11th Edition: Released on 30 December 2024

GENERAL PROTOCOL

for EU Proficiency Testings on Pesticide Residues in Food and Feed

Introduction

This protocol contains general procedures valid for all European Union Proficiency Testings (EUPTs) organised on behalf of the European Commission, DG-SANTE[1] by the four European Union Reference Laboratories (EURLs) responsible for the area of pesticide residues analysis in food and feed. These EUPTs are organised for laboratories belonging to the Network[2] of National Reference Laboratories (NRLs) and Official Laboratories (OfLs) of the EU Member States. OfLs from EFTA countries and EU-Candidate countries are also welcome to participate in the EUPTs. OfLs from Third countries may be permitted to participate on a case-by-case basis.

The following four EURLs for pesticide residues were appointed by DG-SANTE based on the official controls Regulation (EU) No. 2017/625[3]:

- EURL for Fruits and Vegetables (EURL-FV),
- EURL for Cereals and Feedingstuff (EURL-CF),
- EURL for food of Animal Origin and commodities with high fat content (EURL-AO) and
- EURL for pesticides requiring Single Residue Methods (EURL-SRM).

The aim of these EUPTs is to obtain information regarding the quality, accuracy and comparability of pesticide residue data in food and feed reported to the European Union within the framework of the national control programmes and the EU multiannual co-ordinated control programme[4]. Participating laboratories will be provided with an assessment of their analytical performance that they can use to demonstrate their (ongoing) analytical proficiency and compare themselves with other participating laboratories. By pointing out areas of analytical deficiencies, EUPTs contribute to the continuous improvement of the analytical quality of OfLs, thus helping to increase the confidence on the results generated by them.

---

1 DG-SANTE = European Commission, Health and Food Safety Directorate-General

2 For more information about the EURL/NRL/OfL-Network please refer to the EURL-Web-portal under:"http://www.eurl-pesticides.eu"

3 Regulation (EU) 2017/625 of the European Parliament and of the Council on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products.. Published at OJ of the EU L95 of 07.04.2017

4 European Commission Proficiency Testings for Pesticide Residues in Fruits and Vegetables, Trends in Analytical Chemistry, 2010, 29 (1), 70 – 83.

Table of Content

EUPT- Organisers and Scientific Committee

EUPTs are organised either by single EURLs, or collaboratively by more than one EURL.

An **organising team** (in the following named **organisers**[5]) is appointed by the EURL(s) in charge of a given PT. The organisers are in charge of all administrative and technical PT activities of a proficiency testing (PT) round. These tasks include the PT-announcement, the production of the proficiency testing item (PT-item), the undertaking of homogeneity and stability assessments, the portioning, packing and shipment of the PT-Items, the handling and evaluation of the results and method information submitted by the participants, the drafting of the preliminary and final reports as well as the generation and distribution of EUPT-participation certificates.

To complement the internal expertise of the EURLs, a group of external consultants forming the **EUPT-Scientific Committee** (EUPT-SC)[6] has been established and approved by DG-SANTE. The EUPT-SC consists of expert scientists with many years of experience in PTs and/or pesticide residue analysis. The latest **composition of the EUPT-SC** and the affiliation of each of its members is shown on the EURL-Website. The members of the EUPT-SC are also listed in the Specific Protocol and the Final Report of each EUPT.

The EUPT-SC is made up of the following two subgroups:

a) An independent **Quality Control Group** (EUPT-QCG) and
b) An **Advisory Group** (EUPT-AG).

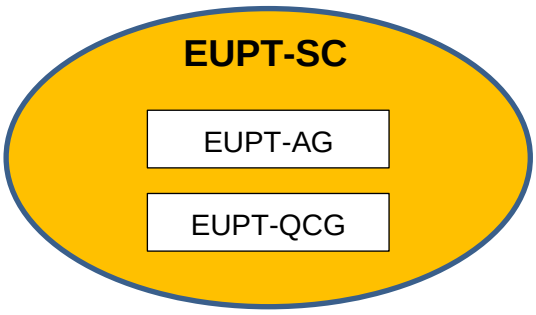

Figure 1: *Composition of EUPT-Scientific Committee*

The EUPT-SC's role is to assist the organisers during the planning and the data evaluation phase of a PT-round. Input from the EUPT-SC is requested, when it comes to e.g. selecting the commodities for the EUPTs of the following season, selecting the analytes to be included in the Target Pesticides List (p. 8), establishing the Minimum Required Reporting Levels (MRRLs) for each of the analytes,

[5] The term organisers is to be considered equivalent to the term PT-provider in ISO 17043:2023

[6] Link to the List of current members of the EUPT Scientific Committee: http://www.eurl-pesticides.eu/library/docs/allcrl/EUPT-SC.pdf

and statistically evaluating the participants' results (in anonymous form). The EUPT-SC is furthermore consulted when it comes to drafting and updating documents, such as the General and Specific PT Protocols and the Final EUPT-Reports.

The EUPT-QCG has the additional function of supervising the quality of EUPTs and of assisting the EURLs in confidential aspects such as the choice of the analytes to be present in the PT item and the approximate concentrations at which they should be present.

The EUPT-SC typically meets once a year, after all EUPTs of the season have been conducted and preliminarily evaluated by the four pesticide EURLs. The aim of these meetings is to discuss the preliminary evaluation of the EUPT-results, especially where case-by-case decisions are needed. PT plans for the next EUPT season and, if needed, possible changes in the EUPT-General Protocol are also discussed during these meetings. The main topics and decisions on these meetings are documented.

The present EUPT General Protocol (EUPT-GP) was drafted by the EURLs and reviewed by the EUPT-SC. Follow the link to access a website giving an **overview of EUPT-GP versions**. The latest version of the EUPT-GP is highlighted.

EUPT Participants – Eligibility and Obligation for Participation

Within the European Union, all NRLs operating in the same area as the organising EURL, as well as all OfLs whose scope overlaps with that of the EUPT, are legally obliged to participate in EUPTs. The legal obligation of NRLs and OfLs to participate in EUPTs arises from:

- Art 38 (2) of Regulation (EU) No. 2017/625[3]
- Art. 28 (3) of Reg. (EC) No. 2005/396 (for all OfLs analysing for pesticide residues within the framework of official controls of food or feed[7])
- Art. 101 (1)(a) of Regulation (EU) No. 2017/625[3] (for all NRLs).

Every year, shortly before launching the registration period of the first of the four EUPTs in a given EUPT-Season, all OfLs and NRLs are asked to update their routine scope of commodities as well their contact information within the EURL-DataPool. Based on this information the OfLs are classified into those that are obliged and those that are eligible to participate in each of the EUPTs to be conducted within a given year.

[7] Official controls in the sense of Regulation (EU) 2017/625. This includes labs involved in controls within the framework of national and/or EU programs, as well as labs involved in import controls according to Regulation (EU) 2019/1793 (which repealed Regulation (EC) No. 2009/669).

NRLs are responsible for checking whether all relevant OfLs within their network are included in the list of obliged laboratories with their current commodity-scopes and contact information.

OfLs are furthermore urged to keep their own profiles within the EURL-DataPool up-to-date, especially their commodity and pesticide scopes and their contact information.

Labs that are obliged to participate in a given EUPT, but are not able to participate, must provide the reasons for their non-participation. This also applies to any participating laboratories failing to report results.

EUPTs are furthermore open to the following laboratories as long as sufficient material is available:

a) any other OfLs from EU countries that are not covered by the above obligations to participate;

b) NRLs and OfLs from EU-candidate countries and EFTA countries;

c) other laboratories from EU or EFTA countries analysing official organic samples within the frame of Reg. 889/2008/EC;

d) governmental laboratories from Third Countries (countries outside EU)

e) other laboratories from Third Countries as long as they are involved in controls of products destined for export to the EU.

Note on a): Laboratories having been designated as OfLs, according to Art. 37(2)(b) of Regulation (EU) No. 2017/625[3] by a Competent Authority of an EU Member State (MS1) will normally also need to be commissioned with OfL activities in a different EU Member State (MS2) for being eligible for participation. Scan-copies of documents giving information about the period and scope of these OfL activities for MS2 may be requested by the EUPT organizers. The responsible NRL and/or Competent Authority of MS2 may be contacted before deciding whether the laboratory in question is eligible or even obliged to participate in a certain PT. A laboratory whose OfL-appointment in the area of pesticide residue analysis has ceased, will normally loose its eligibility (and obligation) to participate in EUPTs, but participation may be allowed if the responsible NRL and/or Competent Authority of MS1 or MS2 considers its participation essential for judging the proficiency in view of a planned or potential OfL activity in the future.

Laboratories of groups c) and e) will be requested to provide a proof of their function (e.g. scan copy of a document stating official appointment).

Obligation of OfLs and NRLs to double-check Status of EUPT-Participation:

Based on the latest information within the DataPool and considering the selected commodities of the upcoming EUPTs, the OfLs (including the NRLs) are grouped into those for which participation in a given EUPT is **obligatory** and those for which participation is **voluntary** ("OV-grouping").

Upon accessing the EUPT Registration Form within the EURL-DataPool, laboratories can choose the EUPTs they would like to participate and view their OV-grouping status for each of the selected PTs. If a laboratory does not agree with its OV-Grouping, it should promptly contact the corresponding NRL and the EUPT-organisers and give the reasons why it believes it should be grouped differently. The reasons provided by the laboratories will be noted by the organisers and if indicated, the DataPool will be updated accordingly. In any case, the OV-grouping prepared by the EURLs is indicative only, as the real obligation to participate in a given EUPT arises from the above-mentioned EU-regulations, not the DataPool entries or any lab's claims. Additional requirements may arise from accreditation bodies or local rules and regulations.

Within the DataPool, NRLs have the possibility to view data relevant to OfLs within their network (OV- grouping, registration progress) and are responsible for checking whether the OV- grouping of all relevant OfLs within their network is correct.

OfLs that are obliged but not able to participate in a given EUPT must provide the reasons for their non-participation. This also applies to any participating laboratories that fail to submit PT-results.

Participation fee and Invoicing

By completing the registration for participation in a given EUPT, a laboratory agrees to proceed with a timely payment of the participation fee after being accepted for participation and after the invoice issued by the organiser is received. The invoice fee covers the costs of production, handling and delivery of the PT-materials. The organisers will issue digital invoices in PDF format only, and without any electronic signature. By registering to an EUPT the laboratories also accept that the pdf invoice, issued by the organisers and sent via e-mail to the participant, is sufficient for triggering the payment of the participation fee. The EURLs retain the right to decline any request for supplementary forms or additional paperwork in connection to the payment. The laboratories should note that additional costs may incur if such extra services are requested, depending on the incurring extra workload. Extra costs may also incur if new modified invoice is requested, e.g. because of missing or erroneous information caused by errors or omissions by the registered laboratory during registration. OfLs not paying the EUPT participation fee will be initially reminded, and then warned that information concerning their laboratory may be blacked out in the final report of the concerned EUPT and the certificate of participation may not be issued to them, and that their participation in subsequent

EUPTs could be denied. In case of a repetitive non-payment, the EUPT organisers may inform the corresponding NRL and/or the competent authority responsible for the OfL.

Confidentiality and Communication

The proprietor of all EUPT data is DG-SANTE and as such has access to all information.

For each EUPT, the laboratories are given a unique code (lab code), initially only known to themselves and the organisers. In the final EUPT-Report, the names of participating laboratories will not be linked to their laboratory codes. It should be noted, however, that the organisers, at the request by DG-SANTE, may present the EUPT-results on a country-by-country basis. It may therefore be possible that a link between codes and laboratories could be made, especially for those countries where only one laboratory has participated. Furthermore, the EURLs reserve the right to share EUPT results and codes amongst themselves: for example, for the purpose of evaluating overall laboratory or country performance as requested by DG-SANTE.

As laid down in Regulation (EU) No. 2017/625[3], NRLs are responsible for evaluating and improving their own OfL-Network. On request from the NRLs, the EURLs will provide them with the PT-codes of the participating OfLs belonging to their OfL-Network. This will allow NRLs to follow the participation and performance of the laboratories within their network.

Communication between participating laboratories during the test, on matters concerning a PT exercise, is not permitted from the start of the PT exercise until the preliminary report distribution.

For each EUPT the organising EURL prepares a specific EUPT-Website where all PT-relevant documents in their latest version are linked. In case of important modifications on any of these documents, the participating laboratories will be informed via e-mail. In any case, as soon as the PT-period starts the participants are encouraged to visit the particular EUPT-Website, to make sure that they are using the latest versions of all PT-relevant documents.

The official language used in all EUPTs is English.

Announcement / Invitation Letter

Approximately 3 months before the distribution of the PT items to the participants the EURLs will publish an Announcement/Invitation letter on the EURL-web-portal and distribute it via e-mail to the NRL/OfL mailing list available to the EURLs. This letter will inform about the commodity to be used for preparing the PT item, as well as links to the tentative EUPT-Target Pesticides List and the tentative EUPT-Calendar.

Target Pesticides List and PT-Residue Definitions

The Target Pesticides List contains all analytes (pesticides and metabolites) to be sought for, along with the Minimum Required Reporting Levels (MRRLs) valid for the specific EUPT. The MRRLs are typically based upon the lowest MRLs found either in Regulation (EC) No. 2005/396 or in Regulation (EU) No. 2016/128 (Baby Food Directive).

The residue definition in an EUPT may differ from the legal one if this is deemed necessary by the organisers for ensuring a better evaluation of the results. Participants must express their results as defined in the Target Pesticides List of the respective EUPT. Separately quantifiable analytes are typically listed separately unless stated otherwise.

Specific Protocol

For each EUPT, the organising EURL will publish a Specific Protocol at least 2 weeks before the PT item is distributed to the participating laboratories. The Specific Protocol will contain all the information previously included in the Invitation Letter but in its final version, information on payment and delivery, instructions on how to handle the PT item upon receipt and on how to submit results, as well as other relevant information.

Assessing the Homogeneity of the PT Item

A suitable homogeneity of the EUPT item is of high importance as it ensures that portion-to-portion variability has only a negligible impact on the evaluation of the participant's performance. The PT item is tested for homogeneity, typically after bottling and before distribution to participants, but in justifiable cases the tests for homogeneity assessment may also be conducted after the distribution of the material to the participants[8]. The homogeneity assessment usually involves analysis of two replicate analytical portions, taken from at least ten randomly chosen units (bottled portions) of treated PT item. Measurements should be conducted in random order with the aim of minimizing the risk of misinterpreting signal drifts within a measurement sequence as concentration shifts linked to the bottle numbering, i.e. the order of the bottle filling.

The homogeneity test data are statistically evaluated according to ISO 13528:2022, Annex B[9] or to the International Harmonized Protocols jointly published by ISO, AOAC and IUPAC[10]. The results of all homogeneity assessment are presented to the EUPT-SC. In special cases, where the above

[8] To minimize the risk of PT item not being acceptably homogeneous, the organisers may opt to conduct a small-scale preliminary homogeneity test prior to bottling the PT item for shipment. The pre-tests may focus on a selected fraction of the analytes, and may also serve for verifying the presence and the approximate levels of the analytes spiked.

[9] ISO 13528:2022: 'Statistical methods for use in proficiency testing by interlaboratory comparisons", International Organization for Standardization.

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

criteria are not met, the EUPT-SC, considering all relevant aspects (e.g. the homogeneity results of other analytes spiked at the same time, the overall distribution of the participants' results (CV*), the analytical difficulties faced during the tests, and knowledge of the analytical behaviour of the compound in question), may decide to overrule the test. The reasons of this overruling have to be transparently explained in the Final EUPT-Report. For certain analytes with comparable properties, an equivalent distribution within the sample can be expected if they were spiked/used simultaneously. The homogeneity test of one or more of these analytes may thus be skipped or simplified. The organisers should keep an eye on the participants' results of such analytes not tested for homogeneity in order to detect at an early stage any signs that could raise doubts about the homogeneity of the material (e.g. an atypically broad distribution of the results compared to other analytes). In such a case, the EUPT-SC may decide that a proper homogeneity assessment should still be performed to clarify the situation.

Assessing the Stability of the Analytes Contained in the PT Item

The PT item will also be tested for stability - according to ISO 13528:2022, Annex B[9]. The time delay between the first and the last stability test (stability assessment period) must exceed the period of the EUPT-exercise. Typically, the first analysis is carried out shortly before the shipment of the PT items and the last one shortly after the deadline for submission of results. If justifiable, the stability assessment period may precede the PT period, partly overlap with it or postdate it. Close proximity to the PT-period is to be favoured, however, to minimize the risk that matrix properties alter in a way that will affect analyte stability. To better recognise trends and gain additional certainty, one or more additional tests may be conducted by the organisers in the interim. At least 6 sub-samples (analytical portions) should be analysed on each test day (e.g. 2 analytical portions withdrawn from three randomly chosen containers OR 6 portions withdrawn from a single container). In principle, all analytes contained in the PT item should be checked for stability. However, in individual cases, where sufficient knowledge exists that the stability of a certain analyte is very unlikely to be significantly affected during storage (e.g. based on experience from past stability tests or knowledge of its physicochemical properties), the organisers, after consultation with the EUPT-QCG, may decide to omit a specific stability test. The EUPT-SC will finally decide whether analytes for which the stability test was not undertaken will be included in the Final EUPT-Report, considering all relevant aspects, such as the distribution of the participant's results (CV*).

An analyte is considered to be adequately stable if $|y_i - y| \leq 0.3 \times \sigma_{pt}$, with y_i being the mean value of the results of the last stability test, y being the mean value of the results of the first stability test and σ_{pt} being the standard deviation used for proficiency assessment (typically 25 % of the assigned value by default).

The results of all stability tests are presented to the EUPT-SC. In special cases, where the above stability test criteria are not met, the EUPT-SC considering all relevant aspects (e.g. the past experience with the stability of the compound, the overall distribution the participants' results, the measurement variability, analytical difficulties faced during the test and knowledge about the analytical behaviour of the compound in question) may decide to overrule the test. The reasons of this overruling will be transparently explained in the Final EUPT-Report.

The organisers may also decide to conduct additional stability tests at storage conditions other than those recommended to the participants e.g. at ambient temperature.

If insight about insufficient analyte stability is gained before the end of the PT-period, the EUPT-QCG will be contacted in order to decide whether the EUPT-SC should be involved in the discussion (as confidential information is involved), whether the PT-participants should be informed about this insight and whether the affected analytes should be removed from the target list.

Stability during shipment: Considering knowledge about the expected susceptibility of analytes in the PT item to possible losses, the organisers will choose suitable shipping conditions to minimize such losses, e.g. shipment of frozen samples, addition of dry ice. As shipment duration can vary from labs/countries to labs/countries, it is recommended that the organisers keep track of the shipment duration and then decide whether it is reasonable to conduct additional stability tests at conditions simulating shipment. Should critical losses be detected for certain analytes, the EUPT-SC will be informed (or the EUPT-QCG before or during the test). Case-by-case decisions may be made by the EUPT-SC, considering all relevant aspects including the duration and conditions of the shipment to the laboratory as well as the feedback by the laboratory. Follow-up measures in case of instability during shipment may include the exclusion of the affected results from the population used for establishing the assigned value (x_{pt}) and the non-calculation of z scores for the affected analytes in order to avoid unfair penalization of the laboratories involved.

If the PT entails analytes that are expected to have a high risk of degradation within the PT item, the organisers should conduct model tests prior to the final preparation of the test item in order to gain insight about the stability behavior of the analytes intended to be spiked during homogenization, transport and storage of the samples. Based on the results of these experiments measures should be taken to minimize the risk of certain analytes failing to meet the stability criteria, which may include adjusting the conditions of homogenization and/or storage and/or shipment or even deciding not to spike the material with certain analytes.

Methodologies to be used by the Participants

Participating laboratories are instructed to use the analytical procedure(s) that they would routinely employ in official control activities (monitoring etc.). Where an analytical method has not yet been established routinely, this should be stated. This can be done via the EURL data submission tool (in the following named Webtool) by answering the question whether the concerned analyte is included within the routine scope of the laboratory and the question about the analytical experience with the compound.

General Procedures for Reporting Results

Participating laboratories are responsible for reporting their own quantitative results to the organiser within the stipulated deadline. Any analyte targeted by a participating laboratory should be reported as "analysed" in the Webtool. In EUPTs by EURLs responsible for MRM compounds (FV, CF, AO) this is done before shipment of the PT test Item. In EUPT-SRMs this is done in the period during which the platform is open for result submission. Each laboratory will be able to report only one result for each analyte detected in the PT item. The concentrations of the analytes detected should be expressed in 'mg/kg' unless indicated otherwise in the specific protocol of the respective EUPT.

For reporting, concentration values ≤ 0.01 mg/kg are recommended being rounded to two significant figures (e.g. 0.0078; 0.010) and values > 0.01 mg/kg to three significant figures (e.g. 0.123; 1.23; 12.3 mg/kg). No penalties will apply where a laboratory reports deviating numbers of significant figures, but in case of less significant figures, zeros will be assumed after the last significant figure (e.g. 0.1 = 0.100 and 0.11 = 0.110). For the calculation of z scores the values will be used as reported. In the preliminary and final report the results will be shown with up to three significant figures.

Laboratories should not report results below their own reporting limits (RLs). Any reported numerical result that is lower than the RL will be marked as a 'False Reporting' (FR) but it will be allocated a z score as any other numerical result. Such results will be, furthermore, included in the results population for establishing the assigned value (x_{pt}), unless they are eliminated for other reasons (e.g. laboratory status, use of biased methodology).

Correction of Results for Bias

According to the DG-SANTE Guidelines, the result of an analyte needs to be adjusted for method bias if the bias exceeds 20%. Unless the method used inherently accounts for method bias (see cases a – c below), laboratories are required to report the recovery (in percent), and whether their results was corrected mathematically using a recovery factor reflecting the reported recovery.

The EUPT-Panel will examine whether results, for which no correction for bias was undertaken, should be omitted from the population used for calculating the assigned value.

When the laboratory uses any of the following approaches inherently accounting for method bias, this needs to be indicated in the appropriate fields within the Webtool. In such cases, reporting of the recovery rate is not mandatory.

a) use of stable isotope labelled analogues of the target analytes as Internal Standard (ILISs), added to the analytical portion at an early stage of the procedure

b) 'procedural calibration' approach

c) 'standard addition' approach with additions of analyte(s) to the analytical portions before extraction.

Methodology Information

All laboratories are requested to provide information on the analytical method(s) they have used. The Webtool, which serves for submitting analytical results, is typically also used for collecting method information.

The collection of method information is considered very important by the EUPT-SC as it facilitates the interpretation of results and the identification of analytical patterns associated with systematically biased results. A compilation of the methodology information submitted by all participants may be presented in an Annex of the Final EUPT-Report or in a separate report. Where the initial method information provided by the participating laboratories is not sufficient for evaluating methodology-related errors or where additional information critical for results evaluation is needed, the EURLs and/or the EUPT-Panel may decide to conduct specific follow-up surveys among the concerned laboratories. If no sufficient information on the methodology used is provided, the organisers reserve the right not to accept the analytical results reported by the participants concerned or even refuse participation in the following PT.

Where necessary, the methods are evaluated and discussed within the EUPT-SC, especially in those cases where the result distribution is not unimodal or very broad (e.g. CV^* > 35 %).

Where certain methodologies or analytical steps are suspected to lead to biased or otherwise erroneous results, the PT-organisers will substantiate this suspicion by own experiments and discuss the issue with the EUPT-SC. Laboratories affected will be informed, e.g. via direct contact and/or via EURL-workshops or trainings and/or through the inclusion of recommendations within the Final EUPT Report.

Cases where reporting limits (RL) of laboratories exceed the MRRL indicate insufficient sensitivity and may be highlighted in the final report as with “PS” for poor sensitivity.

Results Evaluation

The procedures used for the treatment and assessment of results are described below.

- ***False Positive (FP) Results***

These are results of analytes on the Target Pesticides List that are reported at or above their respective MRRL although they were: (i) “not detected”[11] by the organiser, even after repeated analyses, and/or (ii) “not detected” by the overwhelming majority (e.g. > 95 %) of the participating laboratories that had targeted the specific analytes. In certain instances, case-by-case decisions by the EUPT-SC may be necessary.

Any results reported lower than the MRRL will not be considered as false positives, even though these results should not have been reported. If these results are additionally lower than the lab’s reporting limit, they will be attributed with FR (‘False Reporting’).

- ***False Negative (FN) Results***

These are results of analytes reported by the laboratories as ’analysed’ but without reporting numerical values although they were: a) used by the organiser to treat the PT item and b) detected by the organiser as well as the majority of the participants that had targeted these specific analytes at or above the respective MRRLs. Numerical results < RL (RL= Reporting Limit of the laboratory) may be judged as false negatives and may be also regarded as “not correctly found” when it comes to categorization in A and B based on scope. Such results wouldn’t be reported in a routine laboratory environment. Case-by-case decisions by the EUPT-Panel will be taken by the EUPT-SC in such cases.

Where the RL of a laboratory for a certain analyte present in the PT item exceeds the assigned value, with the laboratory not reporting a numerical value, the result may still be judged as a false negative, despite this reporting being unobjectionable in a routine working environment. The FN judgement should in this case penalize the laboratory for not being able to achieve sufficient sensitivity for the analyte in question.

[11] The term “not detected” is also used in the Webtool. In this context this term entails also all cases where no numerical result were reported (e.g. because the level determined was < MRRL and/or < RL)

In cases of the robust mean of the participant results being less than 3 times higher than the MRRL, false negatives will typically not be assigned. The EUPT-SC may decide to make case-by-case decisions in this respect after considering all relevant factors such as the result distribution and the RLs of the affected labs. In case where the not fixing a valid assigned value is due to other reasons, e.g. because the uncertainty of the assigned value (UAV) criteria were not met and/or because of a bimodal distribution of the participant results, the EUPT-SC will decide on case-by-case basis whether FNs should be assigned for the respective analyte or not.

- ***Estimation of the Assigned Value (x_{pt})***

To minimise the influence of out-lying results on the statistical evaluation, the assigned value x_{pt} (= consensus concentration) will typically be estimated using the robust mean estimate of the participant results (x^*) as described in ISO 13528:2022[12], taking into account the results reported by EU and EFTA countries laboratories only. In special justifiable cases, the EUPT-Panel may decide to include results submitted by laboratories not belonging to the EU-/EFTA-OfLs network or even to use only the results of a subgroup of ('expert') laboratories that have previously repeatedly demonstrated good performance for the specific or similar compounds.

Furthermore, the EUPT-Panel may decide to eliminate certain results traceably associated with bias or gross errors for establishing the assigned value (see 'Omission or Exclusion of results' below).

In special justifiable cases, the EUPT-Panel may furthermore decide to use the spiked concentration of an analyte as the best estimate of the assigned value. In such cases, a detailed explanation of the reasons behind this decision will be given and a comparison with calculations involving robust statistics will be undertaken.

In reports, assigned values will be rounded to 3 significant figures if ≥ 0.01 mg/kg and to 2 significant figures if < 0.01 mg/kg (i.e. 0.0078; 0,123; 1.23; 12.3 mg/kg). For the calculation of z scores, the organisers may opt to use assigned values rounded to more significant figures than those stated above.

Since the assigned values of the EUPT analytes are typically generated using robust mean concentrations of participant results, which are generated by a variety of analytical standards and methods, the assigned values of EUPTs are typically metrologically not traceable.

[12] ISO 13528:2022 'Statistical methods for use in proficiency testing by interlaboratory comparisons", International Organization for Standardization. Therein a specific robust method for determination of the consensus mean and standard deviation without the need for removal of deviating results is described (Algorithm A in Annex C).

- ***Omission or Exclusion of Results***

Results reported by laboratories from non-EU and non-EFTA Member States are typically excluded from the population used to derive the assigned value (for exceptions see 'Estimation of the assigned value').

Before estimating the assigned value, results associated with obvious mistakes have to be examined to decide whether they should be removed from the population. Such gross errors may include incorrect recording (e.g. due to transcription errors by the participant, decimal point faults or transposed digits, incorrect unit), calculation errors (e.g. missing factors), analysis of a wrong sample/extract (e.g. a spiked blank), use of wrong concentrations of standard solutions, incorrect data processing (e.g. integration of wrong peak), inappropriate storage or transport conditions (in case of susceptible compounds), and the use of inappropriate analytical steps or procedures that demonstrably lead to significantly biased results (e.g. employing inappropriate internal standards or analytical steps or conditions leading to considerable losses, due to degradations, adsorptions, incomplete extractions, partitioning etc.). Where the organisers (e.g. after the publication of the preliminary report) receive information that certain participant's results are associated with gross errors, the affected results will be examined on a case-by-case basis to decide whether, or not, they should be excluded from the population used for robust statistics. Results may also be omitted e.g. if an inappropriate method has been used even if they are not outliers.

In case of traceable calculation errors by the participants (e.g. use of wrong factors to express the result as required by the PT's residue definition[13]), and in case of non-reporting results that can be calculated from reported values (e.g. summed result not calculated and not reported), the EUPT-Panel may decide to correct or complement results within the population by applying (the correct) factors. The new population of results may then be used for establishing the assigned values. The z score of the concerned results will, however, be calculated using the originally reported values.

Although robust statistics are applied for estimating assigned values and robust standard deviations, certain results showing a strong bias compared to the rest of the population may be, in certain cases, eliminated before applying robust statistics[14]. To identify such strongly biased results, a preliminary consensus calculation of the robust mean (prelim-x*) may be conducted and any results being ≥ 3-fold the prelim-x*[15] may be potentially eliminated. This approach may need to be iterated if the population still entails obvious outliers.

[13] irrespective of who is accounted responsible for the confusion

[14] Please see ISO 13528:2022 Chapter 6.6." Outlier techniques for individual results', therein 6.6.3, Note 3.

[15] Corresponds to preliminary z scores ≥ 8 using the FFP-approach

The result population remaining after the elimination of certain results as described above may be then used to establish the actual assigned value (x_{pt}) and the robust standard deviation (s*) according to the consensus approach described above. The z scores of all results, including those corrected or removed, are to be recalculated using the new assigned value.

All decisions to omit/exclude results will be discussed with the EUPT-SC and the reasoning for the omission of each result clearly stated in the Final EUPT-Report. However, z scores will be calculated for all results irrespective of the fact that they were omitted from the calculation of the assigned value.

Omitted results might be interesting as they might give indications about possible source(s) of errors. The organisers will thus ask the relevant lab(s) to provide feedback on possible sources of errors (see also "follow-up activities").

- ***Uncertainty of the Assigned Value ($u(x_{pt})$)***

The uncertainty of the robust mean values (x_{pt}) is calculated according to ISO 13528:2022 as:

$$u(x_{pt}) = 1.25 \times \frac{s^*}{\sqrt{p}}$$

where s* is the robust standard deviation and p is the number of results.

A broad results distribution (high s*) and/or a limited number of results (p) will increase the uncertainty of the robust mean $u(x_{pt})$ values exceeding 0.3×σ_{pt} (see ISO 13528:2022) will typically mean that the robust mean is too uncertain for the purpose and cannot be straightforwardly taken up as the assigned value. In each of these cases, investigations for elucidating the reasons behind the high uncertainty should be undertaken. Taking into account all relevant aspects[16] the EUPT-SC may decide that the analyte results should be re-evaluated based on a refined or extended result population or an alternative approach. If, despite these considerations and irrespective of the outcome of the UAV test the EUPT-SC concludes that, the assigned value of a specific analyte is too uncertain for a valid evaluation, it may decide that the results for the analyte in question should not be evaluated or only evaluated for informative purposes.

- ***Considering the UAV when Calculating z Scores***

Where the vast majority of the results is close to the robust mean and narrowly distributed but the UAV-test is still marginally failing[17] (e.g. where $u(x_{pt})$ is up to 0.4×σ_{pt} = 10% in absolute terms), the

[16] e.g. information about methodologies used by the participants (especially if these are likely to produce biased results), multimodality, number of submitted results, homogeneity data, stability data

[17] e.g. due to a combination of few results, and sporadic biased results.

EUPT-Panel may consider to calculate z' scores using the following formula, which considers the uncertainty of the assigned value:

$$z' = \frac{x_i - x_{pt}}{\sqrt{{\sigma_{pt}}^2 + u^2\,(x_{pt})}}$$

where $u(x_{pt})$ being the uncertainty of the assigned value and σ_{pt} being the standard deviation of the assigned value that may be set equal to $FFP\text{-}\sigma_{pt}$ (see below). z' scores will be shown for Informative purposes only.

In special cases[18], the EUPT-SC may consider useful to proceed with the calculation of z scores for both extremes of the assigned value as derived by applying the UAV (i.e. $x_{pt} \pm (x_{pt})$). This upper and lower bound calculation of the z scores will also be for informative purposes only. The aim of this calculation is to help laboratories having performed well in a PT demonstrate their good performance even in cases where the UAV-test has not passed the criteria. Example: x_{pt} = 1.0 mg/kg, (x_{pt}) = 0.1. Taking into account the calculated uncertainty, the AV should range between 0.9 and 1.1 mg/kg. If the result of a laboratory is 0.7 mg/kg, the z score calculates to -1.2 using x_{pt} = 1.0 mg/kg, For the upper limit of x_{pt} = 1.1 the z score calculates to -1.76 and for the lower limit of x_{pt} =0.9 the z score calculates to -0.72. This means that, even at worst-case scenario, the laboratory's result remains within the acceptable range.

- ***Standard Deviation for Proficiency Assessment (Target Standard Deviation)***

The standard deviation for proficiency assessment (σ_{pt}) will be calculated using a Fit-For-Purpose approach with a fixed Relative Standard Deviation (FFP-RSD). Based on experience from previous EUPTs[19], a percentage of 25 % is currently used as FFP-RSD for all analyte-matrix combination, and the Fit-For-Purpose target standard deviation ($FFP\text{-}\sigma_{pt}$) is calculated as follows:

$$FFP\text{-}\sigma_{pt} = 0.25 \times x_{pt}$$

The EUPT-SC reserves the right to also employ other FFP-RSDs or other approaches for setting the assigned value on a case-by-case basis, considering analytical difficulties and experience gained from previous proficiency testings.

For informative purposes the robust relative standard deviation (CV^*) of the participants results is calculated according to ISO 13528:2022; Chapter 7.7 following Algorithm A in Annex C (so called "consensus approach from participant results").

[18] E.g. where the population of results is narrow, but the UAV tests fails due to a few deviating results in combination with a relatively small number of results, e.g. <20

[19] Comparative Study of the Main Top-down Approaches for the Estimation of Measurement Uncertainty in Multiresidue Analysis of Pesticides in Fruits and Vegetables. J. Agric. Food Chem., 2011, 59(14), 7609-7619. DOI:10.1021/jf104060h

- ***z Scores***

This parameter is calculated using the following formula:

$$z_i = \frac{(x_i - x_{pt})}{FFP\text{-}\sigma_{pt}}$$

where x_i is the value reported by the laboratory, x_{pt} is the assigned value, and $FFP\text{-}\sigma_{pt}$ is the standard deviation using the FFP approach. Z scores shown in the preliminary and Final EUPT-Report will be rounded to one decimal place. For the calculation of combined z scores (see below) the original z scores will be used and the combined z scores will be rounded to one decimal place after calculation.

For practical reasons, any z scores > 5 will be typically reported as '> 5' and a value of '5' will be used to calculate combined z scores (p. 19). Following ISO 17043:2023[20], z scores will be classified as follows:

$\|z\| \leq 2.0$	Acceptable
$2.0 < \|z\| < 3.0$	Questionable
$\|z\| \geq 3.0$	Unacceptable

All false negatives will be assigned a z score of -4. These z scores will typically appear in the z score histograms and will be used in the calculation of combined z scores.

- ***Collection of Measurement Uncertainty (MU) Figures***

For each EUPT the participating labs are asked to voluntarily report the MU figure they would report in routine analyses. The EUPT-SC will decide how to evaluate these figures and whether indications will be made to the laboratories in this regard.

- ***Categorization of Laboratories***

The EUPT-SC will decide if and how to classify the laboratories into categories based on their scope and/or performance. Currently, a scope-based classification into Category A and Category B is employed. Laboratories that have:

a) analysed at least 90% of the compulsory analytes in the target pesticides list,
b) reported numerical results for at least 90 % of the compulsory analytes present in the PT item
c) reported no false positives

[20] ISO/IEC 17043:2023. Conformity assessment – General requirements for the competence of proficiency testing providers

are considered to have demonstrated 'sufficient scope' and will be therefore classified into Category A. For the 90% criterion, the number of analytes needed to be correctly analysed to have sufficient scope will be calculated by multiplying the number of compulsory analytes from the Target Pesticides List by 0.9 and rounding to the nearest full number with 0.5 decimals being rounded downwards (see some examples in Table 1).

Table 1: *Number of analytes from the Target Pesticides List needed to be targeted or analytes present in the PT item that need to be correctly detected and quantified to have sufficient scope.*

No. of compulsory analytes present in the PT item / target pesticides list (N)	90 %	No. of compulsory analytes needed to be correctly detected and quantified / targeted to have sufficient scope (n)	n
3	2.7	3	N
4	3.6	4	
5	4.5	4	N - 1
6	5.4	5	
7	6.3	6	
8	7.2	7	
9	8.1	8	
10	9.0	9	
11	9.9	10	
12	10.8	11	
13	11.7	12	
14	12.6	13	
15	13.5	13	N - 2
16	14.4	14	
17	15.3	15	
18	16.2	16	
19	17.1	17	
20	18	18	
21	18.9	19	
22	19.8	20	
23	20.7	21	
24	21.6	22	
25	22.5	22	N - 3
26	23.4	23	

- ***Overall Performance of Laboratories - Combined z Scores***

For evaluation of the overall performance of laboratories the average of the squared z scores (AZ^2)[21] and/or the average of the absolute z scores (AAZ) can be calculated for informative purposes. To minimize the influence of outlying results, the calculation of AZ^2 and AAZ will not be conducted in the case of < 10 and < 5 results, respectively, and z scores higher than 5 will be set as 5. Combined z scores are typically only calculated for laboratories within Category A and considering results of

[21] Laboratory assessment by combined z score values in proficiency tests: experience gained through the EUPT for pesticide residues in fruits and vegetables. Anal. Bioanal. Chem., 2010, 397, 3061–3070. DOI:10.1007/s00216-010-3877-3

compulsory analytes, but the organisers may deviate from this if considered reasonable, provided that a minimum number of results (z scores) have been reported. Combined z scores may be also calculated using results of across PTs.

Considering the cut-off of high z scores at 5, the AZ^2 is calculated as follows:

$$AZ^2 = \frac{\sum_{i=1}^{n} z_i^2}{n}$$

Where n is the number of z scores to be considered in the calculation.

Based on the AZ^2 achieved, the laboratories are classified as follows:

$AZ^2 \leq 2.0$	Good
$2.0 < AZ^2 < 3.0$	Satisfactory
$AZ^2 \geq 3.0$	Unsatisfactory

Combined z scores are considered to be of lesser importance than individual z scores. The EUPT-SC retains the right not to calculate AZ^2 if it is considered as not being useful or if the number of results reported by any participant is considered being too low.

In the case of EUPT-SRMs, where only a few results per laboratory may be available, the average of the absolute z scores (AAZ) may be calculated for informative purposes, but only for labs that have reported enough results to obtain 5 or more z scores. For the calculation of the AAZ, z scores higher than 5 will also be set as 5. The z scores appointed to false negatives will be also included in the calculation of the combined z scores. In general, laboratories should aim to achieve AAZ scores < 0.9, which corresponds to an average bias of 22.5 %[22].

Laboratories within Category B will be typically ranked according to the total number of analytes they correctly reported to be present in the PT item. The number of acceptable z scores achieved may be presented, too.

Publication of Results

The EURLs will publish a preliminary report, containing tentative assigned values and z score values for all analytes present in the PT item, within 2 months of the deadline for result submission. An early

[22] At 22,5% average bias (i.e. AAZ=0.9) and assuming a precision of 10%, the uncertainty calculates to 24.6% (error propagation formula), which is just acceptable. At a precision of 15%, the maximally tolerable average bias calculates to 20%, which translates to an AAZ of 0.8. The uncertainty of the bias was not considered in these calculations.

distribution of the preliminary report, entailing preliminary assigned values (prAV), will allow an early investigation of possible errors by the participants.

The Final EUPT-Report will be published after the EUPT-SC has discussed the results. Taking into account that the EUPT-SC meets normally only once a year (typically in late summer or autumn) to discuss the results of all EUPTs organised by the EURLs earlier in the year, the Final EUPT-Report may be published up to 12 months after the deadline for results submission. Results submitted by non-EU/EFTA laboratories might not always be included in all tables or figures in the Final EUPT-Report.

Certificates of Participation

Together with the Final EUPT-Report, the EUPT organiser will deliver a Certificate of Participation to each participating laboratory showing the z scores achieved for each individual analyte, the classification into Categories, and if deemed necessary also combined z scores. The certificates of participation will be uploaded onto the EURL-DataPool and can be accessed by the concerned laboratories only.

Feedback and Complaints

Participants have the right to complain about any aspect concerning the PT (e.g. about the on-line tools used for registration and data submission, the organisation and communication with the participants, the timing of the PT, transcription errors and the result evaluation if it is not compliant with the provisions of the general protocol). Complaints about a non-arrival of a PT item or about the bad condition of the PT item upon arrival should be done through the Webtool shortly as indicated in the specific protocols. The EURLs will track the complaints and will try to accommodate all substantiated complaints in due time. After the publication of the final EUPT report, the organizers reserve the right not to consider any complaints arriving more than two months after its publication.

Appeals and complaints concerning the principles of organisation and statistical analysis of the results according to the General Protocol should be made prior to the start of a PT. By signing up to an EUPT, the participant agrees with the provisions of the General Protocol valid for the PT-season in question.

At any time before, during or after the PT participants have the possibility to contact the organisers and make improvement suggestions or indicate general errors. After the distribution of the Final EUPT-Report, participating laboratories may be given the opportunity to give their feedback to the organisers and make suggestions for future improvements through a survey.

Correction of Errors

Should errors be discovered in any of the documents issued prior to the EUPT (Calendar, Target Pesticides List, Specific Protocol, General Protocol), the corrected documents will be uploaded onto the website and in the case of substantial errors, the participants will be informed. **Before starting the exercise, participants should make sure to download and carefully study the latest version of these documents.**

If substantial errors are discovered in the Preliminary EUPT-Report the organisers will distribute a new corrected version, therein it will be stated that the previous version is no longer valid. The online version on the PT website will be replaced.

Where substantial errors are discovered in the Final EUPT-Report the EUPT-SC will decide whether a corrigendum will be issued and how this should look like. The online version of the Final EUPT report will be replaced by the new one and all affected labs will be contacted.

If a new version of any EUPT document is released, each page of the new version must be marked in a way distinguishing it from previous versions, e.g. with the version number.

Where errors are discovered in EUPT-Certificates, the revised certificates will be issued and uploaded to the DataPool. The concerned laboratories will be informed and asked to download the corrected ones.

Follow-up Activities

Laboratories are expected to undertake follow-up activities to trace back the sources of erroneous or strongly deviating results (typically those with $|z| > 2.0$), including all false positives. In exceptional cases, follow-up activities may even be indicated for results within $|z| \leq 2.0$, e.g., if two errors with opposed tendency cancel each other leading to acceptable results, or where the procedure used turns out being significantly biased.

Upon request, the laboratory's corresponding NRL and EURL are to be informed of the outcome of any investigative activities for false positives, false negatives and for results with $|z| \geq 3.0$. Concerning z scores between 2.0 and 3.0 the communication of the outcome of follow-up activities is optional but highly encouraged where the source of deviation could be identified and could be of interest to other labs.

In accordance with the instructions from DG-SANTE, the "Protocol for management of underperformance in comparative testing and/or lack of collaboration of National Reference Laboratories (NRLs) with EU Reference Laboratories (EURLs) activities" is to be followed.

NRLs will be considered as **underperforming in relation to scope** if in at least two of the last four EUPTs falling within their responsibility area they: a) haven't participated, or b) targeted less than 90% of the compulsory analytes in the target lists (80% for SRM-compounds), or c) detected less than 90% of the compulsory compounds present in the PT items (80% for SRM-compounds). Additionally, NRLs that obtained AZ^2 higher than 3 (AAZ higher than 1.3 for SRM-compounds) in two consecutive EUPTs of the last four EUPTs, will be considered as underperforming in accuracy. As soon as underperformance of an NRL is detected, a two-step protocol established by DG-SANTE will be applied[23]:

Phase 1:

- Identifying the origin of the bad results (failure in EUPTs).
- Actions: On the spot visits and training if necessary and repetition of the comparative test if feasible and close the assessment of results by the EURL.

Phase 2:

- If the results still reveal underperformance, the Commission shall be informed officially by the EURL including a report of the main findings and corrective actions.
- The Commission shall inform the Competent Authority and require appropriate actions to be taken.

Underperformance rules for the OfLs will be established at a later stage.

Disclaimer

The EUPT-SC retains the right to change any parts of this EUPT – General Protocol based on new scientific or technical information. Any changes will be communicated in due course.

[23] Article 101 of Regulation (EU) 2017/625