Eurachem

A focus for analytical chemistry in Europe

Planning and Reporting Method Validation Studies

Supplement to Eurachem Guide on the Fitness for Purpose of Analytical Methods

Second edition 2025

Eurachem
A focus for analytical chemistry in Europe

Planning and Reporting Method Validation Studies

Supplement to Eurachem Guide on the Fitness for Purpose of Analytical Methods

Second edition

2025

Acknowledgements

This document has been produced by members of the Eurachem Method Validation Working Group (MVWG). The MVWG members at the time of publication are listed below. The MVWG acknowledges the input of former members to the 1st edition of this document. Production of this supplement was funded in part by the UK government Department for Science, Innovation & Technology (DSIT).

Working Group Members

Abdelkader Boubetra	BIPEA (FR)
Barbara Pohl	Gesellschaft Deutscher Chemiker (DE)
Burçu Biniçi	UME (TR)
Elvar Theodorsson	Linköping University (SE)
Emanuela Gregori	Istituto Superiore di Sanità (IT)
Ernst Halder	Eurachem Switzerland (CH)
Ewa Bulska	Eurachem Poland (PL)
Fatma Akcadag	UME (TR)
Francisco Raposo	CSIC (ES)
Helen Cantwell (Chair)	The State Laboratory (IE)
Helen Gika	Aristotle University (GR)
Isabelle Vercruysse	BELAB (BE)
John Clancy	Henkel Ireland (IE)
Karen Moore	Public Analyst Laboratory (IE)
Lorens Sibbesen	Labquality International (DK)
Marina Patriarca	Eurachem Italy (IT)
Marios Kostakis	National & Kapodistrian University of Athens (GR)
Nathalie Guigues	LNE (FR)
Pedro Pablo Morillas Bravo	Canal de Isabel II (ES)
Perihan Yolci Ömeroglu	Bursa Uludag University (TR)
Pieter Dehouck	European Commission (EU)
Piotr Robouch	European Commission (EU)
Stephan Walch	CVUA Karlsruhe (DE)
Stephen L. R. Ellison	LGC (UK)
Stig Valdersnes	Eurachem Norway (NO)
Vicki Barwick (Editor)	LGC (UK)

Recommended citation

This publication should be cited* as: "V. Barwick (ed.), Planning and Reporting Method Validation Studies – Supplement to Eurachem Guide on the Fitness for Purpose of Analytical Methods (2nd ed. 2025). Available from www.eurachem.org."**Subject to journal requirements*

Planning and Reporting Method Validation Studies

English edition

Second edition 2025

Contents

Foreword to the second edition

Following the publication of the 3rd edition of the guide Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics in February 2025, this supplementary document has been reviewed and updated. The main changes are:

- all cross references updated;
- descriptions of performance characteristics updated in line with the 3rd edition of the Fitness for Purpose Guide;
- revision of working range section in line with the 3rd edition of the Fitness for Purpose Guide;
- additional information on the extent of validation and the risk-based approach.

Foreword to the first edition

The Fitness for Purpose of Analytical Methods - A Laboratory Guide to Method Validation and Related Topics (2nd ed.) was published in 2014. Since then the Method Validation Working Group has identified areas where extra guidance would be appropriate. This extra guidance has been prepared in the form of supplementary documents. This supplementary document is not intended to be used in isolation; it should be used in conjunction with the Guide.

1 Abbreviations and symbols

The following abbreviations, acronyms and symbols occur in this supplement.

ANOVA	analysis of variance
(C)RM	(certified) reference material
ILC	interlaboratory comparison
IQC	internal quality control
LOD	limit of detection
LOQ	limit of quantification
PT	proficiency testing
%RSD	percent relative standard deviation

k	coverage factor used to calculate expanded measurement uncertainty
k_Q	multiplier used in calculating LOQ
s'_0	standard deviation used for calculating an LOD or LOQ
s_I	intermediate precision standard deviation
s_r	repeatability standard deviation

2 Introduction

This supplement is intended to serve as guidance for the planning and reporting of validation studies. The aim is to provide a clear plan for the entire validation study, covering the performance characteristics that will be studied, the target value for each performance characteristic, the materials that will be analysed, the level of replication and order of the experiments, any statistical analysis that will be used, and how the method will be judged as being fit for purpose. Note that the example planning and reporting document contained in this supplement (see section 5) should not be considered as a definitive template. A laboratory should produce its own template(s) taking into account any specific regulatory or accreditation requirements.

The planning and reporting document is structured in such a way that when the experimental work has been completed, it can be easily converted into a validation report.

The document contains the following sections:

- Title page: Includes the method title and reference, and an overview of the method status and purpose of study.
- Analytical requirement: To provide information on the required scope of the method and its application, the purpose of the study, the performance characteristics to be studied, the method performance requirements, any existing performance data and the materials available for the study.
- Performance characteristics: There is a separate section for each performance characteristic. These sections should include the detail of the validation study (the performance criteria, materials to be analysed, number and order of the measurements, how the data will be evaluated, and how the performance will be assessed).
- Summary: On completion of the validation study, provide a summary of the values and/or other information obtained for each performance characteristic and a final statement on whether the aims of the study have been achieved and whether the method is fit for purpose.
- Approval: Sign off of the validation plan and the validation report.
- Learning points: On completion of the study, highlight any key information that has arisen from the validation, such as critical steps in the method or requirements for future quality control.

The document provides guidance on how to complete each section of the validation plan. It also includes references to the relevant sections of the Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics for guidance on the number of measurements required and data analysis [1].

3 Points to consider when planning a validation study

Appendix 1 provides a checklist to assist with validation planning.

3.1 The method to be validated

Before starting a validation study a detailed written procedure (such as a standard operating procedure) describing the method to be evaluated should be available. The formal validation should be considered separately from any method development activities. It is the 'final' version of the method – after completion of method development – that is validated.

3.2 Critical steps in the method and instrument requirements

Before starting the validation study the analyst should be familiar with the method and aware of any critical steps that require particular attention. Any specific requirements relating to measuring instruments and other equipment should also be considered.

3.3 The calibration function

Although the calibration function is not a performance characteristic of a test method, ensuring that the calibration function is fit for purpose is a crucial pre-requisite to the assessment of the performance characteristics and hence to the determination of fitness for purpose of the test method. Establishing a calibration function is part of method development. However, assessing that calibration function should form part of a validation or verification study. See section 5.2 of the Fitness for Purpose Guide [1] for further information.

3.4 Supporting information

There may be existing information available which can help with planning the validation study and/or demonstrating the fitness for purpose of the method. This includes data from participation in interlaboratory comparisons (ILC), such as proficiency testing (PT) schemes, results from internal quality control (IQC) and results from previous routine use of similar methods.

3.5 Extent of the validation

One of the main issues facing laboratories when planning a validation is deciding which performance characteristics should be studied and the level of replication required. The Eurachem Guide [1] and an IUPAC Technical Report [2] provide guidance on this topic. When determining the extent of validation, the laboratory should take into account factors such as regulatory requirements or sectoral guidelines, the status of the method being validated, the laboratory's experience with similar methods, the required scope of application of the method, and the criticality of decisions based on results generated by the method. A risk-based approach can be used to assess whether the extent of validation proposed in the plan is sufficient. This involves identifying the risks associated with the plan (e.g. if a limited number of matrices and/or replicates have been studied due to technical/time constraints), the potential impact any limitations in the plan could have on the fitness-for-purpose of the method, and the risks associated with reporting unreliable results.

3.6 Order of evaluation of performance characteristics

With careful planning it is possible to obtain information on a number of performance characteristics from a single set of experiments (see for example the experimental plan outlined in Appendix 2). However, there are some characteristics which should ideally be evaluated before carrying out a full precision or bias study. Selectivity is generally studied very early in the validation process as without knowledge that the selectivity is acceptable, other performance characteristics will be of little value. In some situations it may be advantageous to carry out a ruggedness study before the full precision and bias studies as it will provide information on the critical steps in the measurement process that need to be controlled. However, regulatory requirements in some sectors (reference 3, for example) stipulate that a ruggedness study should be carried out as the final stage of the validation.

3.7 Materials to be analysed

Guidance on the types of materials (e.g. reference materials (RMs), test samples) which can be analysed is given in the sections for the individual performance characteristics. When planning the study, the scope of the

method should be taken into account. The aim is for the validation to cover a representative range of sample types in terms of matrix and analyte level. This may require the analysis of a number of different materials including certified reference materials (CRMs), spiked samples and test samples. It is important to note that matrix effects should be addressed as part of the method validation (see section 5.2.3 of the Fitness for Purpose Guide for further information). It is also important to establish how much of each material will be required during the validation to ensure that sufficient material will be available.

3.8 Experimental design

Choosing suitable experimental designs is a key part of validation planning. With appropriate planning it is possible to maximise the amount of information obtained from a particular experiment. For example, it may be possible to obtain information on more than one performance characteristic. There are a number of experimental designs which can be used in a validation study. These include:

- Simple replication: This involves making a series of measurements on a single material. It is useful for estimating precision (particularly repeatability). If a reference value is available (e.g. if the material being analysed is a CRM) the results from a simple replication study can also be used to evaluate bias.
- Linear calibration: This type of design is commonly used for instrument calibration, and studies of linearity and working range. This type of design involves observations at a range of levels (usually different analyte concentrations).
- Nested design (also known as a hierarchical design): This is an experimental design in which each level of a given factor appears in only a single level of any other factor. For example, in a study of repeatability and intermediate precision, replicate measurements obtained in a short period of time are 'nested' within days or analytical runs. Figure 1 shows an example of a single factor nested design.

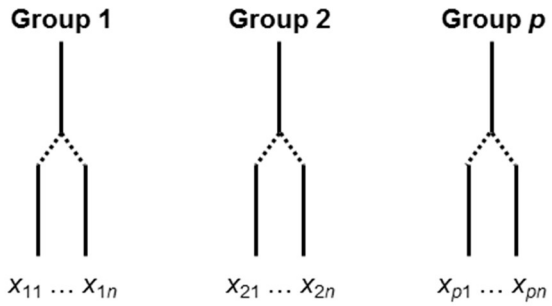

Figure 1: Example of a nested design for an experiment from which different precision measures can be evaluated if the groups represent different analytical runs (ideally carried out on different days).

This type of design is discussed in Appendix 2. The results from this type of experiment can be analysed using one-way analysis of variance (ANOVA), as described in Annex C of the Eurachem method validation guide [1].

- Fractional factorial design: This is a factorial design* from which some carefully chosen combinations of levels have been removed. This reduces the total number of measurements required in a study while still providing useful information. A fractional factorial design commonly used in method validation is a simple seven-factor design, known as a Plackett-Burman design [4]. (*Factorial designs allow the study of multiple parameters at two or more levels. A full factorial design is one in which all combinations of levels are studied.)

4 Notes on completing the validation plan for each performance characteristic

In the example planning and reporting document (section 5) each performance characteristic has a separate section where information relating to the performance criteria, planned experiments and data analysis should be documented. The information in the 'Description' section for each performance characteristic reflects the descriptions and definitions given in the 3rd edition of the Fitness for Purpose Guide. Once the study is complete the same sections can be used to summarise the data and record the outcomes of the validation. The notes below give guidance on the information to be included in each section.

4.1 Performance criteria

Specify the criteria against which the performance characteristic will be assessed (e.g. target values for precision, bias or limit of detection (LOD)).

4.2 Experiments

Outline the experiments that will be carried out to evaluate the performance characteristic. Include information on:

- The materials that will be analysed – e.g. (C)RMs, test samples, calibration standards
- The experimental design, including:
  - The number of replicate measurements that will be made on each material
  - The measurement conditions and order of analysis (e.g. if the measurements are to be made on different days, and/or by different analysts, and/or using different measuring instruments).

4.3 Evaluation of data

Outline how the data will be evaluated. Include information on:

- Any statistical parameters to be calculated from the data (e.g. mean, standard deviation)
- How values for performance characteristics are to be calculated form the data
- Any statistical tests that will be used
- How the 'fitness for purpose' of the performance characteristic will be assessed.

4.4 Notes

- Include any other information relevant to the evaluation of the performance characteristic.
- Include information on any historical performance data that may be available.
- Include information on any risk assessment carried out in relation to the proposed validation experiments.

4.5 Conclusions

- On completion of the evaluation of the performance characteristic, this section should include a statement of whether the performance criteria have been met.

5 Example planning and reporting document

Method Title

The determination of A {*analyte or measurand*} in the presence of B {*interference*} in C {*sample type/matrix*} using D {*principle*}

Include method reference number if applicable

A: What quantity is being measured?

B: Are there any known interferences that can be accommodated by the method?

C: What sample types/matrices will be analysed using the method?

D: What measurement technique/measuring instrument will be used?

Method status

Is the method, e.g. a published standard method (unmodified), based on a published standard method (with modification), a method developed in-house?

Purpose of the study

Outline the purpose of the study, e.g. to validate a new in-house method, to verify the performance of a published standard method, to validate the extension of the scope of the method.

Analytical requirement	
Analyte	Specify the analyte(s) (e.g. copper, creatinine, hexavalent chromium).
Measurand	State the measurand (the quantity intended to be measured). E.g. is it the 'total' concentration of the analyte(s) present that is of interest, the 'amount extracted' under specified conditions, or the result obtained from a specified (standard) measurement procedure? State the units in which the measurement results will be reported. State required range (e.g. concentration range in samples).
Matrix and form	State the matrix/matrices of the samples and their physical form.
Purpose of measurement	Specify why the measurements are required (e.g. to check compliance with a particular regulation or a manufacturing specification).

Purpose of the study	State the purpose of the study, for example: • Full validation of a method developed in-house • Verification of implementation of a published method for which data on performance characteristics are available • Validation of change of scope of a method • Comparison of methods to assess whether a proposed replacement method has equivalent performance to an existing method • Re-validation following change in operating conditions • Re-validation after period of non-use.
Performance characteristics	List the performance characteristics (e.g. selectivity, LOD, LOQ, precision, etc.) to be evaluated during the study. Justify any omissions (e.g. ruggedness not relevant as a published standard method is being used) including risk assessments where appropriate.
Performance requirements	How does the method need to perform to deliver results that are fit for purpose? Summarise the performance target values for the performance characteristics to be evaluated during the study. State and justify how the performance requirements were defined. Performance target values may be: • Defined in standards/regulations • Stated in a published standard method (can the stated performance be achieved?) • Related to a product specification in manufacturing quality control • Based on performance of similar procedures that are known to be fit for purpose • Based on the performance characteristics of an existing method when the purpose of the validation study is to evaluate a replacement method • Defined as the current state-of-the-art (what is the method capable of?).

Other considerations	Is there any historical data on method performance available? Is sampling/subsampling required (and will this be done within the laboratory)? Are there any restrictions on sample size or availability? Is the analyte dispersed or localised within the samples? Are there any known interferences? List any CRMs that are commercially available with a matrix and property values that are similar to the test samples. Identify any other (C)RMs that may be used during the validation study (e.g. pure substance reference materials used for preparing spiked samples).
	See section 4.6 of the Eurachem Guide [1] for further information on specifying the analytical requirement.

Performance characteristic	**Selectivity**
Description	The extent to which the method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behaviour [5].

Performance criteria	Demonstrate that other components likely to be present in the test samples do not affect the measurement results.
Experiments	• Analyse procedural and sample blanks [6]. • Analyse test samples and RMs by candidate and other independent (confirmatory) methods. • Analyse test samples containing suspected interferences and the analytes of interest • For multiple possible interferences, consider a ruggedness study to screen for the effect of a number of interferences. o Compare results for test samples with and without the interferent present to establish whether there is a significant effect on results.
Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.1 and Quick Reference 1 (Selectivity).** **Section 5.9 and Quick Reference 8 (Ruggedness).**

Notes	
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Limit of detection (LOD)**
Description	Lowest concentration of the analyte that can be detected by the method at a specified level of confidence [1]).

Performance criteria	State required LOD (this is generally expressed in the same units as the measurement results). If the concentration of the analyte in test samples is expected to be well above the LOD, an indicative value is required to demonstrate that this is the case.
Experiments	• Make replicate measurements on a suitable material (the aim is to obtain an estimate of the precision near zero): o If blank samples give a response, analyse a blank sample (a sample containing matrix components but none of the analyte of interest). o If a blank is not available, or the standard deviation of blank measurements is zero, analyse a low concentration test sample or low concentration spike. • In both cases, replicate the whole measurement procedure, including any sample preparation. Typically the measurements are made under repeatability conditions. **See section 5.3 (Limit of detection and limit of quantification) and Quick reference 3 (Limit of detection) of the Eurachem Guide [1] for guidance on the number of replicates.**
Evaluation of data	See the following section of the Eurachem Guide [1] for further information: **Section 5.3, Quick reference 3 and Annex B.**

Notes	It may also be necessary to establish the instrument LOD prior to the full LOD study, to establish the instrument's capabilities. In this case a prepared sample is analysed (i.e. only the end measurement step is replicated, not the sample preparation). For methods with a scope covering very different matrices it may be necessary to determine the standard deviation and calculate the LOD for each matrix separately. If the LOD is a critical performance characteristic it is recommended that the estimate obtained during the validation study is confirmed during routine use of the method.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Limit of quantification (LOQ)**
Description	Lowest level of analyte that can be determined with acceptable method performance [1].

Performance criteria	Typically the lower end of the working range. Demonstrate that the LOQ is compatible with the working range specified in the analytical requirement.
Experiments	Typically, LOQ calculations are based on the standard deviation estimate obtained from the LOD study.
Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.3 (Limit of detection and limit of quantification) and Quick reference 4 (Limit of quantification)**

Notes	If the LOQ is a critical performance characteristic it should be confirmed that any estimate obtained via calculation (for example using LOQ = $k_Q \times s'_0$ as described in the Eurachem Guide) is achievable. This can be done by analysing a sample in the precision study with a concentration close to the calculated LOQ.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Working range**
Description	The interval over which the method provides results with an acceptable uncertainty [1].

Performance criteria	The working range of the method should be specified in the method scope. • Demonstrate that the method can be used over the interval specified in the method scope. • Confirm that the proposed instrument calibration procedure specified in the method is adequate. o Note that establishing a calibration function is part of method development but assessing the suitability of the calibration function forms part of a validation
Experiments	Method working range • Calibrate the instrument according to the calibration procedure proposed in the candidate method. • Analyse a blank sample plus a number of reference materials or spiked samples with concentrations spaced evenly across the range of interest. • The samples can have different matrices (e.g. if there are certain sample types that would always have a low concentration of the analyte and others that would always be higher). • If possible, carry out at least duplicate measurements on each sample. • Randomise the order of analysis of the samples if possible. **See section 5.2 and Quick reference 2 (Calibration function); and 5.4 and Quick Reference 5 (Working range) of the Eurachem Guide [1] for guidance on the number of standards and level of replication.**
Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.4 and Quick Reference 5.**

Notes	If data are available from bias and precision studies that cover the range of interest, a separate method working range study may not be required.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Bias**
Description	Quantitative measure of trueness (where trueness is an expression of how close the mean of an infinite number of results is to a reference value [1]). Estimated as the difference between the mean of a set of measurement results and a reference value.

Performance criteria	State the acceptable bias, specified in terms of bias (or relative bias) or recovery.

Experiments	Evaluation of bias requires comparison of measurement results with a reference value. There are three main approaches:
	• Analysis of certified reference material(s)
	  o In cases where the measurand is defined by the method (also referred to as empirical methods) the CRM should be certified using the method being validated
	• Analysis of spiked sample(s)
	  o Analyse the unspiked matrix to confirm it is blank or establish baseline concentration
	• Comparison with alternative method
	  o Measure RM or test sample using candidate method and alternative method.
	More than one material may need to be analysed to representatively cover the scope of the method.
	Simple replication studies and nested designs are commonly used in the evaluation of bias.
	See section 6.5 and Quick reference 6 (Trueness) of the Eurachem Guide [1] for guidance on the number of replicates.

Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.6 and Quick reference 6 (Trueness)**

Notes	In general, the analysis of a CRM is the preferred approach if a suitable material is available. Comparison of results against an alternative method gives a measure of bias relative to that method. The alternative method may be a reference method or, if the intention is to replace one method with another and there is a need to demonstrate equivalent performance, a method currently in use in the laboratory. The alternative method may itself be biased, in which case the experiment will not provide an absolute measure of trueness. Where an empirical methods is being validated, comparison with a reference method is not applicable.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Precision: Repeatability**
Description	Precision: A measure of how close results from replicate measurements are to one another under specified conditions [1]. Repeatability: Measure of the variability in results when measurements are performed in a single laboratory by a single analyst using the same equipment over a short timescale [1].

Performance criteria	State target repeatability (expressed as a standard deviation s_r or relative standard deviation $\%RSD_r$).
Experiments	The following materials are suitable for precision studies: • Surplus test samples • Spiked samples • RMs More than one material may need to be analysed to representatively cover the scope of the method. When evaluating precision, a reference value is not required. Precision studies can therefore be based on the analysis of surplus test samples. Note that using RMs to estimate precision can underestimate the precision achieved for test samples. This is because RMs are usually much more homogeneous than routine test materials. However, if RMs or spiked samples are used during a precision study, it will also be possible to evaluate bias. Simple replication studies and nested designs are commonly used in the evaluation of precision. A simple replication study under repeatability conditions will provide an estimate of repeatability for the material studied. A nested design will also allow the evaluation of intermediate precision (see intermediate precision section). **See section 5.7 (Precision) and Quick Reference 7 (Repeatability, intermediate precision and reproducibility) of the Eurachem Guide [1] for guidance on the number of replicates.** **See also Appendix 2 in this document.**
Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.7, Quick Reference 7 and Annex C (Analysis of variance (ANOVA)).**

Notes	If no prior information is available about the precision of the method it is advisable to complete a limited repeatability study (simple replication) before carrying out a full repeatability/intermediate precision study.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Precision: Intermediate precision**
Description	Precision: A measure of how close results from replicate measurements are to one another under specified conditions [1]. Intermediate precision: Measure of the variability in results when measurements are made in a single laboratory but under conditions that are more variable than repeatability conditions [1] .

Performance criteria	State target intermediate precision (expressed as a standard deviation s_I or relative standard deviation $\%RSD_I$).
Experiments	Identify suitable materials –surplus test samples, spiked samples, RMs, – covering the scope of the method (analyte level and sample matrix). An efficient approach for obtaining an estimate of intermediate precision is to use a nested design. Data from such a study will provide the information required to evaluate both repeatability and intermediate precision, using one-way ANOVA. Nested design For each material: • Analytical runs (carried out under repeatability conditions) repeated on different days. • If possible, runs are made using different analysts and equipment. • A minimum of 2 replicates per material per run is required. o Number of replicates within each run will need to be increased if the number of runs is decreased (to give sufficient data for the repeatability estimate). Conversely, 2 replicates is acceptable if the number of runs is increased. • Consider preparing fresh reagents/calibrations standards, etc. between the runs. • Randomise the order of analysis of the different materials within a run if possible. **See section 5.7 (Precision) and Quick Reference 7 (Repeatability, intermediate precision and reproducibility) of the Eurachem Guide [1] for guidance on the number of replicates.** **See also Appendix 2 in this document.**
Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.7, Quick Reference 7 and Annex C (Analysis of variance (ANOVA)).**

Notes	There are many different ways of planning a nested design (number of 'groups' of data and number of replicates per group). The aim is to have sufficient data (degrees of freedom) for a reasonable estimate of the within- and between-group variation. For example, 6 groups with 3 replicates per group results in 5 degrees of freedom for the between-group variance estimate and 12 degrees of freedom for the within-group term. However, 11 groups with 2 replicates per group gives 10 degrees of freedom for the between-group variance estimate and 11 degrees of freedom for the within-group term. If the study involved different laboratories the precision estimate obtained will represent reproducibility rather than intermediate precision.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance characteristic	**Ruggedness (robustness)**
Description	Ability of a measurement procedure to maintain acceptable performance under minor changes in operating conditions [7].

Performance criteria	Identify the experimental parameters likely to vary during the application of the method that might have an effect on the measurement results. Some typical parameters are listed below: • Mass of sample • Time • Temperature • pH • Concentration/volumes of reagents. Determine whether pre-defined variations of those parameters have a significant effect on measurement results.
Experiments	To screen the effect of a number of parameters simultaneously, experimental design tools provide an efficient solution. For example, a Plackett-Burman design (a type of fractional factorial design) allows 7 parameters to be studied in 8 experiments. **See section 5.9 and Quick Reference 8 (ruggedness) of the Eurachem Guide [1] for guidance on planning ruggedness studies.**
Evaluation of data	See the following sections of the Eurachem Guide [1] for further information: **Section 5.9 and Quick Reference 8.**

Notes	A ruggedness study is not generally required for standard (published) methods or well established methods. A ruggedness study does not require a CRM (although one can be used if available). Since the ruggedness study assesses changes in results when the method parameters are varied, the exact concentration of the analyte in the sample used does not need to be known. A ruggedness study can therefore be carried out using test samples.
Conclusions	On completion of the evaluation of the performance characteristic, state whether the performance criteria have been met.

Performance parameter	**Measurement uncertainty***
Description	Parameter associated with the result of a measurement, that characterises the dispersion of values that could reasonably be attributed to the measurand [8].

Main sources of uncertainty	**Sources of information**
Produce a list of main sources of uncertainty: • Input quantities appearing in the equation used to calculate the measurement result • Other steps in the measurement procedure (e.g. sample extraction and clean-up) • Environmental conditions • Instrument parameters.	Make use of data from validation study and/or internal quality control: • Intermediate precision estimate • Bias estimate and its uncertainty • For uncertainty sources not adequately covered by precision/bias data, obtain additional information: o Manufacturer's information o Published data o Additional experiments.
Express uncertainty estimates as standard deviations (or relative standard deviations). Obtain combined standard uncertainty using the 'square root of the sum of the squares' rule. Report as expanded uncertainty – multiply combined standard uncertainty by coverage factor, k. Typically k=2 for an expanded uncertainty at a confidence level of approximately 95 %.	

Notes	Measurement uncertainty is covered in section 5.8 of the Eurachem Guide [1]. For detailed information on uncertainty estimation, see the Eurachem/CITAC guide on Quantifying uncertainty in analytical measurement [9]. *Strictly, measurement uncertainty is not a performance characteristic of a particular measurement procedure but a property of the results obtained using that measurement procedure. Measurement uncertainty is a crucial part of every measurement result and reflects the effects of the performance characteristics.
Conclusions	Include a statement on whether the measurement uncertainty is fit for purpose. Guidance on setting a target for the measurement uncertainty is available in the Eurachem/CITAC Guide, 'Setting and using target uncertainty in chemical measurement' [10].

Summary *(on completion of the study)*

Performance characteristics	Include a summary of the values/evidence obtained for each performance characteristic and a statement on whether the performance criteria have been achieved.
Comments	Include any additional comments on the validation as a whole.
Conclusion	Include a final statement on whether the aims of the study have been achieved and whether the method is considered to be fit for purpose.

Approval

Final sign-off	The validation plan should be approved before starting any experimental work. Once the study has been completed the final step is for the validation to be 'signed-off' and the method approved as fit for purpose.

Learning points from the validation *(on completion of the study)*

	It is also helpful to document any specific learning points identified during the validation. These may include: • Information on critical steps in the method • Requirements for quality control when the method is in routine use.

Appendix 1: Checklist for a validation study

Method validation should always be a planned activity. This supplement and the associated Eurachem Guide [1] provide guidance on planning and completing validation studies, with a focus on choosing suitable materials and appropriate experimental designs. In addition to these considerations, the planning process should also involve careful evaluation of the status of the laboratory with regards to its readiness for performing the study. This includes assessing the availability of staff with appropriate knowledge and experience, ensuring access to appropriate equipment and understanding the complexity of the task to be accomplished. Laboratories must take responsibility for their own validation studies and develop protocols that meet the requirements of a particular study.

The following checklist aims to help laboratories to ensure that all the key aspects have been addressed during the planning process and to identify any actions that need to be taken.

A. Analytical requirement			**Comments/actions**
A.1	Analyte specified?	☐ YES ☐ NO	
A.2	Measurand specified?	☐ YES ☐ NO	
A.3	Matrix and form of samples specified?	☐ YES ☐ NO	
A.4	Expected levels/required working range specified?	☐ YES ☐ NO	
A.5	Purpose of method well understood?	☐ YES ☐ NO	
A.6	Use of results clearly specified, including assessment of the criticality of decisions based on the results?	☐ YES ☐ NO	
A.7	Any specific regulatory requirements?	☐ YES ☐ NO	
A.9	Performance characteristics to be studied identified?	☐ YES ☐ NO	
A.10	Target values for performance characteristics stated?	☐ YES ☐ NO	
A.11	Expected frequency of use of the method known?	☐ YES ☐ NO	
A.12	Any deadline for start of routine use of method?	☐ YES ☐ NO	
Other comments/actions:			
B. Purpose of validation study			**Comments/actions**
B.1	Purpose of validation exercise stated?	☐ YES ☐ NO	
B.2	Method to be validated for use in another laboratory?	☐ YES ☐ NO	
Other comments/actions:			

C. Knowledge of selected method			**Comments/actions**
C.1	Method/similar methods well known in lab?	☐ YES ☐ NO	
C.2	Clear and unambiguous method description available (e.g. standard operating procedure)?	☐ YES ☐ NO	
C.3	Any known/foreseen critical steps?	☐ YES ☐ NO	
C.4	Any supplemental standard operating procedures required?	☐ YES ☐ NO	
C.5	Any health/safety issues?	☐ YES ☐ NO	
Other comments/actions:			
D. Specific requirements for performing the method			**Comments/actions**
D.1	Any specific requirements for sample handling/storage?	☐ YES ☐ NO	
D.2	Any specific requirements for sample preparation?	☐ YES ☐ NO	
D.3	Any specific requirements for equipment calibration?	☐ YES ☐ NO	
D.4	Any specific requirements for environmental monitoring?	☐ YES ☐ NO	
Other comments/actions:			
E. Competence for validation			**Comments/actions**
E.1	Responsible person for the study appointed?	☐ YES ☐ NO	
E.2	Analyst(s) carrying out validation familiar with the method?	☐ YES ☐ NO	
E.3	Supplementary training required?	☐ YES ☐ NO	
E.4	Supervision during validation required?	☐ YES ☐ NO	
Other comments/actions:			
F. Equipment and facilities			**Comments/actions**
F.1	Particular equipment required for sample preparation?	☐ YES ☐ NO	
F.2	Required measuring equipment available?	☐ YES ☐ NO	
F.3	Measuring equipment properly calibrated?	☐ YES ☐ NO	
F.4	Measuring equipment properly maintained?	☐ YES ☐ NO	
F.5	Facilities appropriate for the application of the method?	☐ YES ☐ NO	
F.6	Environmental conditions under control?	☐ YES ☐ NO	
Other comments/actions:			

G. Tools available for validation			**Comments/actions**
G.1	Suitable blanks available?	☐ YES ☐ NO	
G.2	RMs/CRMs available?	☐ YES ☐ NO	
G.3	Spiking of samples possible/required?	☐ YES ☐ NO	
G.4	Surplus test samples available?	☐ YES ☐ NO	
G.5	Stability of validation materials under control?	☐ YES ☐ NO	
G.6	Reference method(s) available?	☐ YES ☐ NO	
Other comments/actions:			
H. Evaluation of individual performance characteristics			**Comments/actions**
H.1	Performance target specified?	☐ YES ☐ NO	
H.2	Materials to be analysed specified and sufficient material available?	☐ YES ☐ NO	
H.3	Experimental plan defined (number of replicates, order of analysis)?	☐ YES ☐ NO	
H.4	Data analysis defined (including statistical tests)?	☐ YES ☐ NO	
H.5	Criteria for assessing fitness for purpose specified?	☐ YES ☐ NO	
Other comments/actions:			
I. Supplementary information to support assessment of method performance			**Comments/actions**
I.1	Any historical data available (e.g. IQC or results from routine application of method)?	☐ YES ☐ NO	
I.2	Possible to participate in PT during validation?	☐ YES ☐ NO	
I.3	Possible to participate in/arrange other ILC?	☐ YES ☐ NO	
Other comments/actions:			
J. Approval of validation plan			**Comments/actions**
J.1	Validation plan signed off by appropriate person?	☐ YES ☐ NO	
Other comments/actions:			
K. On completion of study			**Comments/actions**
K.1	Assessment of fitness for purpose completed for each performance characteristic and method as a whole?	☐ YES ☐ NO	
K.2	Validation report signed off?	☐ YES ☐ NO	
K.3	Final method documentation (e.g. standard operating procedure) prepared and signed off?	☐ YES ☐ NO	
K.4	Ongoing quality control requirements established?	☐ YES ☐ NO	
Other comments/actions:			

Appendix 2: Experimental plan – example of a nested experimental design

There are many ways in which experiments can be designed to provide the data required for a validation study. With careful planning it is possible to obtain data on a number of performance characteristics within a single set of experiments. The plan shown in Table 1 is an example of a nested design. The replicate measurements on each material are grouped by analytical run (1 to *p*) and are carried out under repeatability conditions (with *n* replicates per run). If the runs are carried out on different days (and by different analysts using different equipment, if possible) an estimate of repeatability and intermediate precision can be obtained. By including CRMs (1 to *q*) and/or spiked samples (1 to *z*) in the experiment it is also possible to obtain an estimate of the bias. The materials included within a study will depend on the scope of the method, the scope of the validation and the materials available. Typical materials are included as examples. The aim should be to cover the scope of the method in terms of analyte level and sample matrix.

Key to Table 1

p	number of runs
n	number of replicates within each run
Blank sample	sample containing none of the analyte(s) of interest
Sample 1…Sample *m*	surplus test samples used to evaluate precision
CRM 1…CRM *q*	certified reference materials used to evaluate bias. Can also be used to evaluate precision
Low level spiked sample	sample spiked at level close to expected/previously estimated LOQ to confirm it is achievable
Spiked sample 1…spiked sample *z*	spiked samples used to evaluate recovery (note that it will also be necessary to analyse the materials before spiking). Can also be used to evaluate precision.

The plan allows for a number of different runs (*p*) with replication (*n*) within each run. The runs should be carried across different days, but all the runs do not have to be completed on different days. For example, if 10 runs are planned, it would be acceptable to complete the study in 5 days by carrying out two runs per day. However, the factors that are varied between each run (analyst, measuring instrument, etc.) should be considered to ensure a reliable estimate of intermediate precision is obtained.

To obtain reliable estimates of the performance characteristics, between 5 and 14 degrees of freedom are desirable for the estimates of within-run and between-run variability. Data from the experimental plan shown above can be analysed using one-way ANOVA (see Appendix C of the Fitness for Purpose Guide [1]). Applying ANOVA, the degrees of freedom for the within-group term is $p(n-1)$ (assuming the same number of replicates within each group), while for the between-group term it is $p-1$. There are therefore a number of combinations of number of runs and within-run replication that will provide sufficient data.

Table 1 – Example of a nested experimental design

		Material/Performance characteristic							
Run	**Rep.**	**Blank sample** • LOD • LOQ	**Sample 1** • Repeatability • Intermediate precision	**Sample *m*** • Repeatability • Intermediate precision	**CRM 1** • Bias • Repeatability • Intermediate precision	**CRM *q*** • Bias • Repeatability • Intermediate precision	**Low level spiked sample** • Confirmation of LOQ	**Spiked sample 1** • Recovery • Repeatability • Intermediate precision	**Spiked sample *z*** • Recovery • Repeatability • Intermediate precision
1	1	Blank sample	Sample 1	Sample *m*	CRM 1	CRM *q*	Low level spiked sample	Spiked sample 1	Spiked sample *z*
1	…	Blank sample	Sample 1	Sample *m*	CRM 1	CRM *q*	Low level spiked sample	Spiked sample 1	Spiked sample *z*
1	*n*	Blank sample	Sample 1	Sample *m*	CRM 1	CRM *q*	Low level spiked sample	Spiked sample 1	Spiked sample *z*
…									
p	1	Blank sample	Sample 1	Sample *m*	CRM 1	CRM *q*	Low level spiked sample	Spiked sample 1	Spiked sample *z*
p	…	Blank sample	Sample 1	Sample *m*	CRM 1	CRM *q*	Low level spiked sample	Spiked sample 1	Spiked sample *z*
p	*n*	Blank sample	Sample 1	Sample *m*	CRM 1	CRM *q*	Low level spiked sample	Spiked sample 1	Spiked sample *z*

Bibliography

For a list of current references relating to quality in analytical measurement, please refer to the Eurachem *Reading List* available under the *Publications* section of the Eurachem website, www.eurachem.org.

1 H. Cantwell (ed.), Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics (3rd ed., 2025). Available from www.eurachem.org.

2 M. Thompson, S. L. R. Ellison, R. Wood, Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC technical report), Pure Appl. Chem., 2002, 74 (5), 835.

3 2002/657/EC Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L221/8-36, 17.8.2002.

4 W. J. Youden, E. H. Steiner, Statistical Manual of the AOAC, AOAC International, 1975, ISBN 0-935584-15-3.

5 Selectivity in analytical chemistry (IUPAC recommendations 2001), Pure Appl. Chem., 2001, 73(8), 1381

6 H. Cantwell (ed.) Blanks in Method Validation – Supplement to Eurachem Guide The Fitness for Purpose of Analytical Methods, (2nd ed,. 2025). Available from www.eurachem.org.

7 Compendium of terminology in analytical chemistry ('Orange Book), 4th edition, D. Brynn Hibbert (Ed), The Royal Society of Chemistry (2023), ISBN (print): 978-1-78262-947-4, https://doi.org/10.1039/9781788012881

8 BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM 1995 with minor corrections) (www.bipm.org) (Printed as ISO/IEC Guide 98-3:2008)

9 S. L. R. Ellison and A. Williams (eds.), Eurachem/CITAC Guide CG4: Quantifying Uncertainty in Analytical Measurement (3rd ed., 2012), ISBN 978-0-948926-30-3. Available from www.eurachem.org.

10 R. Bettencourt da Silva and A. Williams (eds.), Eurachem/CITAC Guide: Setting and Using Target Uncertainty in Chemical Measurement (1st ed., 2015), ISBN 978-989-98723-7-0. Available from www.eurachem.org.