

# Guide for participants



This guide is based on the ProfTest Syke guide 21-112 Guide for laboratories version 03 (3 May 2024). This guide has been updated based on the current operation procedures, also technical issues has been updated and clarified especially in chapter 5.9 and in appendix 2.

---

## Contents

|     |   |    |
|-----|---|----|
| 1   | Introduction  | 3  |
| 2   | Interlaboratory comparisons as part of the laboratory management system               | 3  |
| 3   | Syke as a proficiency test provider   | 3  |
| 3.1 | Research infrastructure unit of Finnish Environment Institute                         | 3  |
| 3.2 | Advisory group  | 3  |
| 3.3 | Interlaboratory comparisons organized by ProfTest Syke                                | 3  |
| 3.4 | Confidentiality and handling of personal information                                  | 4  |
| 3.5 | Participant feedback  | 4  |
| 3.6 | Subcontracting and cooperation  | 5  |
| 4   | Organizing of proficiency tests at ProfTest Syke                                      | 5  |
| 4.1 | Planning and marketing  | 5  |
| 4.2 | Realization of the PTs and ILCs, including the samples                                | 5  |
| 4.3 | Sample delivery   | 6  |
| 4.4 | Processing of results   | 6  |
| 4.5 | Performance evaluation  | 6  |
| 4.6 | Reporting of results  | 7  |
| 4.7 | Costs and invoicing   | 7  |
| 4.8 | Client support and troubleshooting  | 8  |
| 5   | Participation in ProfTest Syke interlaboratory comparisons                            | 8  |
| 5.1 | Contact person  | 8  |
| 5.2 | Registration  | 9  |
| 5.3 | Cancellation of registration  | 9  |
| 5.4 | Receipt of samples  | 9  |
| 5.5 | Sample storage  | 10 |
| 5.6 | Analysis  | 10 |
| 5.7 | Reporting results to ProfTest Syke  | 10 |
| 5.8 | Preliminary results report  | 11 |
| 5.9 | Final report  | 12 |
| 6   | Further information for participants  | 12 |
| 7   | Revisions and distribution of the Guide for participating laboratories                | 12 |
| 8   | References  | 13 |
|     | Appendix 1. Concepts and definitions  | 14 |
|     | Appendix 2. Statistical procedures for testing the samples and processing the results | 16 |
|     | Appendix 3. Reporting results of individual participant                               | 23 |

## 1 Introduction

The proficiency tests (PT) and other interlaboratory comparisons (ILC) organized by the Finnish Environment Institute (Syke) are provided under the name of ProfTest Syke. Most commonly organized ProfTest Syke interlaboratory comparisons are the proficiency tests. This guide is mainly for proficiency test participants participating in tests for chemical analysis, but, when applicable, may also be adapted for other interlaboratory comparisons arranged by ProfTest Syke.

This guide aims to provide an overview of organizing the ProfTest Syke ILCs, to assist the understanding of the instructions given for each ILC.

## 2 Interlaboratory comparisons as part of the laboratory management system

Most Finnish laboratories have a management system based on the standard SFS-EN ISO/IEC 17025 [1], which requires effective quality control procedures for monitoring the validity of analytical results. Widely used and accepted way to monitor the validity of the analytical results is to participate in interlaboratory comparisons or in proficiency testing schemes. The primary aim of ILCs is to help individual participants to monitor the reliability of their test results and to take corrective actions where necessary to improve the quality of results. The participation in ILCs also increases the trust of the participant's clients by increasing the awareness of the quality of the results and their comparability. The important concepts and definitions for organizing the ILCs are shown in Appendix 1. Eurachem has compiled a guide for selecting, using and interpreting proficiency testing schemes for laboratories [2], and the Finnish Accreditation Service (FINAS) has published its own policy on ILCs [3].

## 3 Syke as a proficiency test provider

### 3.1 Research infrastructure unit of Finnish Environment Institute

Finnish Environment Institute is a national environmental reference laboratory established under the Environmental Protection Act. The duties of the reference laboratory include providing interlaboratory comparisons for analytical laboratories and other producers of environmental information. Research infrastructure unit is responsible of the reference laboratory activities within Syke. Syke is accredited by the Finnish Accreditation Service as a testing laboratory T003 and a calibration laboratory K054 (SFS-EN ISO/IEC 17025) as well as a proficiency testing provider (PT01, SFS-EN ISO/IEC 17043, [finas.fi/sites/en](https://finas.fi/sites/en)). ProfTest Syke interlaboratory comparisons are widely utilized for environmental measurements, environmental sampling as well as other relevant sectors.

### 3.2 Advisory group

The interlaboratory comparison services in Finnish Environment Institute, ProfTest Syke, are guided by an advisory group comprised of representatives from different relevant sectors. The advisory group provides expert support, proposes improvements, and represents client perspective. It also facilitates information flow between the participants and the provider of interlaboratory comparisons. The members of the group are informed on the ProfTest Syke website ([syke.fi/proftest/en](https://syke.fi/proftest/en)).

### 3.3 Interlaboratory comparisons organized by ProfTest Syke

ProfTest Syke provides interlaboratory comparisons both nationally and internationally. The yearly amount of interlaboratory comparisons varies depending on the needs of the participants as well as on the available resources at ProfTest

Syke. Yearly the interlaboratory comparison have altogether over 300 participants. Number of participants of an individual ILC varies from 5 to 55. More information of the ILCs provided by ProfTest Syke is given on the website ([syke.fi/proftest/en](http://syke.fi/proftest/en)).

### 3.4 Confidentiality and handling of personal information

The provider handles the participant results confidentially. A **permanent laboratory code** is assigned for each participant when participating first time in ProfTest Syke interlaboratory comparison. Customer information and results are handled in the ProfTest Syke's electronic client interface, ProfTestWEB ([www.p5.ymparisto.fi/Labtest/en](http://www.p5.ymparisto.fi/Labtest/en)). The permanent laboratory code is shown on the 'Customer information' page on ProfTestWEB.

To assure the confidentiality, ProfTest Syke does not use this permanent laboratory code when reporting results of any ILC. For each ILC the participants get randomly and separately chosen **participant codes** (Participant id). The participant code for each test is available via the customer profile on ProfTestWEB. When especially needed, the permission to inform the participant code of a particular ILC is requested from the participant, if the participant is producing results for the Finnish environmental authorities. If the participant codes are provided, the environmental authorities are reminded of the confidentiality. Generally, if needed, the participant informs their participant code directly to the environmental authorities.

When placing an order for an ILC, general participant information is collected from the participant: the name of the contact person and participant (ie. e.g. the name of the laboratory), address and billing information. The participants can review and update their information via ProfTestWEB. If the contact person wants to remove his/her information from ProfTest Syke database, this could be done by sending the request via email: [proftest@syke.fi](mailto:proftest@syke.fi). The personal information related to the ILCs is handled by the personnel of ProfTest Syke and the designated persons of the technical administration (system maintenance and user permissions).

### 3.5 Participant feedback

Participant feedback plays an important role in improving the interlaboratory comparison services at ProfTest Syke. The feedback and questions are invited to be delivered at any time via email: [proftest@syke.fi](mailto:proftest@syke.fi). Feedback could be given also via ProfTestWEB. Further, feedback could be delivered also via the members of the advisory board. ProfTest Syke also arranges a feedback questionnaire for participants every few years. Besides questions on customer satisfaction, the service provider seeks opinions also on proficiency test timetables, sample concentration ranges, sample types, and the content of reports.

All feedback will be replied in shortest possible time.

All feedback related to the interlaboratory comparison services is documented and exploited when arranging future proficiency tests and improving the activities. For each ILC, the related feedback and comments are included in the final report of the ILC.

General feedback related to the ProfTest Syke services could be sent directly to Tero Eklín, director of Research infrastructure unit ([firstname.lastname@syke.fi](mailto:firstname.lastname@syke.fi)). Where necessary, disagreements arising between the organizer and the participants are aimed to be settled through negotiation and conciliation.

### 3.6 Subcontracting and cooperation

Commonly the ProfTest Syke ILCs are organized together with Syke's testing and calibration laboratory. Subcontracting and/or cooperation is used when needed analyses are not available at Syke or needed resources are not available at the time in question. Subcontracting could concern for example sample collection, preparation of samples and sample testing as well as analytical expertise. Neither the permanent laboratory codes nor the participant codes are ever disclosed to subcontractors or to cooperation partners, nor do they ever evaluate participant performance.

The same competence requirements are applied for subcontracted functions than those of the organizer. All details of subcontracting and competence requirements are documented.

In interlaboratory comparisons the domestic cooperation partners are the operators whose activities on the basis of the laws or regulations include the reference laboratory activities or other equivalent obligations. For example, in the proficiency test of radon in groundwater the cooperation partner is Radiation and Nuclear Safety Authority (STUK). Cooperation partner can also be a subcontractor, when subcontracting includes a large entity (for example, a wide analytical expertise). Cooperation partners' competence requirements are the same as those of proficiency test provider.

## 4 Organizing of proficiency tests at ProfTest Syke

### 4.1 Planning and marketing

The annual program of ProfTest Syke proficiency tests is published in October–November of the previous year on the ProfTest Syke website ([syke.fi/proftest/en](http://syke.fi/proftest/en)). The participants of ProfTest Syke proficiency tests are informed of the publication of annual program by email. ProfTest Syke also promotes the upcoming PTs and ILCs via LinkedIn ([linkedin.com/in/proftestsyke/](https://www.linkedin.com/in/proftestsyke/)) and via Eptis database ([eptis.bam.de](http://eptis.bam.de)).

For proficiency tests and interlaboratory comparison organized infrequently or seldom, pre-registration or other information might be requested prior the test is properly planned.

### 4.2 Realization of the PTs and ILCs, including the samples

The registration opens about two months before the planned realization of the ILC. The registration opens on ProfTestWEB and ProfTest Syke sends an information letter by email to receivers (mainly participants of former tests) who are potentially interested to participate in the ILC. The information letter is also available on the ProfTest Syke website ([syke.fi/proftest/en](http://syke.fi/proftest/en)). More information could be asked from ProfTest Syke customer service ([proftest@syke.fi](mailto:proftest@syke.fi)).

ProfTest Syke reserves the right to cancel the ILC if the number of participants is significantly lower than anticipated. Participants will be informed of cancellation at the latest two weeks before the planned time of realization of the proficiency test.

Sample preparation starts well before the realization of the ILC. The number of subsamples to be prepared is confirmed after the closure of the registration. It is of high importance for the ILCs to prepare both homogenous and stable samples. While most samples are delivered ready for analysis, in certain cases participants are requested to complete the sample preparation, e.g. by adding the solution containing the measurand(s) provided with the sample. This procedure is applied when the

sample contains unstable measurands (e.g. BOD<sub>7</sub>).

Sample homogeneity is tested using at least one of its measurands (see Appendix 2, part 3 Homogeneity test). E.g. the sample containers for nitrogen compounds are tested by determining total nitrogen, as it resembles best the possible inhomogeneity caused by particles. Stability of the measurand is tested if it is not known to be stable based on the literature or experience (see Appendix 2, part 9 Stability test).

#### **4.3 Sample delivery**

The transport date and estimated arrival date of samples is informed to the participants in the information letter. Samples are generally delivered within 24 hours and special arrangements may be made to govern timely deliveries to participants abroad. The consignment number (or reference number) is informed to the participants abroad, thereby enabling shipments to be tracked via internet. The provider follows the stability of samples during the shipment when the measurands have poor stability (e.g. temperature control or weighing prior and after delivery).

#### **4.4 Processing of results**

The results of participants are processed applying ISO 13528 standard [5] and IUPAC Technical report [6]. First the normality of data is studied (see Appendix 2, part 7 Normality) and outliers are removed based on the outlier tests performed (Appendix 2, part 2 Outlier tests). The assigned value for the measurand is usually either the calculated value (synthetic samples) or the robust mean, the mean or the median of results reported by the participants. The certified value of certified reference material (CRM) or a value determined using CRM may also serve as the assigned value. In special cases the assigned value may be the consensus value of the results of expert laboratories selected in advance or it may be metrologically traceable result. The expanded uncertainty is evaluated for the assigned value (see Appendix 2, part 10 Uncertainty and reliability of assigned value). If the number of participant results is low (fewer than 6) or the results are widely scattered, either the assigned value and method for performance evaluation are estimated separately or assigned value is not set.

#### **4.5 Performance evaluation**

The performance evaluation is usually based on the z scores (see Appendix 2, part 11 z score), where the provider sets the standard deviation (accepted deviation from the assigned value, based on predefined criteria) for proficiency assessment. The standard deviation for proficiency assessment is estimated i.a. based on the concentration of the measurand, the type and complexity of analytical method employed (different e.g. when determining pH or mineral oil content of water), the results of homogeneity and stability tests, the uncertainty of the assigned value, the standard deviation of results, and the long-term variation in the former proficiency tests. The standard deviation for proficiency assessment can also be based on the legislative requirements. Preliminary values for standard deviation for proficiency assessment are provided in the sample letter and the values are reviewed and finalized when processing of results.

If the standard deviation for proficiency assessment set by the provider is not appropriate for the participant's purpose, the participant may recalculate the z score using the formula shown in Appendix 3.

The reliability of the assigned value is tested by comparing its uncertainty to standard

deviation for proficiency assessment (Appendix 2, part 10). The reliability of the standard deviation for proficiency assessment is tested by comparing it to the standard deviation of the participants' results (Appendix 2, part 11).

When the participant has reported their results together with the uncertainty information, the zeta scores and their comparison to the z scores are given to the participants as part of the preliminary results report (Appendix 2, part 12).

When there are only few reported results for a measurand ( $n < 6$ ), the performance could be evaluated by the means of  $D\%$  (*Difference*) or  $E_n$  (*Error, normalized*) scores.  $D\%$  and  $E_n$  scores describe the difference between the participant results and assigned value.  $E_n$  score includes the expanded uncertainties of the participant result and the assigned value.

#### 4.6 Reporting of results

The processing, evaluation, and reporting of the results are based on the information reported by the participants. ProfTest Syke is not responsible for the correctness of the information reported by the participants (e.g. the accreditation status of the results). The correctness of the reported results of participants could affect to the correctness of the final report.

ProfTest Syke publishes the preliminary results report of the ILCs mostly within 1-2 weeks after receiving the results. The participant code for each test is available via the customer profile on ProfTestWEB, on the page of the test and the code is official when the preliminary results report is published.

The final report of the ILC is published within 2–5 months after receiving the results. The final report is published in English, when more than 10% of the participants are from abroad, or when English report is otherwise more applicable. In other cases, the report is published in Finnish. The report includes summaries of sample preparation and testing, more detailed information is available from the provider if needed.

#### 4.7 Costs and invoicing

Providing interlaboratory comparisons is a commercial service of Syke governed by the Act on Criteria for Charges Payable to the State (1992) and its subordinate statutes. All prices are subject to valid VAT (value added tax) unless the payer is classified as a government department.

Costs are calculated on the basis of e.g. equipment, labour, delivery, and similar expenses. The costs take also into account the general change in price levels. Usually the price is divided into basic fee (same for all participants) and separate fees for the samples. Basic fee for participation includes the sample delivery costs within Europe. Participants from outside Europe are kindly instructed to contact the provider to get more information of the delivery costs.

A cost estimate is prepared for each interlaboratory comparison at the time when preparing the annual program and it is reviewed when dispatching the information letter. The estimated costs may change, for example, if the test program is modified by the request of participants, or due to a substantial increase in costs.

The invoicing is done based on the participants' orders and invoices are dispatched after publishing the preliminary results report. The provider defrays the delivery costs of the damaged or missing samples while the costs of providing and delivering additional samples must be borne by the participants.

If participant withdraws their registration later than **two weeks prior the foreseen sample delivery day**, the participant will be charged **70 % of the basic fee**. The participation fee must be paid in full when the participant has registered and received the samples but does not deliver the test results to the organizer. Each participant must defray its own analysis costs and the possible customs fees and similar.

**The participant may provide several results for measurands for each interlaboratory comparison. The participant receives separate evaluation for each additional result set and the provider will charge an additional fee of 40 % of the basic fee for each additional data set.**

The samples are pre-tested. However, in case of sample preparation failure observed after sample delivery, there is no charge for the participants. If possible, a new sample will be delivered to the participant at the standard charge.

#### 4.8 Client support and troubleshooting

The frequently asked questions are collected and answered on ProfTest Syke website ([syke.fi/proftest/en](http://syke.fi/proftest/en)).

The provider and analytical experts assist the participants in solving problems related to unforeseen performance. They may be contacted after the preliminary results report has been delivered, especially if the interlaboratory comparison results indicate a need for corrective actions in the participating laboratory. The analytical experts for each ILC are listed in the letters and reports.

Participant may request parallel analysis together with the testing laboratory of Syke research infrastructure unit or with some other laboratory. These analysis requests will be charged separately.

If a participant has discovered a problem within their analysis, they may use the possible spare batch of sample material for their internal quality assurance. Participants may order samples also later to resolve problems or to test methods. ProfTest Syke stores the samples until the publication of the final report, and samples of stable measurands are stored for two years. Samples are subject to a delivery charge as well as the sample fee defined for the ILC.

On request ProfTest Syke could provide individual participants a summary of specific determinations (in the form of z scores) spanning a period of several years. The charged fee for the summary of participant's performance over the longer period will be the cost of data retrieval.

## 5 Participation in ProfTest Syke interlaboratory comparisons

### 5.1 Contact person

ProfTest Syke maintains a register of the participants of the interlaboratory comparisons. Participants must appoint a contact person and preferably a deputy for the communication with the ILC provider. The contact person will serve as the addressee for samples and ILC results and will be advised of other substantial information related to the ILC. The given contact information is used for sample delivery as well as for invoicing the participation. The contact person can view and update their information via the customer profile in ProfTestWEB.

The provider must be notified when a new contact person is appointed. If the contact person wants to remove his/her information from the register, a request could be



delivered via email: [proftest@syke.fi](mailto:proftest@syke.fi).

## 5.2 Registration

Participants register (creates order) into an open ILC via the electronic client interface, ProftestWEB ([wwwp5.ymparisto.fi/labtest/en](http://wwwp5.ymparisto.fi/labtest/en)) according to the given timetable. The interface could be found via Proftest Syke website ([syke.fi/proftest/en](http://syke.fi/proftest/en) → Current proficiency tests).

When the participant has already used ProftestWEB, username and password is used to log in. Then the contact information is filled in automatically to the New order form. Via the Orders page on ProftestWEB, it is also possible to register to an open ILC without login. In such case, after the order is sent, the provider gives the participant access to the interface.

When registering, the participant orders the needed samples by selecting them on the order form. The participant may order several samples when needed. The cost for samples is indicated on the order form. The participant could also deliver additional set(s) of results. The participant is advised to contact the provider in such case, the provider then creates multiple result forms for the participant. A supplementary charge is added for this (see Chapter 4.7).

At the time of registration, participant should deliver the invoicing information including the VAT number of foreign participant's institute, their own order number (if needed), client code, or invoicing address, when it is not the address of the sample delivery address.

The registration is accepted by Proftest Syke and the acceptance is shown as a date stamp on the information of the current test (Tests → Orders).

## 5.3 Cancellation of registration

The registration is binding. However, in exceptional cases the participant may cancel their registration free of charge **no later than two weeks before the sample delivery date**. For later cancellation, see the fee in Chapter 4.7.

## 5.4 Receipt of samples

At the participating laboratory, the contact person must ensure that the staff is notified of the incoming samples to prevent them from being incorrectly stored for too long. The proficiency test provider must be notified immediately if the samples have not arrived within the specified period.

**The cover letter for samples (sample letter), delivered together with the samples, should be read carefully before analysing any samples.** The letter is available also on the page of current test in ProftestWEB.

The recipient should check the contents of the sample package when the samples arrive, and the arrival of the samples is reported to the provider. Further, the provider should be notified immediately of any broken sample containers or missing samples to ensure that new samples are sent promptly. Sample arrival is reported to the provider via electronic "Sample arrival" form available via QR code or link in sample letter. The link is available also on ProftestWEB, on the page of current test. The form should be filled and delivered to the provider within the requested time. The time of receiving the samples is filled into the form as well as other information requested. The form is designed to help the provider to monitor the delivery process and any problems that may arise.

The participant should label the sample bottles according to their own standard procedures. Participants should note that sample labels on the sample bottles do not withstand e.g. thermal treatment in water (pH determination) or autoclaving ( $N_{\text{tot}}$ ).

## 5.5 Sample storage

The sample letter includes storage instructions. Samples should generally be stored in refrigerator (4 °C) until the time of analysis. Instructions are given separately in special cases (e.g. dried solid samples: storage at 20 °C).

## 5.6 Analysis

Samples are analysed **within the laboratory where they are delivered to and using the normal procedures** of the participant. When necessary, the provider may issue special instructions for sample pretreatment and measurements.

If the participant deviates from the instructions and recommendations issued with the sample, the deviation and the reason for it should be informed when reporting the results. It is particularly important to inform the provider about the deviations from the recommended time of analysis, as these deviations could affect the evaluation of laboratory performance. If the participant has difficulties with the measurement deadlines, they should contact the provider to rearrange the timetable (if possible).

The provider requests participants to report either one test result or multiple results of parallel analysis. Parallel testing is a repeat of the whole analysis from beginning to end, **including the sample preparation steps**. When parallel results are not requested by the provider, the participant will perform the analysis with as many parallel tests as are normally conducted for the measurement.

The test analysis is also subject to normal quality assurance procedures.

## 5.7 Reporting results to Proftest Syke

The results for the ILCs are reported mainly via ProftestWEB ([wwwp5.ymparisto.fi/labtest/en](http://wwwp5.ymparisto.fi/labtest/en)). In special cases, e.g. for rarely conducted ILCs, a case-specific results sheet (Excel) or other means of result reporting could be used. In such cases the participants are separately instructed.

The results are to be reported according to the given timetable enabling the provider to report the preliminary results report on time. While overdue results are generally excluded from result processing (unless otherwise agreed), participants remain liable for the participation charge.

The results are to be reported with one more significant number than specified in the analytical instructions. Results are reported with as many parallel results and in the units as requested.

The used test methods are reported by choosing the appropriate method from the drop-down menu on the Save results page. If no method is appropriate, then "Other method" is selected and briefly described. Literature reference does not suffice, as the provider does not necessarily have access to all references. Details of analytical methods are important, as they enable the provider to compare the results of various methods. Sample pre-treatment details are particularly crucial, for example, when interpreting the results of organic analyses.

**When reporting results, special attention should be paid to result units, to the requested number of parallel results, to the amount of significant numbers, and**

**to ensuring that result is entered on the correct line.** These points have proved to be the most common sources of error when reporting results. When the participant has not followed the given instructions, in general, their result will be excluded when determining the assigned value.

## 5.8 Preliminary results report

The preliminary results report is available on ProfTestWEB, on the page of the test. The preliminary results report is also sent via email to the contact person of the participant. Participant's **participation code** is available on ProfTestWEB, on the page of the test. When required, the participation code could be obtained from the provider.

In general, the preliminary results report is available both in Finnish and in English. If the ILC participants are solely from Finland, the preliminary results report is available only in Finnish. In some larger ILCs having participants from abroad, the preliminary results report could be available only in English.

The purpose of the preliminary results report is to:

- Provide feedback to participants on their results and performance in the ILC at the earliest, and
- Enable the participant to verify that no errors have occurred in reporting the result data. Therefore, the preliminary results report is mostly provided both in Finnish and in English.

The following appendices are usually provided with the preliminary results report:

- Result tables for individual participants (see Appendix 3 for an example) including the results reported by participant.
  - When results are reported as parallel results, the preliminary results report has the mean value.
- Definitions of statistical parameters.
- Summary of the ILC.
- Summary of z scores.
- Summaries of z and zeta scores (Appendix 2, part 12).
- Summary of D% and E<sub>n</sub> scores, when applicable.

Participants should check that their results are correct in the data treatment. Participants may comment the preliminary results report within the given commenting period. After publication of the preliminary results report, the participant results will be corrected only in exceptional cases, but details of errors will assist the performance evaluation. Exceptions could be the errors caused by the provider or errors in reporting units in cases where the number of results is too low for statistical data processing. For the results corrected after the publication of the preliminary results report a written evaluation will be given in the final report.

Participants are kindly requested to report the causes of deviant results, as these may help other participants encountering similar deviations. Additionally, it enables the provider to classify the causes of deviant results in the final report.

## 5.9 Final report

The final report for each interlaboratory comparison is published in the publication series Reports of Finnish Environment Institute and stored permanently to HELDA, the open digital repository maintained by the University of Helsinki ([helda.helsinki.fi/syke](https://helda.helsinki.fi/syke)). The participants are informed of the published final report and the link is found via ProftestWEB as well as on Proftest Syke website ([syke.fi/proftest/en](https://syke.fi/proftest/en)).

If the published final report contains factual errors, those will be swiftly corrected. The made corrections are marked in the header of the corrected page for information, and also described on the title page of the report. If the correction is significant, the final report will be published with a new ISBN number in the HELDA publication database. With minor corrections, same ISBN number remains. The participants of scheme will be informed about the corrections made and, if necessary, a link to the new final report will be delivered.

## 6 Further information for participants

Information on interlaboratory comparisons arranged by other providers is available from Eptis, the European information system ([www.eptis.bam.de](http://www.eptis.bam.de)).

Nordtest has published two useful guides in English: A Handbook for Chemical Analytical Laboratories [7] and a Handbook for calculation of measurement uncertainty in environmental laboratories [8]. A measurement uncertainty software application based on the latter handbook is developed by Syke's Calibration and contract laboratory and is available on their webpage [9]. Both guides are available in several languages.

## 7 Revisions and distribution of the Guide for participating laboratories

This guide is available on the Proftest Syke [website](#) and will be revised as necessary. Participants are responsible for discarding any outdated versions. Revised version will be advertised on the Proftest Syke [website](#).

## 8 References

1. SFS-EN ISO/IEC 17025, 2017. General requirements for the competence of testing and calibration laboratories. Finnish Standard Association (SFS), Helsinki.
2. EURACHEM Guide, 2011. Selection, use and interpretation of proficiency testing (PT) schemes ([eurachem.org/index.php/publications/guides/usingpt](http://eurachem.org/index.php/publications/guides/usingpt)).
3. FINAS A2/2020. Principles for the assessment of the quality assurance and the proficiency testing practices in laboratories ([finas.fi](http://finas.fi)).
4. ISO/IEC 17043, 2010. Conformity assessment — General requirements for proficiency testing. Finnish Standard Association (SFS), Helsinki.
5. ISO 13528, 2022. Statistical methods for use in proficiency testing by interlaboratory comparisons.
6. Thompson, M., Ellison, S.L. R., Wood, R., 2006. The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry laboratories (IUPAC Technical report). Pure Appl. Chem. 78: 145-196. (<http://dx.doi.org/10.1351/pac200678010145>).
7. Nordtest Report TR 569, Edition 5.1, 2018. Internal Quality Control – Handbook for Chemical Analytical Laboratories. ([nordtest.info](http://nordtest.info)).
8. Nordtest Report TR 537, Edition 4, 2017. Handbook for calculation of measurement uncertainty in environmental laboratories ([nordtest.info](http://nordtest.info)).
9. Näykki, T., Virtanen, A. and Leito, I., 2012. Software support for the Nordtest method of measurement uncertainty evaluation. Accred. Qual. Assur. 17: 603-612. MUKIT website: [syke.fi/envical](http://syke.fi/envical).
10. ISO 5725-2, 1994. Accuracy (trueness and precision) of Measurement Methods and Results - Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method.
11. ISO 16269-4, 2010. Statistical interpretation of data — Part 4: Detection and treatment of outliers (Annex A).
12. Lisinger, T.P.J., Kandler, W., Krska, R., Grasserbauer, M., 1998. The influence of different evaluation techniques on the results of interlaboratory comparisons. Accred Qual Assur 3: 322-327 (<https://doi.org/10.1007/s007690050254>).

## **Appendix 1. Concepts and definitions**

### **Assigned value, reference value**

Value attributed to a particular quantity and accepted, sometimes by convention, as having an appropriate uncertainty for a given purpose.

### **Certified reference material, CRM**

A reference material, accompanied by a certificate or other official document, one or more of whose property values are certified by a technical procedure.

### **Homogeneity**

All delivered samples have the same composition.

### **Interlaboratory comparison**

Organization and performance evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions.

### **Normality**

The extent to which the observed distribution approximates to a normal distribution in a test result.

### **Outlier**

Extreme value locating far from the rest of the domain values. Outliers are determined using the Cochran, Grubbs, or Hampel statistical tests.

### **Precision**

The closeness of results when measurements are repeated several times under stipulated conditions. The smaller the random error distribution, the more precise the method.

### **Proficiency testing**

Evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons.

### **Provider**

Organization which takes responsibility for all tasks in the development and operation of a proficiency testing scheme.

### **Reference laboratory**

A laboratory that issues reference values with a known uncertainty for a given material.

### **Reference material, RM**

Material or substance of whose property values are sufficiently homogenous and well established to be used for calibrating an apparatus, assessing a measurement method, and assigning values to materials.

### **Repeatability**

Identical test results from repeated tests performed within a short period by the same operator, or by another operator using the same method, on identical test items, using the same equipment or different equipment in the same laboratory.

**Replicate determination**

Two or more parallel determinations, where the determination is repeated from beginning to end (including the pre-process stages).

**Reproducibility**

Measurement conformity where test results are obtained using different methods, different equipment, in different laboratories, by different operators and at intervals that are long in relation to a single test. The reproducibility deviation is usually greater than the repeatability deviation. It is generally used in proficiency testing schemes.

**Stability**

Samples remain unchanged (stable) until they are analysed.

**Standard deviation for proficiency assessment**

Measure of dispersion used in assessing proficiency, based on the available information.

**Traceability**

The relation of measured results through an unbroken chain of measurements to the appropriate national or international standards.

**Trueness**

The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value.

**Uncertainty of measurements**

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

## Appendix 2. Statistical procedures for testing the samples and processing the results

### 1 ANOVA test

When participants report several replicate results the ANOVA test (analysis of variance) can be used to estimate the standard errors within and between participant results [10].

The repeatability standard error  $s_w$  (within participant results) is calculated using the participants' replicate results. Also, the between participants' results standard error  $s_b$  is calculated. The reproducibility standard error  $s_t$  is calculated according to the equation:

$$s_t = \sqrt{s_w^2 + s_b^2}$$

### 2 Outlier tests

Outlier tests are used to identify the results that differ statistically significantly from the other results in the data set (in practice, the values outside the 95 % confidence level).

The parallel results are tested with Cochran's test and the deviation of the participant result (or the mean of parallel measurements) from the data set is tested with the Grubbs or Hampel test.

#### *Cochran's test*

Cochran's test is designed to assess the within-laboratory deviation, i.e. to determine excessive discrepancies between participants [10]. Participants are numbered 1, 2, ...,  $p$  and iterated distributions  $s_1, s_2, \dots, s_p$ . The test value is:

$$C = \frac{s_{\max}^2}{\sum_{i=1}^p s_i^2}, \text{ where}$$

$s_i$  = the standard deviation of the replicate (parallel) results

$s_{\max}$  = the maximum standard deviation of the replicate results

$p$  = the number of the result series.

Cochran's test is performed when there are parallel results from at least three participants in the result data.



### Grubbs test

With the Grubbs test the deviation of either one result (smallest or highest, Grubbs) or two results (biggest or smallest, Grubbs 2) are tested, one outlier at a time. The possible outliers are tested iteratively with the GESD test [11].

The used software sets value  $K$  as the maximum number of possible outliers,  $K$  being 25 % from the total number of results, 20 numbers at maximum. Then the numbers furthest from the average are examined in order, so that the numbers already processed are not included in the calculation of the average when examining the following deviations.

Value  $G_k$  is calculated for each possible outlier to be tested  $x_k, k = 1, \dots, K$ :

$$G_k = \max_{i \in I_k} \frac{|x_i - \bar{x}(I_k)|}{s(I_k)},$$

where  $I_k$  is the set of the remaining numbers after outlier removal during previous rounds,  $\bar{x}(I_k)$  is the mean calculated from the set of the remaining numbers and  $s(I_k)$  is their standard deviation. Result  $x_k$  is outlier if  $G_k$  is greater than critical value at the 5 % significance level, calculated according to the standard.

In the Grubbs 2 test, the value  $G$  is obtained by comparing the variance of the entire data set to the variance obtained when the two extreme (smallest or highest) values are removed. The results are outliers if  $G$  is greater than the critical value at the 5 % significance level [11].

### Hampel test

Hampel test is based on the median of the data set and the absolute value of a single value. The median  $x_{med}$  (see part 6) of the results  $x_1, x_2, \dots, x_p$  is calculated together with the absolute residuals ( $d_i$ ) of the single results from the median ( $d_i = |x_{med} - x_i|$ ). The median of the absolute residuals MAD (Median Absolute Deviation) is then calculated. The result  $x_i$  is an outlier if  $d_i > 5.06 \times MAD$  [12].

When interpreting the results of the outlier tests, the standard deviation for proficiency assessment ( $s_{pt}$ ) is taken into account. The outlier test is performed when the data consists of at least seven results.

### Robust analysis

The use of robust statistics also allows discarding of extreme results before calculating the final robust mean (see part 8, [5]).

## 3 Homogeneity test

For homogeneity testing 3–15 bottles (circa 10 % of the total amount) from the prepared sample series are used and at least one measurand is determined.

Test results are assessed by analysing the variance between groups (ANOVA), with at least two parallel analyses performed for each sample. Finally, the F-test is used to decide whether the discrepancies between the concentrations of measurand in different bottles are significant [5, 6].

#### 4 Mean

The mean value of results is calculated using the formula:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \text{ where}$$

$\bar{x}$  = the mean value of results

$x_i$  = the single result

$n$  = the number of results.

#### 5 Standard deviation

The standard deviation is the size of result distribution around the mean and is calculated using the formula:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}, \text{ where}$$

$s$  = the standard deviation

$x_i$  = the single result

$\bar{x}$  = the mean value of results

$n$  = the number of results

The standard deviation can also be expressed as a percentage (relative standard deviation).

#### 6 Median

The median is the middle result of a series arranged in order of ascending size (when  $n$  is odd number) or the mean of the two middle results (when  $n$  is even).

#### 7 Normality test

The normality of the result material is tested using the Kolmogorov-Smirnov test, where the results  $x_1, x_2, \dots, x_p$  are combined in an empirical cumulative distribution function of the  $x$  value. The number of results  $x_i$  smaller than  $x$  is calculated and normalized by dividing by the number of results  $p$ . The derived cumulative distribution is compared to the standard cumulative distribution function (the maximum deviation of these is computed and compared to the test value distribution).

#### 8 Robust mean and robust standard deviation

The robust mean is commonly used in evaluating assigned values for proficiency tests and is also recommended in international guides [5, 6]. The impact of deviations on the robust mean is theoretically smaller than on the arithmetic mean.

Although highly deviant values are commonly not discarded when computing the robust mean, their impact is reduced by down-weighting and recalculating [6]. Experience has shown, however, that the robust mean can also be affected by some extreme values (e.g.

values differing from the data more than  $5 \times s_{rob}$  or more than 50 % from the robust mean) [6]. In such cases these extreme values may be discarded before final calculation of the robust mean.

The robust mean and robust standard deviation are calculated using Algorithm A, as set out in standard ISO 13528 [5]:

The data items are sorted in increasing order:  $x_1, x_2, \dots, x_i, \dots, x_p$ .

Initial values for  $x^*$  and  $s^*$  are calculated as:

$$x^* = \text{median of } x_i \text{ (} i = 1, 2, \dots, p \text{)}$$

$$s^* = 1.483 \times \text{median of } |x_i - x^*| \text{ (} i = 1, 2, \dots, p \text{)}$$

The mean  $x^*$  and  $s^*$  are updated as follows:

Calculate  $\varphi = 1.5 \times s^*$ . A new value is then calculated for each result  $x_i$  ( $i = 1, 2, \dots, p$ ):

$$x_i^* = \begin{cases} x^* - \varphi, & \text{if } x_i < x^* - \varphi \\ x^* + \varphi, & \text{if } x_i > x^* + \varphi \\ x_i & \text{otherwise.} \end{cases}$$

The new values of  $x^*$  and  $s^*$  are calculated from:

$$x^* = \sum_{i=1}^p x_i^* / p$$

$$s^* = 1.134 \sqrt{\sum_{i=1}^p (x_i^* - x^*)^2 / (p - 1)}$$

To determine the final robust estimates  $x_{rob}$  and  $s_{rob}$  the robust mean  $x^*$  and the robust standard deviation  $s^*$  may be derived by an iterative calculation, i.e. by updating the values of  $x^*$  and  $s^*$  several times until the process converges.

## 9 Stability test

The stability of samples is tested when the analysed compound has poor stability e.g. during transport of samples (e.g. determining pH, BOD<sub>7</sub>, chlorophyll *a*). Stability is tested after keeping the samples cool (4 °C) and at room temperature (20 °C) during the period of transport. Both samples are tested and the results are processed using the difference in results obtained by analysing samples kept at different temperatures. The difference should be smaller than  $0.3 \times$  standard deviation [5, 6]:

$$D = |c_{20^\circ} - c_{4^\circ}| < 0.3 \times s_{pt}, \text{ where}$$

$c_{20^\circ}$  = the concentration after storing at 20 °C

$c_{4^\circ}$  = the concentration after storing at 4 °C

$s_{pt}$  = the standard deviation for proficiency assessment.

## 10 Uncertainty and reliability of the assigned value

The uncertainty evaluation related to the characterization of the concentrations depends on the evaluation of the assigned value. When using CRM as test sample, the uncertainty of the assigned value is taken directly from the certificate of the reference material. The uncertainty of the theoretical concentration of the synthetic sample is calculated by using GUM calculation where the uncertainties of the sample preparation steps are used and combined. When using consensus value as assigned value, uncertainty for synthetic

sample could be evaluated using robust standard deviation of the result data.

The uncertainty of an assigned value evaluated using the participant results may be estimated as follows:

If the assigned value is calculated as the mean value, then the expanded uncertainty ( $U_{pt}$ ) is calculated as a mean error at the 95 % confidence level [5]:

$$U_{pt} = 2 \cdot s / \sqrt{n}, \text{ where}$$

$s$  = the standard deviation and  $n$  = the number of the results. If the assigned value is calculated as the robust mean, then the uncertainty is calculated using the robust standard deviation at the 95 % confidence level [5]:

$$U_{pt} = 2 \cdot 1,25 \cdot s_{rob} / \sqrt{n}, \text{ where}$$

$s_{rob}$  = the robust standard deviation and  $n$  = the number of the results.

The standard uncertainty of the assigned value ( $u_{pt}$ ) is compared to the standard deviation for the proficiency assessment ( $s_{pt}$ ) with the following criterion [6]:

$$u_{pt}/s_{pt} \leq 0.3$$

The assigned value is reliable when the criterion is fulfilled. If  $0.3 < u_{pt}/s_{pt} \leq 1$ , where  $0.3 < 1 < 0.7$ , then the assigned value has high uncertainty. If  $u_{pt}/s_{pt} > 1$ , z scores will not be reported [5, 6].

When metrologically traceable result (e.g. ID-ICP-MS) is used as assigned value, the standard uncertainty of the measurement (GUM calculated) is used as the standard uncertainty of the assigned value.

## 11 z score in performance evaluation and reliability of the standard deviation for proficiency assessment

Performance for a single result is calculated as follows [4]:

$$z = \frac{(x_i - x_{pt})}{s_{pt}}, \text{ where}$$

$x_i$  = the result of the individual participant

$x_{pt}$  = the assigned value

$s_{pt}$  = the standard deviation for proficiency assessment

A result may be considered [4]:

satisfactory if  $|z| \leq 2.0$

questionable if  $2.0 < |z| < 3.0$

unsatisfactory if  $|z| \geq 3.0$ .

An example of the z scores reporting is shown in Appendix 3.

The reliability of the standard deviation for proficiency assessment and the reliability of the corresponding z score are estimated by comparing the standard deviation of test results ( $s$  tai  $s_{rob}$ ) with the standard deviation for proficiency assessment ( $s_{pt}$ ). If e.g.  $s_{rob} < 1.2 \times s_{pt}$ , then the z scores may be considered reliable [6].

## 12 zeta score and its interpretation

In the preliminary result report, the zeta values are provided for the results, for which measurement uncertainty is reported at the 95% confidence interval ( $k = 2$ ) [4]:

$$\text{zeta} = (x_i - x_{pt}) / \sqrt{u_i^2 + u_{pt}^2}, \text{ where}$$

$x_i$  = the result of individual participant

$x_{pt}$  = the assigned value

$u_i$  = the uncertainty of participant result

$u_{pt}$  = the standard uncertainty of the assigned value

If participant's reported measurement uncertainty is realistic, then the z and zeta scores are similar. Neither is the discrepancy large if the difference  $x_i - x_{pt}$  is small, in which case the result for participant will be near the assigned value. Participant performance is not evaluated on the basis of the zeta score, but the participant could use it when evaluating the measurement uncertainty.

### How to interpret these results?

| z score          | zeta score       | Action to take  |
|------------------|------------------|---|
| Satisfactory     | Satisfactory     | No action, the result is good.  |
| Satisfactory     | Not satisfactory | The claimed uncertainty is too low, but it fills the requirement of the proficiency test.   |
| Not satisfactory | Satisfactory     | The result is within your claimed uncertainty, but not within the limits of proficiency test. The uncertainty might therefore be too high and should be checked against the uncertainty requirement of your client. |
| Not satisfactory | Not satisfactory | The result is too much biased and the reason should be clarified.   |

## 13 D% values and En scores

When the number of reported results is low ( $n < 6$ ), the performance of the participant could be estimated by means of D% values ('Difference'). D% values are calculated as the difference between the participant's result and the assigned value. D% value can be interpreted as the measurement error for the results to the extent to which the assigned value can be considered the reference quantity value.

$$D_i\% = \frac{100(x_i - x_{pt})}{x_{pt}}\%, \text{ where}$$

$x_i$  = participant's result,  $x_{pt}$  = assigned value

The assessment of the D% values could be done by e.g. comparing the results with the quality guidelines or by numeric assessment.

When the number of reported results is low ( $n < 6$ ) and the uncertainty is set for the assigned value, the performance could be estimated by means of  $E_n$  scores (*'Error, normalized'*, Appendix 4). These are used to evaluate the difference between the assigned value and participant's result within their claimed expanded uncertainty.  $E_n$  scores are calculated:

$$(E_n)_i = \frac{x_i - x_{pt}}{\sqrt{U_i^2 + U_{pt}^2}}, \text{ where}$$

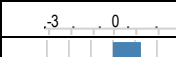
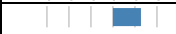







$x_i$  = participant's result,  $x_{pt}$  = assigned value,  $U_i$  = the expanded uncertainty of a participant's result and  $U_{pt}$  = the expanded uncertainty of the assigned value.

Scores of  $E_n$   $-1.0 < E_n < 1.0$  should be taken as an indicator of successful performance when the uncertainties are valid. Whereas scores  $E_n \geq 1.0$  or  $E_n \leq -1.0$  could indicate a need to review the uncertainty evaluation, or to correct a measurement issue.

### Appendix 3. Reporting results of individual participant

The interlaboratory comparison report includes a result printout for each participant specifying the z scores obtained together with the main statistically derived parameters as shown below.

#### Example of results reported separately to each participant and calculation of z score

| Participant 5               |      |        |   |         |                |                       |                      |      |      |      |     |                   |
|-----------------------------|------|--------|---|---------|----------------|-----------------------|----------------------|------|------|------|-----|-------------------|
| Measurand                   | Unit | Sample |  | z score | Assigned value | 2×s <sub>pt</sub> , % | Participant's result | Md   | Mean | s    | s%  | n <sub>stat</sub> |
| N <sub>NH4</sub>            | µg/l | B2N    |  | 1.322   | 73.3           | 15                    | 80.6                 | 73.3 | 74.1 | 3.9  | 5.3 | 26                |
| N <sub>NO2+NO3</sub>        | µg/l | B2N    |  | 0.844   | 154            | 10                    | 161                  | 153  | 153  | 5.4  | 3.5 | 25                |
| N <sub>tot</sub>            | µg/l | B2N    |  | 0.590   | 452            | 15                    | 472                  | 451  | 451  | 25.7 | 5.7 | 26                |
| pH                          |      | B2H    |  | -0.934  | 7.97           | 2.5                   | 7.88                 | 7.99 | 7.98 | 0.1  | 1.1 | 30                |
| P <sub>PO4</sub>            | µg/l | B2P    |  | -0.500  | 21.6           | 10                    | 21.1                 | 21.7 | 21.5 | 0.8  | 3.5 | 24                |
| P <sub>PO4, dissolved</sub> | µg/l | B2P    |  | 0.256   | 21.1           | 10                    | 21.4                 | 21.2 | 21.0 | 1.1  | 5.4 | 21                |
| P <sub>tot</sub>            | µg/l | B2P    |  | -1.602  | 26.6           | 10                    | 24.5                 | 26.4 | 26.6 | 2.0  | 7.7 | 24                |
| P <sub>tot, dissolved</sub> | µg/l | B2P    |  | -2.056  | 25.2           | 10                    | 22.6                 | 25.0 | 25.2 | 1.9  | 7.6 | 19                |

where:

Measurand            The tested parameter

z score                Calculated z score (satisfactory result  $-2.0 \leq z \leq 2.0$ )

Assigned value        Assigned value

2×s<sub>pt</sub> %                Standard deviation for proficiency assessment (95 % confidence level)

Participant's result    Result of an individual participant (when parallel results are reported, the mean value of those) <sup>1)</sup>

Md                      Median value

s                         Standard deviation (absolute)

s%                        Standard deviation as percent (relative)

n<sub>stat</sub>                    Number of participants in statistical processing

<sup>1)</sup> In performance evaluation, the z score is calculated from the precise result reported by the participant. In the result sheet of the report, the Participant's result might slightly differ from the reported value due to the number of visible decimals or due to rounding.

#### z score:

In the example above, the assigned value for N<sub>tot</sub> in sample B2N was 452 µg/l (= x<sub>pt</sub>) and the standard deviation for proficiency assessment s<sub>pt</sub> (2×s<sub>pt</sub> %, at the 95 % confidence level) was 15 %, thus s<sub>pt</sub> = 7.5 % of the assigned value.

The result of the participant 5 was 472 µg/l (= x<sub>i</sub>)

$$z = (x_i - x_{pt}) / s_{pt} = (472 - 452) / (0.075 \times 452) = 0.590.$$