E-mail: codex@fao.org - www.codexalimentarius.org

GUIDELINES ON PERFORMANCE CRITERIA FOR METHODS OF ANALYSIS FOR THE DETERMINATION OF PESTICIDE RESIDUES IN FOOD AND FEED

CXG 90-2017

Adopted in 2017.

OBJECTIVE

1. The purpose of these guidelines is to define and describe the performance criteria, which should be met by methods to analyse pesticide residues in foods and feed (hereafter referred to as food). It addresses the characteristics/parameters to provide scientifically acceptable confidence in the analytical method that is fit for the intended use and may be used to reliably evaluate pesticide residues for either domestic monitoring and/or international trade.

2. This document is applicable to both single residue methods and multi-residue methods (MRMs) that analyse target compounds in all food commodities per the residue definition.

3. These guidelines cover qualitative and quantitative analyses, each having their own method performance criteria. Performance criteria of methods for analyte identification and confirmation are also addressed.

PRINCIPLES FOR THE SELECTION AND VALIDATION OF METHODS

A. Defining the Purpose of the Method and Scope

4. The intended purpose of the method is usually described in a statement of scope, which defines the analytes (residues), the matrices, and the concentration ranges. It also states whether the method is intended for screening, quantification, identification, and/or confirmation of results.

5. In regulatory applications, the maximum residue limit (MRL) is expressed in terms of the residue definition. Residue analytical methods should be able to measure all components of the residue definition.

6. *Fitness-for-purpose* is the extent to which the performance of a method meets the end-user's needs, and matches the criteria (data quality objectives) agreed between the laboratory and the end-user (or client) of the data, within technical and resource constraints. *Fitness-for-purpose* criteria could be based on some of the characteristics described in this document, but ultimately will be expressed in terms of acceptable combined uncertainty[1].

7. Selection of methods is based on analytes and the intended purpose of the analyses[2].

B. Supplementing other Codex Alimentarius Commission Guidelines

8. The Codex Alimentarius Commission (CAC) has issued a guideline[3] for laboratories involved in the testing of foods for import/export which recommends that such laboratories should:

   (a) use internal quality control procedures, such as those described in the "Harmonized Guidelines for Internal Quality Control in Analytical Chemistry Laboratories;"

   (b) participate in appropriate proficiency testing schemes for food analysis which conform to the requirement laid out in "The International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories (Pure Appl. Chem., vol 78, No. 1, pp.145-186, 2006);" and

   (c) whenever available, use methods which have been validated according to principles provided by the CAC.

9. The analytical methods should be used within the internationally accepted, approved, and recognized laboratory Quality Management System[4] to be consistent with the principles in the document for quality assurance (QA) and quality control (QC) referenced above.

C. Method Validation

10. The process of method validation is intended to demonstrate that a method is *fit-for-purpose*. This means that when a test is performed by a properly trained analyst using the specified equipment and materials and exactly following the method protocol, accurate, reliable, and consistent results can be obtained within specified statistical limits for sample analysis. The validation should demonstrate the identity and concentration of the analyte, taking into account for matrix effects, provide a statistical characterization of recovery results, and indicate if the frequency of false positives and negatives are acceptable. When the method is followed using suitable analytical standards, results within the established performance criteria should be obtained on the same or equivalent sample material by a trained analyst in any experienced residue testing laboratory. To ensure method performance remains appropriate over time, method validation should be continuously assessed (e.g. recovery spikes).

---

[1] Harmonized IUPAC Guidelines For Single-Laboratory Validation of Methods of Analysis, Pure & Appl. Chem., 74(5), 2002; 835 – 855

[2] OECD Guidance Document on Pesticide Residue Analytical Methods, ENV/JM/MONO (2007)17

[3] *Guidelines for the Assessment of the Competence of Testing Laboratories Involved in the Import and Export Control of Food* (CXG 27-1997)

[4] General requirements for the competence of testing and calibration laboratories, ISO/IEC 17025 (2005).

PERFORMANCE PARAMETERS FOR ANALYTICAL METHODS

11. The general requirements for the individual performance criteria of a method are summarized below[1,5]

A. Method Documentation

12. After validation, the method documentation should provide, in addition to performance criteria (data quality objectives), the following information:

    (a) Identity of the analytes included in the residue definition.

    (b) Concentration range covered by the validation;

    (c) Matrices used in the validation (representative commodity categories, e.g., similar agricultural products based on characteristics including moisture, fat, and sugar content, pH);

    (d) Protocol describing the equipment, reagents, detailed step-by-step procedure including permissible variations (e.g. “heat at 100 ± 5 °C for 30 ± 5 min”), calibration and quality procedures, special safety precautions required, and intended application and critical uncertainty requirements;

    (e) quantitative result of the expanded measurement uncertainty (MU) for the method should be calculated in the validation procedure and reported, if required.

B. Selectivity

13. Ideally, selectivity should be evaluated to demonstrate that no interferences occur which significantly affect the analysis. It is impractical to test the method against every potential interferant, but it is required that common interferences are checked by analysing a reagent (process) blank for every batch of reagents. When reagents and/or solvents are changed between batches of samples, additional reagent blank evaluations could be performed. Background levels of plasticizers, septa bleed, cleaning agents, reagent impurities, laboratory contamination, carry-over, etc. tend to show up in reagent blanks and must be recognized by the analyst when they occur. Also, analyte-to-analyte interferences must be known by checking individual analytes in mixed standard solutions. Matrix interferences are evaluated by analyses of samples known to be free of the analytes and a matrix blank is required with each batch of samples or a standard addition approach to quantification is adopted (see Section E).

14. As a general principle, selectivity should be such that interferences have no impact on method performance. The ultimate test of selectivity involves the rates of false positives and negatives in the analyses. To estimate rates of false positives and negatives during method validation, an adequate number of blanks per matrix [not from the same source] should be analysed along with spiked matrices at the analyte reporting level.

C. Calibration

15. With the exception of errors in preparation of calibration materials, calibration errors are usually a minor component of the total uncertainty, and can be safely assigned into other categories. For example, random errors resulting from calibration are part of the uncertainty, while systematic errors cause analytical bias, both of which are assessed as a whole during validation and on-going quality control. Nevertheless, there are some characteristics of calibration that are useful to know at the outset of method validation because they affect optimization of the final protocol. For example, it must be known in advance whether the calibration curve is linear or quadratic, passes through the origin, and is affected by the sample matrix or not. The described guidelines in this document relate more to validation, which may be more detailed than the calibration undertaken during routine analysis.

16. Replicate measurements are needed to provide an empirical estimate of uncertainty. The following calibration procedures are recommended for the initial method validation:

    (a) determinations at five or more concentrations should be performed (consider multiple injections per concentration);

    (b) the reference standards should be evenly spaced over the concentration range of interest and the calibration range should encompass the entire concentration range likely to be encountered;

---

5 OECD Guidance Document for Single Laboratory Validation of Quantitative Analytical Method-Guidance used in support of pre-and post-registration data requirements for plant protection and biocidal products ENV/JM/MONO(2014)20

(c) the reference standards should be dispersed over the whole sequence, or encompass the beginning and end of the run to demonstrate that calibration integrity is maintained over the entire sequence; and the fit of the calibration function must be plotted and inspected visually and/or by calculation of the residuals (differences between the actual and calculated concentrations of the standards), avoiding over-reliance on correlation coefficients. If residuals of the calibration curve deviate by more than ±20 – 30 % (30% for calibration concentrations near the instrument LOQ), statistical consideration of outliers should be made, possibly leading to re-analysis of the sequence if quality control criteria are not met.

D. Linearity

17. Linearity can be tested by examination of a plot of residuals produced by linear regression of the responses on the concentrations in an appropriate calibration set. Any curved pattern suggests a *lack of fit* due to a nonlinear calibration function. If this is the case, another function such as quadratic should be tested and applied, using at least five concentration levels. Despite its current widespread use as an indication of quality of fit, the coefficient of determination (R^2) may be misleading because it places greater significance on standards with higher concentrations. In this case, an appropriate weighting factor such as 1/x or $1/x^2$ should be considered to minimize the potential impact of the relative concentration range

18. In general, the use of weighted-linear regression or weighted-quadratic function is recommended rather than linear regression for low part per billion (µg/kg) concentration determinations. Ideally, the value of the intercept should be close to zero to reduce errors in calculating residue concentrations at low levels, although the calibration curve should not be forced through the origin without justification

E. Matrix Effects

19. Matrix-matched calibration is commonly used to compensate for matrix effects. Extracts of blank matrix, preferably of the same or similar type as the sample, should be used for calibration. An alternative practical approach to compensate for matrix effects in gas chromatographic (GC) analyses is the use of chemical components (analyte protectants) that are added to both the sample extracts and the calibration solutions in order to (ideally) maximize equally the response of pesticides in calibrants in solvent and sample extracts. Alternative ways to compensate for matrix effects involve the use of standard addition, isotopically labeled internal standards (IS), or chemical analogues. However, these approaches are often difficult in MRMs because there are too many residues in different matrices at different levels to devise routine procedures, and the lack of isotopically-labelled standards for so many analytes. Ideally, if isotopically labelled standards are available, such standards should represent the range of target compounds and recoveries should fall within the criteria for samples spiked with non-isotopically labelled standards. If solvent-only calibration is used, a measurement of matrix effects must be made to demonstrate equivalence of results by comparing responses of matrix-matched with solvent-only standards.

F. Trueness and Recovery

20. Trueness is the closeness of agreement between a test result and the accepted reference value of the property being measured. Trueness is stated quantitatively in terms of “bias,” with smaller bias indicating greater trueness. Bias is typically determined by comparing the response of the method to a certified (if available) reference material with a known value assigned to the material. Multi-laboratory testing is recommended ideally. Where the uncertainty in the reference value is not negligible, evaluation of the results should consider the reference material uncertainty as well as the statistical variability from analysing the reference material. In the absence of certified reference materials[1,5] guidelines recommend use of an available reference material that is well characterized for the purpose of the validation study.

21. Recovery refers to the proportion of analyte determined in the final result compared with the amount added (usually to a blank) sample prior to extraction, generally expressed as a percentage. Errors in measurement will lead to biased recovery figures that will deviate from the actual recovery in the final extract. Routine recovery refers to the determination(s) performed in quality control spikes in the analysis of each batch of samples.

G. Precision

22. Precision is the closeness of agreement between independent (replicate) test results obtained under stipulated conditions. It is usually specified in terms of standard deviation (SD) or relative standard deviation (RSD), also known as coefficient of variation (CV). The distinction between precision and bias depends on the level at which the analytical system is viewed. Thus, from the viewpoint of a single determination, any deviation affecting the calibration used in the analysis would be seen as a bias. From the point of view of the analyst reviewing a year’s work, the analytical bias will be different every day and should act like a random variable with an associated precision, incorporating any stipulated conditions for the estimation of this precision.

23. For single-laboratory validation, two types of precision sets of conditions are relevant: (a) repeatability, the variability of measurements within the same analytical sequence, and (b) within-laboratory reproducibility, the variability of results among multiple sets of the same sample. It is important that the precision values are representative of likely test conditions. First of all, the variation in conditions among the runs should represent what would normally happen in the laboratory during routine use of the method. This can be done by on-going method performance validation/verification. For instance, variations in reagent batches, analysts, and instruments should be measured in ongoing quality control. Secondly, the test material used should be typical, in terms of matrix and (ideally) the state of comminution, of the materials likely to be encountered in real applications.

24. In single-laboratory validations, precision often varies with analyte concentration. Typical assumptions are that: (a) there is no change in precision with analyte level, or (b) that the standard deviation is proportional to, or linearly dependent on, analyte level. In both cases, the assumption needs to be checked if the analyte level is expected to vary substantially (i.e. when analyte level approaches LOQ).

25. Precision data may be obtained for a wide variety of different sets of conditions in addition to the minimum of repeatability and between-run conditions indicated here, and it may be appropriate to acquire additional information. For example, it may be useful to the assessment of results, or for improving the measurement, to have an indication of separate operator and run effects between- or within- day, or to have an indication of the precision attainable using one or several instruments. A range of different designs and statistical analysis techniques is available, and careful experimental design is strongly recommended in all such studies. The initial validation should be conducted at the targeted limit of quantification (LOQ) or reporting limit of the method, and at least one other higher level, for example, 2-10x the targeted LOQ or the MRL.

H. Limit of Quantification (LOQ)

26. By long-standing definition among analytical chemists, the LOQ is the concentration at which the average signal/noise ratio (S/N) equals 10 in the analysis. The LOQ in practice can only be estimated because precise determination of the actual LOQ requires many analyses of spiked samples and matrix blanks but the LOQ can change day-to-day due to the performance state of the instrument, among many other factors. Some validation guidelines require that the LOQ be verified to meet method performance criteria via spiking experiments at the LOQ, however day-to-day variations in LOQ tend to force the analyst to greatly over-estimate the actual method LOQ, which can be difficult to implement the strict definition of the LOQ (S/N = 10). Thus spiking at the Lowest Validated Level (LVL) is the more descriptive and proper approach. Furthermore, quantification of analytes should not be made below the lowest validated level (LVL) in the same analytical sequence. The S/N at the lowest calibrated level (LCL) must be ≥10 (conc. ≥ LOQ), which can be set as a system suitability check required for each analytical sequence. A quality control matrix spike can also be included in each sequence to verify that the reporting limit is achieved in the analysis (an action level that is typically ≥the LCL). In essence, the point of the validation is not to determine the LOQ, but to demonstrate that the lowest reported concentration is meeting the need for the analysis. While not useful for quantification, some analysts may wish to calculate the limit of detection (LOD) (S/N = 3) to infer the presence of the analyte at concentrations too low to permit an estimate of analyte concentration.

I. Analytical Range

27. The validated range is the interval of analyte concentration within which the method can be regarded as validated. The LVL is the lowest concentration assessed during validation that meets method performance criteria. It is important to realize that the validated range is not necessarily identical to the useful range of the instrumental calibration. While the calibration may cover a wide concentration range, the validated range (which is usually more important in terms of uncertainty) will typically cover a more restricted range. In practice, most methods will be validated for at least two levels of concentration. The validated range may be taken as a reasonable extrapolation between these points of concentration, but many laboratories choose to validate at a third level to demonstrate linearity. For monitoring residue concentrations with respect to Codex standards, the analytical method must be sensitive enough so that the LVL for each analyte is at or below the current Codex maximum residue limit (CXL). The validation range should cover the existing CXL. When a CXL does not exist, the lowest level may be MRLs established by a national regulatory authority. If no CXL or MRL exists for a given analyte/matrix pair, then 0.01 mg/kg or the LOQ (whichever is greater) generally serves as the desirable LVL. In MRMs, the typical analytical goal is to set the LVL (and reporting level) at 0.01 mg/kg in diverse, yet representative commodities.

J. Ruggedness

28. The ruggedness (often synonymous with robustness) of an analytical method is the resistance to change in the results produced by the analytical method when deviations are made from the experimental conditions described in the procedure. The limits for experimental parameters should be prescribed in the method protocol (although this has not always been done in the past), and such permissible deviations, separately or in any combination, should produce no meaningful change in the results produced. A "meaningful change" here would imply that the method would not meet the data quality objectives defined by the *fitness for purpose*. The aspects of the method that are likely to affect results should be identified, and their influence on method performance evaluated by using ruggedness tests.

29. Examples of the factors that a ruggedness test could address are: small changes in the instrument,, brand/lot of reagent or changes in operator; concentration of a reagent; pH of a solution; temperature of a reaction; time allowed for completion of a process, and/or other pertinent factors.

K. Measurement Uncertainty (MU)

30. The formal approach to measurement uncertainty estimation is a calculated estimate from an equation or mathematical model, around which the true value can be expected to lie within a defined level of probability. The procedures described in method validation are designed to ensure that the equation used to *estimate the result*, with due allowance for random errors of all kinds, is a valid expression embodying all recognized and significant effects upon the result. Further considerations and description of the measurement uncertainty are provided in "Guidelines on Estimation of Uncertainty of Results"[6].

31. It is preferable to express the uncertainty of measurement as a function of concentration and compare that function with a criterion of *fitness for purpose* agreed between the laboratory and the client or end-user of the data. One possibility is to calculate MU from proficiency test data[6].

PERFORMANCE CRITERIA OF SCREENING METHODS

32. Screening methods are usually either qualitative or semi-quantitative in nature, with the objective being to discriminate samples which contain no residues above a threshold value ("negatives") from those which may contain residues above that value ("indicated positives"). The validation strategy therefore focuses on establishing a threshold concentration above which results are "potentially positive," determining a statistically based rate for false detect (positive or negatives), testing for interferences and establishing appropriate conditions of use. The screening concept offers laboratories an effective means to extend their analytical scope to analytes, which potentially have a low probability of being present in the samples. Analytes that occur more frequently should continue to be monitored using validated quantitative MRMs. As in quantitative methods, screening methods should also be checked in terms of selectivity and sensitivity. In some applications, commercial test kits may be useful, but current techniques have rarely met multi-residue screening needs economically in practice. Selectivity and analytical scope are often improved when chromatography or other form of separation is used prior to detection. Another approach is to use screening methods that involve mass spectrometry (MS)-based detection, which is able to distinguish particular chemicals from each other.

33. The selectivity of screening methods must be able to distinguish the presence of the target compound, or group of compounds, from other substances that may be present in the sample material. Selectivity of screening methods is normally less than that of a quantitative method. Screening methods can take advantage of a structural feature common to a group or class of compounds and may be based on immunoassays or spectrophotometric responses which may not unambiguously identify a compound.

34. The validation of a screening method based on a screening detection limit (SDL) can be focused on detectability. For each representative type of matrix (commodity group)[7], a minimal validation should involve analysis of at least 5 samples spiked at the estimated SDL. The samples and at least 5 matrix blanks from different sources (e.g. obtained from different markets or different agricultural fields, etc.). More replicates of greater diversity provide for a better validation. A minimum of two different samples for each type of matrix should be suitable for the intended scope of the laboratory. Additional validation data can be collected from on-going QC-data and method performance verification during routine analysis. The SDL of the qualitative screening method is the lowest level at which an analyte has been detected (not necessarily meeting the MS-identification criteria) in at least 95% of the samples (e.g. an acceptable false-negative rate of 5%).

---

[6] *Guildelines on Estimation of Uncertainty of Results* (CXG 59-2006)

[7] Table 5, *Guidelines on Good Laboratory Practice in Pesticide Residue Analysis* (CXG 40-1993)

PERFORMANCE CRITERIA OF QUANTITATIVE METHODS

35. Selectivity is of particular importance in defining the performance criteria of quantitative methods used in regulatory control programs for pesticide residues in foods. Ideally, the method needs to provide a signal response that is free from interferences from other analytes and matrix compounds that may be present in a sample or sample extract. Chromatographic analyses based on peaks, which are not fully resolved, provide less reliable quantitative results. Use of element-specific detectors or different detection wavelengths or MS-based detectors which are better able to distinguish a particular compound or structure, combined with chromatographic separation, improves the selectivity of quantitative methods.

36. The requirement to recover a range of different pesticide residues in one extraction increases the potential for compromised selectivity in MRMs compared to single residue methods. Using less selective extraction and clean-up procedures is likely to result in greater co-extracted matrix material in the final extract. The nature and quantities of such co-extracted material can vary markedly based on the matrix, method, and analytes of interest. Care is therefore required when setting criteria for the precision and trueness of MRMs to ensure that quantification will not be affected by chemical interferences.

37. In addition to the selectivity of a method, the ability of the method to provide a reliable quantitative result must be demonstrated (i.e. trueness - see section F and precision – see section G). Ideally, the relative standard deviation between the original sample and replicates will be less than 20 percent.

38. Acceptability criteria for a quantitative analytical method should be demonstrated at both initial and on-going validation stages, as being capable of providing acceptable mean recovery values at each spiking level. For validation, it is recommended that a minimum of 5 replicates be analysed (to check the recovery and precision) at the targeted LVL, LOQ, or reporting limit of the method, and at least one additional higher level, for example, 2-10x the LVL or the MRL. If a method is being used for compliance testing (i.e. if a commodity is compliant with an established MRL) the MRL (or CXL) should fall within the validated concentration range. When the residue definition includes two or more analytes, the method should be validated for all analytes.

39. The trueness of a method may be determined by analysis of a certified reference material, by comparison of results with those obtained using another method for which the performance criteria have previously been rigorously established (typically a collaboratively studied method), or by determination of the recovery of analyte fortified into known blank sample material. Acceptable mean recoveries for enforcement purposes should normally range from 70-120% with a RSD ≤20%. For very low concentrations (e.g. <0.01 mg/kg) some laboratories may accept method performance criteria that fall outside of these criteria (e.g. 60 – 120% with a RSD <30%). In certain cases (typically with MRMs), recoveries outside this range may be acceptable, such as when recovery is lower but consistent (e.g. demonstrating good precision). This is more justifiable if the reason for the systematic low bias is well established by chemistry (e.g. known analyte distribution between phases in a partitioning step). However, a more accurate method should be used, if practicable. Recoveries >120% are likely to be attributable to a positive interference or bias that should be investigated.

40. Analysis of incurred matrix to support method validation is encouraged. For interpreting recoveries, it is necessary to recognize that analyte spiked into a test sample may not behave in the same manner as the biologically incurred analyte (pesticide residue). In many situations, the amount of an extracted incurred residue is less than the total incurred residues actually present. This may be due to losses during extraction, intra-cellular binding of residues, the presence of conjugates, or other factors that are not fully represented by recovery experiments using analyte-fortified blank matrices. Often radio-labelled incurred residues or standard reference materials are required to assess recoveries of incurred residues.

41. At relatively high concentrations, analytical recoveries are expected to approach one hundred percent. At lower concentrations, particularly with methods involving extensive extraction, isolation, and concentration steps, recoveries may be lower than at higher concentrations. Regardless of what average recoveries are observed, recovery with low variability is desirable so that a reliable correction for recovery can be made to the final result, when required.

42. In general, residue data do not have to be adjusted for recovery when the mean recovery is within the range of 70-120%. Recovery corrections should be made consistent with the guidelines provided by the CXG 37-2001[8]. This will facilitate direct comparison of data sets. Correcting functions should be established on the basis of appropriate statistical considerations and documented, archived and made available to clients and reviewers. Data should (a) be clearly identified as to whether or not a recovery correction has been applied and (b) if applicable, include the amount of the correction and the method by which it was derived. This will promote direct comparability of data sets. Correction functions should be established on the basis of appropriate statistical considerations, and documented, archived and made available to the client.

[8] Harmonized IUPAC Guidelines for the use of Recovery Information in Analytical Measurement. Pure & Appl. Chem., 71,1999; 337 – 348. CXG 37-2001

43. In accordance with ISO IEC17025[4], participation in a proficiency testing program should be done. Many proficiency testing schemes are available for laboratories worldwide that conduct pesticide residue monitoring. Inter-laboratory testing may also be performed.

PERFORMANCE CRITERIA OF METHODS FOR ANALYTE IDENTIFICATION AND CONFIRMATION

44. By far, gross errors (spurious mistakes made during sample preparation) are the greatest source of misidentifications in MS-based methods. For this reason, all regulatory enforcement actions (above an MRL or for those with no MRL on that commodity) require confirmation of the result via re-extraction of a replicate test portion of the original sample and re-analysis, ideally using different sample preparation and/or analysis.

45. Selectivity is the primary consideration for methods of identification. The method should be sufficiently selective to provide unambiguous identification. MS coupled to a chromatographic separation method is a very powerful combination for identification of an analyte in the sample extract. This method provides information about the structure of the analyte that is not obtainable with chromatography alone. GC-MS and LC-MS tools (full-scan, selected ion mode, high-resolution, tandem MS/MS, hybrid systems, among other advanced techniques) provide many measurable parameters, such as retention times, chromatographic peak shapes, ion intensities and relative abundances/ratios, mass accuracies, and other useful aspects to help make analyte identifications. However, successful methods can be developed and applied using non-MS based techniques (e.g. HPLC with photo-diode array detection, GC with element selective detection), especially if confirmation of the test result is done with alternative column chemistries.[9]

A. MS-Based Identification

46. There are no universally accepted criteria for identification. Table 1 gives examples of criteria.

47. Current practices in qualitative and quantitative analysis of pesticide residues commonly involve chromatography + selected ion monitoring (SIM) or MS/MS techniques. Full-spectral MS is also an acceptable tool that uses spectral library matching factors and/or relative abundances of major ions within the full spectra. The latter case can be treated as ion ratios in the criteria given below using at least 3 ions. In the former case, matching factors should be used for regulatory identification purposes, and the library reference spectra should be obtained from background-subtracted high purity standards on the same instrument using the same conditions as in the sample analysis. The following identification criteria should be met:

(a) Analyte retention time reference values should be determined from contemporaneously analysed (within the same batch) high concentration matrix-matched calibration standards. Otherwise, if it is known that no interferences are present, solvent-based standard solutions can be used

(b) Ion ratio reference values are to be set in the same way as in paragraph 47 a. The different ions used for identification must co-elute and have similar peak shapes. The ion from the calibration standard with the higher average intensity is to be used as the denominator in the ion ratio, expressed in percentage (due to signal fluctuations, matrix effects, etc.deviations of ion ratios up to 30% are acceptable).

(c) The signal to noise ratios for measured peaks must be greater than 3 and/or the signal should exceed the threshold intensity level as compared to the signal of a suitable calibration standard or control encompassing the level of interest.

(d) The ion transitions chosen for identification purposes should make chemical/structural sense (be sure that the ions chosen do not originate from a degradant, impurity, or confusion with a different chemical than the analyte).

(e) All measured reagent and matrix blank samples should be free of carry-over, contamination, and/or interferences with a response >20% of the LOQ. For matrix blank samples, 30% of LOQ may be acceptable.

(f) For MS analyses, it is preferable to monitor ions with a mass/charge ratio greater than 100.

48. The minimum acceptable retention time for the analyte(s) should be at least twice the retention time corresponding to the void (dead) volume of the column. The retention time of the analyte in the extract should correspond to that of the reference value (47a) within ±0.2 min or 0.2% relative retention time, for both gas and liquid chromatography (preferably ± 0.1 min if possible).

[9] *Guidelines on Good Laboratory Practice in Pesticide Residue Analysis* (CXG 40-1993)

49. Methods based on high-resolution mass spectrometry are considered to provide improved reliability through accurate measurement of the mass/charge of the ion than cannot otherwise be obtained using unit-resolution mass spectrometry techniques. Different types and models of mass spectrometric detectors provide different degrees of selectivity, which relates to the confidence in identification. The example criteria for identification provided in Table 1.should only be regarded as guidelines for identification, not as absolute criteria to prove presence or absence of a compound.

B. Confirmation

50. If the initial analysis does not provide unambiguous identification or does not meet the requirements for quantitative analysis, a confirmatory analysis is required. This may involve re-analysis of the extract or the sample. When a CXL/MRL is exceeded, a confirmatory analysis of another portion of the sample is required. For unusual pesticide/matrix combinations, a confirmatory analysis is also recommended.

51. If the initial confirmatory method is not based on an MS technique, the confirmatory methods should involve MS-based analyte identification. Moreover, the confirmatory methods should use an independent approach based on different chemical mechanisms (such as LC and GC separations). In some situations, confirmation by independent laboratories may be appropriate. Examples of analytical techniques that may be suitable to meet criteria for confirmatory analytical methods are summarized in Table 2.

Table 1. Identification criteria for different MS techniques

<table>
<tr><th rowspan="2">MS detector / characteristics</th><th rowspan="2">Typical systems (examples)</th><th rowspan="2">Acquisition</th><th colspan="2">Requirements for identification</th></tr>
<tr><th>minimum number of ions</th><th>other</th></tr>
<tr><td>Unit mass resolution</td><td>quadrupole,
ion trap, TOF</td><td>full scan, limited m/z range, SIM</td><td>3 ions</td><td rowspan="4">S/N ≥ 3[e]

Analyte peaks in the extracted ion chromatograms must fully overlap.

Ion ratio within
±30% (relative)
of average
of calibration standards from same sequence[f]</td></tr>
<tr><td>MS/MS</td><td>triple quadrupole, ion trap, Q-trap, Q-TOF, Q-Orbitrap</td><td>selected or multiple reaction monitoring, mass resolution for precursor-ion isolation equal to or better than unit mass resolution</td><td>2 product ions</td></tr>
<tr><td rowspan="2">Accurate mass
measurement</td><td rowspan="2">High resolution MS:
TOF or Q-TOF
Orbitrap or Q-Orbitrap
FT-ICR-MS
sector MS</td><td>full scan, limited m/z range, SIM,
fragmentation with or without precursor-ion selection, or combinations thereof</td><td>2 ions with
mass accuracy
≤ 5 ppm[a,b,c]</td></tr>
<tr><td>combined single stage MS and
MS/MS with mass resolution for precursor-ion isolation equal to or better than unit mass resolution</td><td>2 ions:
1 molecular ion, (de)protonated molecule or adduct ion with mass acc. ≤ 5 ppm[a,c]
plus
1 MS/MS product ion[d]</td></tr>
</table>

a) preferably including the molecular ion, (de)protonated molecule or adduct ion

b) including at least one fragment ion

c) < 1 mDa for m/z < 200

d) ≤5 ppm

e) in case noise is absent, a signal should be present in at least 5 subsequent scans

f) if the mass accuracy of a precursor and its product ion is ≤ 5 ppm, ion ratio tolerance is optional.

Table 2. Examples of detection methods suitable for the confirmatory analysis of substances

Detection method	Criterion
LC or GC and MS	If sufficient number of fragment ions are monitored
LC-DAD	If the UV spectrum is characteristic
LC – fluorescence	In combination with other techniques
2-D TLC – (spectrophotometry)	In combination with other techniques
GC-ECD, NPD, FPD	Only if combined with two or more separation techniques
LC-immunoaffinity	In combination with other techniques
LC-UV/VIS (single wavelength)	In combination with other techniques

ANNEX

DEFINITIONS

Analyte: The chemical substance sought or determined in a sample *Guidelines on Analytical Terminology* (CXG 72-2009).

Analyte protectant: Compounds that strongly interacts to fill active sites in the gas chromatographic system, thereby reducing the analyte interactions with those active sites and yielding less peak tailing or losses, thus a higher analyte response.

Applicability: The analytes, matrixes, and concentrations for which an analytical method can be used satisfactorily *Guidelines on Analytical Terminology* (CXG 72-2009).

Coefficient of Variation (CV): Often referred to as the Relative Standard Deviation (RSD). This is a measure of precision in quantitative studies comparing the variability of sets with different means.

Confirmation: The combination of two or more analyses that are in agreement with each other, at least one of which meets identification criteria.

Confirmatory method: A method that is capable of providing complementary information in agreement with a previous result. Ideally, a different subsample is analysed with a method involving a different chemical mechanism than in the first analysis, and one of the methods meets analyte identification criteria with an acceptable degree of certainty at the level of interest.

Degradate (degradant, degradation product): Component of a pesticide residue occurring in a commodity as a result of abiotic transformation of the pesticide (e.g. heat, light, moisture, pH, etc.)

False positive: A result wrongly indicating that the analyte is present or exceeds a specified concentration (e.g. CXL/MRL or reporting level).

False negative: A result wrongly indicating that the analyte is not present or does not exceed a specified concentration (e.g. CXL/MRL or reporting level).

Fortification: Addition of analytes for the purposes of determining the recovery (also known as spiking).

Identification: Process of unambiguously determining the chemical identity of all or any components of the residue definition.

Incurred residue: Residue occurring in a commodity resulting from specific use of a pesticide or from consumption by an animal or environmental contamination in the field, as opposed to residues present due to laboratory fortification of samples.

Interference: Intrinsic or extrinsic response unrelated to an analyte (e.g. noise) due to electronic, chemical, or other factors related to the instrumentation, environment, method, or sample.

Interferent: A chemical or other factor causing an interference

Internal standard (IS): A chemical added at a known amount to samples and/or standards in a chemical analysis, including the blank and calibration standards. This substance can then be used for calibration by plotting the ratio of the analyte signal to the internal standard signal as a function of the concentrations. This ratio for the samples is then used to obtain the analyte concentrations. The internal standard used needs to provide a signal that is similar to the analyte signal in most ways but sufficiently different so that the two signals are readily distinguishable from each other.

Limit of Detection (LOD): The lowest concentration or mass of the analyte that can be detected (but not quantified) in a sample. In practice, this is typically the analyte concentration at which the average signal/noise is 3.

Limit of quantification (LOQ): The smallest concentration of the analyte that can be quantified. It is commonly defined as the minimum concentration of the analyte in the test sample that can be determined with acceptable precision (repeatability) and accuracy under the stated conditions of the test. For the scope of this document, this is typically the analyte concentration at which the average signal/noise is 10. [See also paragraph 26].

Linearity: The ability of a method of analysis, within a certain range, to provide an instrumental response or results, proportional to the quantity of analyte to be determined in the laboratory sample *Guidelines on Analytical Terminology* (CXG 72-2009).

Lowest Calibrated Level (LCL): The lowest concentration (or mass), which the determination system is successfully calibrated, through the analysis batch.

Lowest Validated Level (LVL): The lowest validated spiking level meeting the method performance criteria.

Matrix: The material or component (e.g. the food) that is sampled for pesticide residue studies.

Matrix blank: Sample material or sample portion containing no detectable concentration of the analytes of interest.

Matrix effect: An influence of the one or more undetected components from the sample on the measurement of the analyte concentration or mass.

Matrix-matched standards: Standard solutions prepared in final extracts of matrix blanks similar to that of the sample

Metabolite: Component of a pesticide residue occurring in a commodity as a result of biotic transformation (metabolism) of a pesticide in a biological system (e.g. plant, animal).

Multiresidue method (MRM): A method which can determine a large number of compounds typically from different chemical classes

Precision: Degree of variability of a measurement around a mean.

Quantitative method: A method capable of producing analyte concentration (determinative) results with trueness and precision that comply with established criteria.

Recovery: Amount measured as a percentage of the amount of analyte(s) (as per residue definition) originally added to a sample of the appropriate matrix, which contains either no detectable level of the analyte or a known detectable level. Recovery experiments provide information on both precision and trueness and thereby the accuracy of the method.

Relative Standard Deviation (RSD): The standard deviation, divided by the absolute value of the arithmetic mean, expressed in percentage. It refers to the precision of the method (also known as coefficient of variation-CV).

Repeatability: Precision usually expressed as RSD, obtained from the same measurement procedure or test procedure; the same operator; the same measuring or test equipment used under the same conditions; the same location and repetition over a short period of time *Guidelines on Analytical Terminology* (CXG 72-2009).

Reproducibility: Precision (typically expressed as RSD) from observation conditions where independent test/measurements results are obtained with the same method on identical test/measurement items in different test or measurement facilities with different operators using different equipment *Guidelines on Analytical Terminology* (CXG 72-2009).

Residue Definition: the spectrum of compounds to be analysed which may include the parent compound, metabolites, isomers, reaction products and/or degradants. The residue definition is typically determined by a regulatory body.

Ruggedness: A measure of the capacity of an analytical procedure to remain unaffected by small but deliberate various in method parameters and provides an indication of its reliability during normal usage *Guidelines on Analytical Terminology* (CXG 72-2009).

Sample preparation: Involves the extraction of a test portion of the sample, its clean-up and other steps that lead to the sample solution for analysis.

Screening Detection Limit (SDL): Lowest level of fortification that has been shown to have certainty at a 95% confidence level.

Screening Method: A method that meets predetermined criteria to detect the presence, or absence, of an analyte or class of analytes, at or above the minimum concentration of interest.

Selectivity: The extent to which a method can determine particular analyte(s) in a mixture(s) or matrices(s) without interferences from other components of similar behaviour *Guidelines on Analytical Terminology* (CXG 72-2009).

Sensitivity: Quotient of the change in the indication of a measuring system and the corresponding change in the value of the quantity being measured *Guidelines on Analytical Terminology* (CXG 72-2009).

SIM: selected ion monitoring, a mass spectrometry detection technique

Single Residue Method: A method which determines a single analyte or a small group of analytes with similar physico-chemical properties.

Standard addition: The method of standard addition is a type of quantitative analysis approach sometimes used in analytical chemistry whereby a known quantity of analyte is added directly to the aliquots of final extracts.

TOF: Time of flight, a detection methodology used in mass spectrometry.

Trueness: The closeness of agreement between the average of an infinite number of replicate measured quantity value and a reference quantity value *Guidelines on Analytical Terminology* (CXG 72-2009).

Uncertainty: A parameter associated with the result of a measurement that characterizes the dispersion of values that could reasonably be attributed to the measurement.