based around the statistical results of a number of EU-based PT studies involving competent residue laboratories participating in a number of multi-residue studies on fruit and vegetables. The mean relative standard deviations reported from a number of these studies have ranged between 20 to 25% providing a MU approximating to 50%.

---

[10] International Vocabulary of basic and general terms in Metrology, Geneva 1993.

In the absence of other statistical data, a laboratory testing food commodities for compliance with EU pesticide MRL regulations can adopt a default MU of 50% provided it could establish its analytical proficiency through participation in EU or similar PT studies and/or it can demonstrate acceptable long-term precision and bias associated with its test results. In the longer term however, it should be incumbent on the laboratory to verify its adoption of the default MU by independently estimating MU based on in-house precision and validation data.

Precision data derived from the use of the Horwitz relationship

In the absence of data from inter-laboratory studies on a particular method, the reproducibility standard deviation, and hence MU, may be determined from an equation reported by Horwitz which correlates reproducibility standard deviation with analyte concentration. The Horwitz relationship between coefficient of variation (CV) and analyte concentration is based on the results from a large number of food-based collaborative studies reported in the literature. The Horwitz Equation is also a helpful tool to compare in-house MU estimates against the expected value derived from published inter-laboratory studies.

Precision data derived from inter-laboratory studies (Collaborative Studies and PT Studies)

The results reported for inter-laboratory studies are subject to both imprecision and bias. If such studies involve a sufficient number of laboratories and are designed to cover real test conditions (range of analytes and matrices), the reproducibility standard deviations obtained will reflect the typical errors likely to be encountered in practice. PT study data therefore may be used to provide reasonable estimates of measurement uncertainty.

Collaborative studies on methods are generally well defined with well documented instructions on the analytical process and usually only involve expert laboratories with reputable experience in residue analysis. Under these conditions the analytical variance is likely to be the best achievable when applying the method under reproducibility conditions, particularly as error contributions from sample in-homogeneity are likely to be negligible. Providing a laboratory can demonstrate an ability to achieve the analytical performance associated with a particular collaborative study, the reproducibility standard deviation obtained for the study will be a good basis for estimating MU. A competent laboratory however, should be able to improve on the inter-laboratory method precision when conducting the method under within-laboratory reproducibility conditions, and hence reduce the MU.

If certified reference materials (CRMs) are employed in collaborative studies, the study report should provide an estimate of the bias of the method against the ‘certified’ value and this will need to be taken into consideration when estimating the MU.

In PT studies, it is normal for laboratories to employ their own test method for analysis. The method may be a standard method, a modified standard method or a method developed and validated in-house. Furthermore, there is generally greater variability in the analytical competence of the participating laboratories than is the case for collaborative studies. Because of these factors, the reproducibility standard deviation obtained for PT studies is likely to be larger than that anticipated from a method-based collaborative study. MU based on such data may be larger than the estimates reported by many participant laboratories. Nevertheless, an estimate of MU based on a PT study involving laboratories with a range of expertise using a variety of methods may be more pragmatic and useful for judging compliance of food commodities with respect to pesticide residues in international trade. The 50% default MU applied by the EU member states is based on PT data for a range of pesticides and food matrices.

Whether or not a laboratory uses PT data to estimate MU, the information from PT studies is useful to compare and verify estimates based on data such as in-house validation or quality control experiments.

MU derived from in-house validation and quality control data

There is general consensus amongst chemical metrologists that the best source of uncertainty data on the analytical process is derived from the laboratory’s method validation and/or verification studies and long-term quality control data. This is based on the assumption that the laboratory has undertaken validation and/or verification studies and has sufficient experience to have built up long-term bias and reproducibility data on suitable quality control (QC) samples, CRMs, reference materials (RMs) or matrix spikes.

The limited availability of CRMs for pesticide residues in food matrices usually requires laboratories to focus on spiked samples or other suitably characterised samples for internal quality control. The use of matrix-based QC samples such as samples with incurred residues, left-over PT study samples or spiked residue-free laboratory samples provides laboratories with a capability to monitor and control method (and analyst) performance while gathering information on both bias and precision. Control charts are excellent tools for evaluating long-term precision and monitoring statistical control of the analytical process.

Bias, where significant, and the uncertainty of bias, should be considered when estimating MU. This is illustrated in the example discussed under section 5.4.

Bias can best be determined from the use of CRMs. However given the paucity of CRMs for pesticides in food and the large number of pesticides normally incorporated into a multi-residue screen, it is generally necessary to rely on the recoveries of spiked matrix samples to provide information on method bias.

The performance of laboratories in PT studies can further provide a useful indication of the bias of individual laboratories against the consensus values and, in some instances, the spiking level of the PT samples. However, bias should be based on or confirmed by the results from a number of PT studies before it is used as an input in the estimation of MU.

Worked Examples

The following worked examples describe acceptable procedures for estimating MU based on different combinations of in-house validation data, in-house precision data and inter-laboratory data. The Horwitz Equation and results from PT studies further provide useful benchmarks for comparison with in-house MU estimates.

The following worked examples use hypothetical data for chlorpyrifos as a typical pesticide residue and draw heavily on examples presented in Eurolab Technical report No 1/2007 and the Nordtest Report TR537.

5.1 Estimating MU using the Horwitz Equation

The Horwitz Equation expresses reproducibility standard deviation as a function of analyte concentration.

$$u' = 2^{1-0.5 \log c}$$

where u' = relative reproducibility standard deviation

c = concentration of analyte (in g/g)

The relative expanded MU, U' (at 95% confidence level) may then be estimated by

$$U' = 2u'$$

Since the Horwitz Equation is a function of analyte concentration, it will provide a range of MU values depending on pesticide concentration as noted in the following table:

Concentration (mg/kg)	u′(%)	U′ (%)
1.0	16	32
0.1	22.6	45
0.01	32	64

Example 1:

A laboratory measures 0.40 mg/kg chlorpyrifos in a sample of tomato.

The Horwitz Equation predicts a relative reproducibility standard deviation of 18.4% at a concentration of 0.40 mg/kg.

$u' = 18.4\%$

$U' = 2u' = 37\%$

The laboratory would therefore report the result as 0.40 ± 0.15 mg/kg.

The laboratory report should state that the reported uncertainty was an expanded uncertainty with a coverage factor of 2 to give a level of confidence of approximately 95%. Unless stated otherwise, this is generally assumed for results reported with expanded uncertainties.

In the absence of supporting data, the Horwitz Equation should be used with some caution and only as an indicator of the likely uncertainty associated with test results. Advances in analytical methodologies, particularly instrumental techniques, have provided the capability to achieve very low limits of quantitation with much less uncertainty then predicted by the Horwitz Equation. Thompson and Lowthian have reported that laboratories tend to out-perform the Horwitz function at low concentrations. It should be noted however that the Thompson concept limits the maximum value for u′ for concentrations below 0.1 mg/kg to 22% independent of the concentration.

5.2 Estimating MU by application of the EU default value of 50%

Before applying a default MU, laboratories should ensure that they are able to routinely achieve uncertainties not greater than the default value.

Example 2:

A laboratory measures 0.40 mg/kg chlorpyrifos in a sample of tomatoes. An agreed default value of ± 50% is to be applied to the measured result.

Accordingly, the laboratory would report the result as 0.40 ± 0.20 mg/kg.

5.3 Estimating MU based on Intra-laboratory QC and data from PT Studies

5.3.1 Using the assigned (or consensus) value from PT studies

$$U' = 2u' \qquad \text{Equation 1}$$

$$u' = \sqrt{u'(Rw)^2 + u'(bias)^2} \qquad \text{Equation 2}$$

where U' = expanded relative uncertainty

u' = combined relative standard uncertainty

$u'(R_w)$ = relative standard uncertainty due to within-laboratory imprecision (relative intra-laboratory reproducibility standard deviation)

$u'(bias)$ = relative standard uncertainty component due to bias

Example 3:

In this example, $u'(R_w)$ is obtained from within-laboratory QC data, preferably long-term QC data and $u'(bias)$ is estimated from PT data.

Laboratory result for chlorpyrifos in tomato = 0.40 mg/kg.

Relative standard deviation from analysis of in-batch QC samples of tomato spiked at 0.5 mg/kg with chlorpyrifos (one spiked sample per week for previous 3 months) = 15%.

The laboratory has participated in 6 PT studies where the analytes have included chlorpyrifos in different vegetables and fruit matrices. For these studies, the relative differences between the laboratory's result and the assigned value were -15%, 5%, -2%, 7%, -20% and -12%. An average of 16 laboratories participated in each of the PT studies. The average relative reproducibility standard deviation (S_R) reported for chlorpyrifos in the six studies was 25%.

$$u'(bias) = \sqrt{RMS'_{bias}{}^2 + u'_{(C\,ref)}{}^2} \qquad \text{Equation 3}$$

where RMS'_{bias} = root mean square of relative bias value

$u'(C_{ref})$ = average relative uncertainty of the assigned values for chlorpyrifos in the six studies

$$RMS'_{bias} = \sqrt{\frac{\sum(bias)^2}{n}} \quad \text{(n = Number of PT studies)} \qquad \text{Equation 4}$$

$$= \sqrt{\frac{(-15)^2 + (5)^2 + (-2)^2 + (7)^2 + (-20)^2 + (-12)^2}{6}}$$

$$= 11.9\%$$

$$u'(C_{ref}) = \frac{S_R}{\sqrt{m}} \qquad \text{Equation 5}$$

where S_R = average relative standard deviation for chlorpyrifos from the six studies

m = average number of participants per study

$$= \frac{25}{\sqrt{16}}$$

$$= 6.3\%$$

So, $u'(bias) = \sqrt{(11.9)^2 + (6.3)^2} = 13.5\%$

From Equation 2,

$$u' = \sqrt{(15)^2 + (13.5)^2} = 20\%$$

From Equation 1, the expanded relative uncertainty (95% confidence) = 40%.

The Laboratory should report the result as 0.40 ± 0.16 mg/kg.

Notes:

1. The RMS'_{bias} value accounts for both bias and the uncertainty of bias.
2. The calculated MU is a best estimate only since the PT data is for different matrices and different concentrations of chlorpyrifos.
3. If possible, MU should be calculated based on data generated at or near the most critical concentration, for example the Codex MRL.

5.3.2 PT Studies with Certified Reference Materials (CRMs)

If a suitable CRM containing chlorpyrifos is distributed as a sample in a PT study, then there would be no need to calculate $u'(C_{ref})$ from the PT results.

In this case, $u'(C_{ref})$ would be the uncertainty stated for the certified concentration, converted to a relative standard deviation.

For example, if the 95% confidence range for the certified value for chlopyrifos in the CRM was 0.489 ± 0.031 mg/kg, then:

$u(C_{ref})$ (standard deviation) $= \frac{0.031}{2} = 0.0155$ mg/kg, and

$u'(C_{ref})$ (relative standard deviation) $= \frac{0.0155 \times 100}{0.489} = 3.17\%$

In the unlikely event that several CRMs containing chlorpyrifos were distributed in different rounds of the PT studies, then the mean $u_{(C\ ref)}$ would be used to calculate U.

In both cases, RMS'_{bias} would be calculated using Equation 4.

Example 4:

Study No.	CRM	relative bias	$u'(C_{ref})$
1	A	-12%	2.3%
2	B	-15%	1.7%
3	C	-3%	2.0%
4	C	5%	2.0%
5	C	-20%	2.0%
6	A	0%	2.3%

Mean $u'(C_{ref})$ = 2.05%

From Equation 4, RMS'_{bias} = 11.6%

From Equation 3, $u'(bias)$ = 11.8%

Note:

4. The relative uncertainty associated with CRMs is likely to be less than that associated with assigned or consensus values.

If the laboratory's relative standard uncertainty due to analytical imprecision $u'(R_W)$ remained the same i.e.,15%, then from Equations 1 and 2.

$u' = 19\%$

$U' = 38\%$

The laboratory could report the result as 0.40 ± 0.15 mg/kg.

5.4 Estimating MU using Intra – laboratory QC data

Example 5:

- Laboratory result for chlorpyrifos in tomato = 0.40 mg/kg.
- Stated purity of chlorpyrifos calibration material used to prepare the spiking solution = 95 ± 2% (certificate of analysis).
- Fourteen recoveries (%) recorded for in-batch QC samples spiked at 0.5 mg/kg chlorpyrifos over the past 3 months; 90, 100, 87, 89, 91, 79, 75, 65, 80, 82, 115, 110, 65, 73 provided a mean recovery of 86% and a relative standard deviation of 15%.

Assuming the uncertainty stated for the reference material to be an expanded uncertainty U (95% confidence range)

$$u'(C_{ref}) = \frac{2}{2} = 1\%$$

Note:

5. This assumes that the uncertainties associated with the preparation of the spiking solution and the spiking of the tomatoes are both insignificant. This is likely to be the case, but, if not, $u'(C_{ref})$ will nevertheless still be only a very minor contribution to the overall uncertainty.

$u'(R_W) = 15\%$ (relative intra-lab reproducibility standard deviation)

Using Equation 4, and taking bias to be 100 - % recovery,

$RMS'_{bias} = 20\%$

From Equation 3, $u'(bias) = 20\%$

From Equation 2, $u' = 25\%$

From Equation 1, $U' = 50\%$

The laboratory could report the result as 0.40 ± 0.20 mg/kg.

Note:

6. This uncertainty would apply to results not corrected for recovery. If, at the end of the analytical program, the results were corrected for the average recovery achieved over the 3 month period of analysis, then u′(bias) need only reflect the uncertainty associated with the mean recovery. Then u′(bias) may be calculated as the relative standard uncertainty of the recovery factor applied (the relative uncertainty of the mean recovery) combined with the relative standard uncertainty of the spike concentration, $u'(C_{ref})$.

Relative Standard Uncertainty of mean recovery,

$$u'\,\overline{\mathrm{Re}\,c} = \frac{u'(Rw)}{\sqrt{n}} \qquad \text{Equation 6}$$

where

n = the number of replicates from which the mean recovery is calculated

$$u'\,\overline{\mathrm{Re}\,c} = \frac{15}{\sqrt{14}} = 4\%$$

$$u'(bias) = \sqrt{u'(\overline{\mathrm{Re}\,c})^2 + u'(C_{ref})^2} \qquad \text{Equation 7}$$

thus $u'(bias) = \sqrt{(4)^2 + (1)^2} = 4.1\%$

Then, from Equation 2 and 1, using the $u'(R_W)$ value of 15% calculated previously

$u' = 15.5\%$ and

$U' = 31\%$

If results were corrected for recovery, the result should be reported as

0.40 ± 0.12 mg/kg

Note:

7. This example shows that if results are corrected for a mean recovery based on nine or more replicate recovery experiments conducted during the course of an analytical program, using a reference material for which the purity is known with a high level of certainty, a reasonable estimate of measurement uncertainty may be calculated from solely the intra-lab reproducibility standard deviation.