UNODC
United Nations Office on Drugs and Crime

Guidance for the Validation of Analytical Methodology and Calibration of Equipment used for Testing of Illicit Drugs in Seized Materials and Biological Specimens

A commitment to quality and continuous improvement

Photo credits:
UNODC Photo Library

Laboratory and Scientific Section
UNITED NATIONS OFFICE ON DRUGS AND CRIME
Vienna

Guidance for the Validation of Analytical Methodology and Calibration of Equipment used for Testing of Illicit Drugs in Seized Materials and Biological Specimens

A commitment to quality and continuous improvement

UNITED NATIONS
New York, 2009

Acknowledgements

This manual was produced by the Laboratory and Scientific Section (LSS) of the United Nations Office on Drugs and Crime (UNODC) and its preparation was coordinated by Iphigenia Naidis and Satu Turpeinen, staff of UNODC LSS (headed by Justice Tettey).

LSS wishes to express its appreciation and thanks to the members of the Standing Panel of the UNODC's International Quality Assurance Programme, Dr. Robert Anderson, Dr. Robert Bramley, Dr. David Clarke, and Dr. Pirjo Lillsunde, for the conceptualization of this manual, their valuable contributions, the review and finalization of the document.*

*Contact details of named individuals can be requested from the UNODC Laboratory and Scientific Section (P.O. Box 500, 1400 Vienna, Austria).

ST/NAR/41

UNITED NATIONS PUBLICATION
Sales No. E.09.XI.16
ISBN 978-92-1-148243-0

This publication has not been formally edited.

Contents

1. Introduction

1.1 Background

The UNODC Laboratory and Scientific Section provides support to laboratories in introducing and implementing a quality management system through a number of initiatives, including the provision of reference samples of controlled substances, laboratory manuals on recommended methods, training opportunities and the International Collaborative Exercises scheme and by promoting and facilitating the exchange of information, material and data [1].

The validation of analytical methods and the calibration of equipment are important aspects of quality assurance in the laboratory. This manual deals with both of these within the context of testing of illicit drugs in seized materials and biological specimens. Further information on quality assurance is given in other UNODC manuals.

1.2 Purpose of the manual

The manual is intended to provide an introduction to the validation of analytical methods, and also the performance verification of laboratory equipment. It has been designed to provide practical guidance to national authorities and analysts in the implementation of method validation within their existing internal quality assurance programmes.

The procedures described in the manual represent a synthesis of the experience of scientists from several reputable laboratories around the world. Many professional organizations have also developed guidelines for method validation as a component of quality assurance and good laboratory practices, and these have been reviewed in preparing this manual. While there is diversity with respect to detail in method validation protocols according to their context, there is also a common thread of principle underlying all systems. In general, this manual attempts to promote and harmonize national efforts by providing internationally acceptable guidelines. Importantly, it also focuses specifically on the issue of quality assurance and good laboratory practices in drug testing laboratories. It can also serve as an educational document and as a means of encouraging laboratories to consider quality assurance matters.

1.3 Layout and terminology used in this manual

Subsequent parts are concerned with validation of analytical methods and calibration/performance verification of instrumentation and equipment. Method validation and verification aims to ensure that the results produced are fit for their intended purpose while calibration/performance verification of instrumentation and equipment is concerned with ensuring that they are performing correctly. Validation of an analytical system, often referred to as system suitability testing, is concerned with checking the performance of the combination of method and equipment in day-to-day analytical procedures.

The manual is divided into four major parts and a glossary of terms.

PART 1 gives an overview of the theory and practice of method validation and instrument calibration/performance verification.

PART 2 is intended to be a practical guide for analysts. It contains prescriptive recommendations on how to validate qualitative and quantitative methods, for both seized materials and biological specimens. These "quick start" recommendations are to assist in quickly and systematically identifying the validation requirements.

PART 3 is intended to be a practical guide for calibration/performance verification of instrumentation and equipment subdivided into procedures for different instrumentation and equipment.

PART 4 contains examples of standard operating procedures for method validation to assist the laboratory manager in preparing these documents for inclusion in the quality manual of the laboratory.

The ANNEX provides a glossary of selected terms which are particularly relevant to the topics of this manual.

1.4 Use of the manual

The suggested approaches to method validation given in this manual have been chosen on the basis of proven usefulness and value. However, while several skeleton models for method validation are provided which can, in part, be used directly, it is recommended that managers of laboratories should supervise the preparation of in-house validation procedures following the guidelines given. The final choice of the method validation system remains in the hands of the laboratory manager, who should also take responsibility for ensuring that staff comply with the prescribed procedures.

Attention is drawn to the importance of adequately trained staff where matters of quality assurance are concerned. Implementation of a written or formalized quality assurance programme, as required by an external accreditation system, can only be effectively carried out in cooperation with an informed and aware staff.

An important adjunct to the development of an internal quality assurance programme is participation in an external proficiency testing scheme and laboratories are encouraged to take part in proficiency testing programmes and ring tests such as the International Collaborative Exercises (ICE) set up by UNODC within the International Quality Assurance Programme (IQAP). Within the context of validation of analytical methods, the importance of inter-laboratory tests is highlighted below (see part 2.8.)

The Laboratory and Scientific Section would welcome observations on the contents and usefulness of the present manual. Comments may be addressed to:

Laboratory and Scientific Section
United Nations Office on Drugs and Crime
Vienna International Centre, VIC
PO Box 500
1400 Vienna
Austria

Fax: (+43-1) 26060-5967
Email: lab@unodc.org
Website: www.unodc.org

2. Validation and verification of analytical methods

2.1 Introduction: role of validation within quality assurance and good laboratory practices

Methods used in an analytical chemistry laboratory must be evaluated and tested to ensure that they produce valid results suitable for their intended purpose, i.e. they must be validated. Any laboratory which adopts UNODC recommended methods* should either revalidate them or verify them as appropriate to ensure that they work properly in its local environment. Verification involves fewer (see part 2.4 and 2.5 below) experimental operations than validation.

Any method newly introduced into a laboratory should also be documented and all analysts who will use it must receive adequate training and demonstrate their competence in the method before commencing actual casework. Commercial methods also need revalidation, or at least verification. Manufacturers' recommended procedures should be followed as closely as possible. Otherwise, if significant changes are made, full validation is necessary. If a method is modified or applied to a new situation (e.g., different sample matrix), revalidation or verification would be required depending on the extent of the modification and the nature of new situation. Revalidation would be required, for example, when a method designed to work for urine is applied to blood; verification would be required when a chromatographic column of a different nature or dimension is used. No action is required where a modification is only small, for example when a chromatographic column is changed for another of the same type.

The validation or verification of a method follows a standardized set of experimental tests which produce data relating to accuracy, precision etc. The process by which this is done should be written down as a standard operating procedure (SOP). Once methods have been validated or verified, they should be formally authorized for

*UNODC Laboratory and Scientific Section has published a series of manuals on recommended methods for testing major drugs of abuse, they are published under the symbols ST/NAR. The whole series or individual numbers are provided under request.

routine use in the laboratory by the responsible person, for example the laboratory manager [2].

A "Method Authorization Form" or similar document specified in the quality manual records the details of the method and the data on which the evaluation of the method is based, including the following:

- Title of the method
- Analyte(s)
- Sample matrix
- Scientific basis of the method
- Validation study data (accuracy, precision, selectivity, range, LOD etc.)
- Name and position of the authorizing person
- Date

Note that the SOPs for validating or verifying a method, in common with all SOPs in the laboratory quality manual, should also be authorized by the laboratory manager.

Once they have been established, it is essential that all SOPs are followed exactly. If variations are made, the variations must be documented. Any significant variations require that the method be revalidated for these new conditions. For all SOPs the last approved version should be used.

Laboratory documentation for a quality system is complex in nature, and therefore laboratories must have an appropriate document control procedure as recommended in the "Guidance for the Implementation of a Quality Management System in Drug Testing Laboratories" manual [3].

Systems proposed in the literature for the validation process may vary in several respects from these guidelines because validation is necessarily tied to the intended application. One of the benefits of these guidelines is that they have been tailored to the qualitative and quantitative analysis of controlled drug substances, either in seized materials or in biological specimens.

2.2 Evolution of a new method

Schematics for the evolution of a new method are provided in the ISO standards and other publications [4,5,6]. The following scheme is generally applicable.

Step	*Involves....*
PRELIMINARY STEPS	
1. Identify requirements	Establishing end purpose
2. Select candidate method (depends on: availability of equipment and facilities, staff expertise and requirements for staff training, regulatory requirements)	Literature search for existing method; or Identification of a similar method; or Novel approach; or Recommendation of colleagues; or Recommendation of UNODC or other authoritative organization
3. Develop method	Preliminary assessment to establish if it is capable of meeting the requirements
METHOD VALIDATION	
4. Identify the type of method (specific requirements depend on whether it is a qualitative or quantitative method and the techniques involved)	(See part 2.4)
5. Produce validation documentation	Writing up experimental/validation work (See part 2.6)
6. Write user instructions (Method SOP)	(See ST/NAR/25)
7. Obtain management authorization	(See part 2.1)
METHOD PERFORMANCE MONITORING AND REVIEW	
8. Perform quality control to monitor compliance with acceptance criteria for end purpose (see step 1)	Using traceable standards, blanks, spiked samples, control charts, etc. and external proficiency testing programmes
9. Review method and propose changes	Revalidating as appropriate Drafting revisions to SOP's
10. Obtain management authorization	Updating SOPs

2.3 Preliminary steps

The central issue to be tackled before developing a new method is to establish the purpose for which the results will be used. This will then define acceptance criteria for the performance of the method and may well define or restrict the choice of techniques. To give an example, a method for the quantitative analysis of controlled drugs in seized materials will have certain minimum requirements with respect to accuracy and precision, specificity, etc., and these requirements must be satisfied before the method can be accepted for routine use. As another example, a method for the analysis of low concentrations of drug metabolites in biological specimens may require the use of techniques with the highest sensitivity and selectivity, which may only be satisfied by gas chromatography or liquid chromatography in combination with mass spectrometry.

It is important in terms of the laboratory's human and financial resources to avoid unnecessary over-specification of the requirements for which the results will be used since this might lead to prolonged analysis times, increased costs and redundant information.

2.4 Method validation

Useful protocols for method validation in the literature have derived, amongst others, from the Current Good Manufacturing Practice, Code of Federal Regulations, Food and Drug Administration, National Drug Administration, the United States Pharmacopoeia Convention, the American Public Health Association and the International Conference on Harmonization [2]. Additionally, the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG), ENFSI, IUPAC and Eurachem/CITAC have published detailed series of recommendations [5].

Methods can be classified in a number of ways [7], but in the present instance an important distinction should always be made between qualitative and quantitative methods.

Qualitative methods for drugs testing require the following set of validation parameters to be determined:

- Specificity/selectivity
- Limit of detection (LOD)
- Precision (within the laboratory repeatability and/or within the laboratory reproducibility conditions)
- Stability

For those qualitative methods with a pre-defined threshold concentration for reporting results, the following three additional parameters should be determined:

- Linearity
- Accuracy (bias) (under within laboratory repeatability and/or within laboratory reproducibility conditions) at the threshold concentration
- Precision (under within laboratory repeatability and/or within laboratory reproducibility conditions) at the threshold concentration

Quantitative methods for drugs testing require the following set of validation parameters to be determined:

- Specificity/selectivity
- Limit of detection (LOD)
- Precision (under within laboratory repeatability and/or within laboratory reproducibility conditions)
- Linearity and working range
- Accuracy (bias) (under within laboratory repeatability and within laboratory reproducibility conditions)
- Recovery
- Uncertainty of measurement
- Stability

Additional parameters to be determined which are desirable but not essential include lower limit of quantitation (LLOQ), ruggedness and robustness. For qualitative and quantitative methods that are to be used by more than one laboratory, each laboratory should verify the method, and the inter-laboratory precision and accuracy should be determined.

2.5 Method verification

When a laboratory is implementing a method which has already been validated, there is no need to revalidate the method fully, but its performance should be verified for the minimum set of parameters listed below. Usually, verification involves determining fewer parameters and making fewer measurements for each parameter than does validation. Results of verification may differ slightly from those obtained during validation, but whether they are acceptable should be determined by the purpose for which the method will be used.

Qualitative methods for drugs testing require the following set of verification parameters to be determined:

- Specificity/selectivity if sample matrix differs from that used in the method development
- Limit of detection (LOD)
- Precision (under repeatability or reproducibility conditions)

For those qualitative methods with a pre-defined threshold concentration for reporting results, the following additional parameter should be determined:

- Accuracy (bias) at threshold concentration
- Precision at threshold concentration

Accuracy and precision should be determined at the threshold concentration.

Quantitative methods for drugs testing require the following set of verification parameters to be determined:

- Specificity/selectivity and LOD if sample matrix differs from that used in method development
- Accuracy (bias) (under repeatability or reproducibility conditions)
- Precision (under repeatability or reproducibility conditions)

2.6 Validation/verification parameters

Specificity (Selectivity)

This parameter is concerned with the extent to which other substances interfere with the identification and, where appropriate, quantification, of the analyte(s) of interest. It is a measure of the ability of the method to identify/quantify the analytes in the presence of other substances, either endogenous or exogenous, in a sample matrix under the stated conditions of the method.

Specificity is determined by adding materials which might be encountered in samples. For example, a specificity test of an immunological method for biological specimens may use potentially cross-reacting substances; a specificity test of a spot test could include potentially interfering substances which might inhibit or mask the colour reaction; a chromatographic method for the determination of concentrations of drugs of abuse in clinical samples should be free of interferences from the expected concomitantly administered therapeutic drugs. Specificity is concentration-dependent and should be determined at the low end of the calibration range. The validation

should meet the purpose of the method and ensure that the effects of impurities, cross-reacting substances, etc., which may be present in the matrix are known.

Limit of detection (LOD)

This is the lowest analyte concentration that can be detected and identified with a given degree of certainty. The LOD is also defined as the lowest concentration that can be distinguished from the background noise with a certain degree of confidence. There are several methods of estimating the LOD, all of which depend on the analysis of blank specimens and examination of the signal to noise ratio. A minimum requirement for signal to noise of 3 is widely accepted. The LOD is not a robust or rugged parameter and can be affected by minor changes in the analytical system (e.g. temperature, purity of reagents, matrix effects, instrumental conditions). It is therefore important that this parameter is always verified by laboratories adopting previously validated methods.

Precision (under Repeatability and/or Reproducibility conditions)

Precision is a measure of the closeness of the analytical results obtained from a series of replicate measurements of the same measure under the conditions of the method. It reflects the random errors which occur in a method.

Two commonly accepted sets of conditions under which precision is measured are repeatable and reproducible conditions.

Repeatability conditions occur when the same analyst analyses samples on the same day with the same instrument (e.g. gas chromatograph) or materials (e.g. spot test reagents) in the same laboratory. Any variation from these conditions (e.g. different analysts, different days, different instruments, different laboratories) represent reproducibility conditions. Precision is usually measured as the coefficient of variation or relative standard deviation of analytical results obtained from independently prepared quality control standards. Precision is concentration dependent and should be measured at different concentrations within the working range, typically at the lower, mid and upper parts. Acceptable precision at the lower concentrations is 20%. At higher concentrations better precision would be expected. These acceptance criteria may be widened in some instances, for example the analysis of autopsy samples, where matrix effects may be significant.

Linearity and working range

Traditionally, methods are described as linear when there is a directly proportional relationship between the method response and concentration of the analyte in the matrix over the range of analyte concentrations of interest (working range). The working range is predefined by the purpose of the method and may reflect only a

part of the full linear range. Acceptance criteria usually involve a Goodness of Fit test. A high correlation coefficient (r) of 0.99 is often used as criterion of linearity. However, this is not sufficient to prove that a linear relationship exists, and a method with a coefficient of determination of less than 0.99 may still be fit for purpose. These parameters are not applicable to qualitative methods unless there is a threshold concentration for reporting results.

Accuracy (bias)

This is a measure of the difference between the expectation of the test result and the accepted reference value due to systematic method and laboratory error. It is usually expressed as a percentage. Accuracy and precision together determine the total error of the analysis. Accuracy is ideally determined using Certified Reference Materials (CRMs), if available, reference methods, collaborative studies or by comparison with other methods [4].

In practice, CRMs are rarely available for drugs of abuse. For drugs of abuse in biological fluids there are the National Institute of Standards and Technology (NIST) CRMs but these do not cover a large menu of substances. As an alternative, reference standards from an authoritative organization such as UNODC, Drug Enforcement Administration (DEA) or reputable commercial provider can be used.

It is common to estimate accuracy by analysing samples spiked at three different concentrations (low, medium, high) covering the working range. The concentrations of these standards should be different from those used for preparing the calibration curves and they should be prepared from a different stock standard solution. Acceptability criteria for accuracy mirror those for precision.

Recovery

The recovery of an analyte in an assay is the detector response obtained from an amount of the analyte added to and extracted from the matrix, compared to the detector response for the true concentration of the pure authentic standard (seized materials). It may also be understood as the percentage of the drug, metabolite, or internal standard originally in the specimen that reaches the end of the procedure. In the case of biological specimens, blanks of the biological matrix once the final extracts have been obtained may be spiked with the true concentration of the pure authentic standard and then analysed. Recovery experiments should be performed by comparing the analytical results for extracted samples at three concentrations (typically those corresponding to control samples used to evaluate a method's precision and accuracy). Recovery of the analyte need not to be 100%, but the extent of recovery (of the analyte and the internal standard) should be consistent (for all concentrations tested), precise and reproducible (better than 20%)

Uncertainty of measurement [8, 9, 10]

Testing laboratories should have and apply procedures for estimating uncertainty of measurement [1]. Considering uncertainty provides assurance that results and conclusions from methods and analytical schemes are fit for purpose [11].

Metrologically, uncertainty is defined as a parameter associated with the result of measurement that characterises the dispersion of the values that could reasonably be attributed to the measurand. (Measurand: Particular quantity subject to measurement.)

In more practical terms, uncertainty can be defined as a probability or level of confidence. Any measurement we make will have some uncertainty associated with it and the uncertainty interval which we quote will be the range within which the true value lies at a certain level of confidence. Typically we use a 95% confidence interval [12].

Understanding of uncertainty is fundamental to the interpretation and reporting of results [11]. The laboratory shall at least attempt to identify all the components of uncertainty and make a reasonable estimation, and shall ensure that the form of reporting of the result does not give a wrong impression of the uncertainty.

Uncertainty of measurement comprises, in general, many components. The uncertainty is calculated by estimating the errors associated with the various stages of the analysis, e.g. pre-analytical effects, homogenization, weighing, pipetting, injection, extraction, derivatisation, recovery, calibration curves. Validation data e.g. accuracy and precision under, repeatability/reproducibility conditions already account for many of these factors and should be used.

Estimates of overall uncertainty at the 95% confidence level can be calculated using the following formula:

$$U = 2\,x\,\sqrt{u_1^{\,2} + u_2^{\,2} + u_3^{\,2}}$$

where $u_{1,}\, u_2$ etc are the individual component uncertainties.

Individual component uncertainties that are less than 20% of the highest component uncertainty have little impact on the overall uncertainty and can be omitted from the calculation.

Stability

The validation of the method should demonstrate the extent to which the analytes are stable during the whole analytical procedure, including storage before and after analysis. In general, this is carried out by comparing freshly prepared standards of known concentration with similar standards retained for different periods of time and stored under various conditions. See reference [13] and further references therein.

2.7 Method performance monitoring and review

After a method has been validated or verified, and implemented, there is a continuing need within any quality assurance system to monitor that the method is still performing within its specifications. This process of monitoring involves on-going quality control of the method with blanks, controls and calibrators, and testing of the components of the system (this is sometimes referred to as system suitability testing)[14], for example, column performance in terms of resolution and peak shape, detector response and reagent specifications. Clear control limits (e.g. the acceptable variability in detector response) should be specified for the method, together with the corrective actions which should follow if these are exceeded, including recalibration, reverification or revalidation of the method.

2.8 Inter-laboratory collaborative exercises/ proficiency tests

These studies are essential to establish the reliability and compatibility of data that need to be shared. Collaborative exercises can be used as an integral part of method validation to estimate accuracy and precision under reproducibility conditions and to determine ruggedness. Some of these exercises require that the same method is used at each location. Collaborative exercises and proficiency testing schemes can be used to monitor and compare a laboratory's performance against that of other laboratories producing equivalent data. Several external quality control assurance schemes are available for controlled drugs analysis including the UNODC ICE. See also ISO/IEC Guide 43-1 and 43-2 for the accreditation of providers of PT schemes.

2.9 Practical guidelines for validation and verification of methods

2.9.1 Seized materials—qualitative analysis

Parameter	Validation requirements	Acceptance criteria
(a) Colour tests		
Specificity/ Selectivity	Analyse the following samples under the specified test conditions and note the colour obtained in the specified time: • All controlled drugs in the group(s) of interest. All compounds from natural sources or from a synthetic preparation process usually found to be present in seized samples containing the group of drugs of interest. • All substances commonly found as diluents, excipients, etc. in the matrix in which the drug is seized. • Examples of controlled drugs from another group(s). • A range of real or simulated samples of seized materials of known composition for matrix effects. The number of test samples should be as large as possible within practical limits, but a suggested minimum number is 20.	No significant interference (masking of the test) by commonly-occurring substances. • All drugs in the target group giving negative results identified. • Fewer than 5% of real or simulated casework samples containing the analyte(s) at the minimum concentration likely to be found in casework samples encountered by the laboratory give false negatives. • The rate of false negative should be lowered to a minimum (ideally 0%) when the colour test is used as a crucial screening test and no other test will be performed following a negative screening test • Fewer than 10% of real or simulated casework samples which do not contain the target analytes give false positives.

2.9.1 Seized materials—qualitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(a) Colour tests *(Continued)*		
LOD	Analyse pure samples of a selection of drugs in the group of interest at decreasing amounts until they can no longer be detected. Determine the effects of the matrix on the LOD by spiking the substance(s) into a variety of commonly occurring matrices. The quantity of material tested should be the amount specified in the analytical method.	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in casework samples encountered by the laboratory.
Precision under repeatability and reproducibility conditions	Analyse at least 10 replicate samples of known composition at an amount between 1.25 x LOD and 2 x LOD.	No more than one sample in five (20%) should give a false negative result.
(b) Microcrystal tests		
Specificity/ Selectivity	Examine each drug in the group(s) of interest in a range of typical matrices under the specified test conditions, photograph and note the features which characterize it as the particular drug or one of a specified group of drugs.	The analytical method in terms of specificity and selectivity should be fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances)

LOD	Analyse samples of each specific drug in the class(es) of interest in a variety of commonly occurring matrices at a series of dilutions to establish the minimum concentration at which the drug can still be detected with confidence.	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in casework samples encountered by the laboratory.
Precision under repeatability and reproducibility conditions	Analyse at least 10 replicate samples of known composition at an amount between 1.25 x LOD and 2 x LOD.	No more than one sample in five (20%) should give a false negative result.
(c) Spectroscopic techniques (UV, IR, NMR, IMS, MS)		
Specificity/ Selectivity	Analyse under the specified test conditions and identify the characteristic absorptions, resonances or ions for: • Samples of all controlled drugs in the group(s) class(es) of interest • Samples of all compounds from natural sources or from a synthetic preparation process usually found to be present in seized samples containing the class of drugs of interest. • All substances commonly found in the matrix in which the drug is seized, as diluents, excipients, etc. • Examples of controlled drugs from other classes.	The analytical method in terms of specificity and selectivity should be fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances)

2.9.1 Seized materials—qualitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(c) Spectroscopic techniques *(Continued)*		
	• A range of samples of seized materials of known composition for matrix effects. The number of test samples should be as large as possible within practical limits, but a suggested minimum number is 20.	
LOD	Analyse samples of a selection of drugs in the group(s) of interest in a variety of commonly occurring matrices at a range of dilutions to establish the minimum concentration at which the drugs can still be detected with confidence.	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in casework samples encountered by the laboratory.
Precision under repeatability and reproducibility conditions	Analyse at least 10 replicate samples of known composition at an amount between 1.25 x LOD and 2 x LOD.	No more than one sample in five (20%) should give a false negative result.
(d) Thin layer chromatography		
Specificity/ Selectivity	Taking care not to overload the plate, analyse under the specified test conditions, and note the Rf values and the colour obtained in the specified time for:	The analytical method in terms of specificity and selectivity should be fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances)

	• All controlled drugs in the group(s) of interest. • All compounds from natural sources or from a synthetic preparation process usually found to be present in seized samples containing the group of drugs of interest. • All substances commonly found in the matrix in which the drug is seized, such as diluents, excipients, etc. • Examples of controlled drugs from other group(s). • Analyse mixtures of substances of similar Rf and confirm which can be identified in the presence of the others.	
LOD	Analyse samples of a selection of drugs in the group(s) of interest in a variety of commonly occurring matrices at a range of dilutions to establish the minimum concentration at which the drugs can still be detected with confidence.	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in casework samples encountered by the laboratory.
Precision under repeatability and reproducibility conditions	Determine the intra- and inter-laboratory variation in relative Rf values obtained from comparison of the sample Rf value with the Rf value from authentic standards run in parallel with the samples.	No more than one sample in five (20%) should give a false negative result.

2.9.1 Seized materials—qualitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(e) GC/HPLC/Capillary electrophoresis		
Specificity/ Selectivity	Analyse under the specified test conditions, and note the retention times obtained for: • All controlled drugs in the group(s) of interest. • All compounds from natural sources or from a synthetic preparation process usually found to be present in seized samples containing the group of drugs of interest. • All substances commonly found in the matrix in which the drug is seized, such as diluents, excipients, etc. • Examples of controlled drugs from other group(s). Analyse mixtures of substances of similar retention times and confirm which can be identified in the presence of the others.	The analytical method in terms of specificity and selectivity should be fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances)
LOD	Analyse samples of each specific drug in the class(es) of interest in a variety of commonly occurring matrices at a series of dilutions to establish the minimum concentration at which the drug can still be detected with confidence (signal to noise ratio of at least 3:1).	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in casework samples encountered by the laboratory.

Precision under repeatability and reproducibility conditions	Analyse at least 10 replicate samples of known composition at an amount between 1.25 x LOD and 2 x LOD. Determine the variation (RSD) [15] in retention times relative to the internal standard	No more than one sample in five (20%) should give a false negative result. RSD should be better than ±2%
(f) Immunoassay (semi-quantitative)		
Specificity/ Selectivity	Analyse the following using the specified extraction/ pre-treatment process: • All controlled drugs in the group(s) of interest. • All compounds from natural sources or from a synthetic preparation process usually found to be present in seized samples containing the group of drugs of interest. • All substances commonly found as diluents, excipients, etc. in the matrix in which the drug is seized. • Examples of controlled drugs from other classes. • A range of real or simulated samples of seized materials of known composition for matrix effects. The number of test samples should be as large as possible within practical limits, but a suggested minimum number is 20.	The immunoassay specifications in terms of specificity and selectivity should be fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances).

2.9.1 Seized materials—qualitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(f) Immunoassay (semi-quantitative) *(Continued)*		
LOD	Analyse samples of a selection of drugs in the class of interest in a variety of commonly occurring matrices at a range of dilutions to establish the minimum concentration at which the drugs can still be detected with confidence.	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in casework samples encountered by the laboratory.
Precision under repeatability and reproducibility conditions	Determine precision under repeatability conditions by analysing at least 10 replicate samples of known composition at an amount between 1.25 and 2x cut-off concentration.	No more than one sample in five (20%) should give a false negative result.

Note: In GC-MS, the mass spectrometer can be operated in either the repetitive full scan mode or the Selected Ion Monitoring (SIM) mode. In LC-MS-MS, the mass spectrometer can be operated in either the repetitive full scan mode or in the multiple reaction monitoring (MRM) mode.

In the repetitive full scan mode, the mass spectrum is used in conjunction with the retention time or relative retention time to identify and confirm drug(s) present. For identification purposes (relates to specificity and selectivity, accuracy) the retention time and mass spectrum of each drug in the sample are compared with those of an authentic standard analysed under identical conditions, usually in the same batch or within the same day. Alternatively, the mass spectrum can be compared with a library spectrum of the standard using a library search routine. The retention time of the suspected drug should agree closely with that of the standard if available (within a window of approximately ±2% of the retention time) and the mass spectrum should have a good visual match to that of the standard or should achieve a fit factor of 900 or more in the library search (on a scale in which a perfect fit achieves a fit factor of 1,000).

When used in the SIM/MRM mode, at least three ions/transitions are selected for each target analyte, typically including the base peak and molecular ion plus one other diagnostic ion. For identification purposes, the areas of peaks in the selected ion chromatograms at the analyte retention time should have relative intensities which match those of a standard analysed in the same batch under identical conditions, with an allowable error of approximately ±20%. Similar criteria may be used for computer-generated mass chromatograms obtained from data acquired in the repetitive full scan mode, if this mode provides adequate sensitivity, and a sufficient number (more than 12) of mass spectra across the chromatographic peak to permit peak areas to be determined with reasonable accuracy. If the mass spectrometer is operated in the chemical ionization mode, there may only be one ion present and identification of the drug will have to be on the basis of the retention time and fact that the ion is present.

2.9.2 Seized materials—quantitative analysis

Parameter	Validation requirements	Acceptance criteria
(a) Spectroscopic techniques (UV, IR, NMR, IMS, MS)		
Specificity/ Selectivity	As for qualitative analysis	As for qualitative analysis
LOQ	Analyse samples of each specific drug in the class(es) of interest in a variety of commonly occurring matrices at a series of dilutions to establish the minimum concentration at which the drug can still be quantified with confidence.	The requirements should be met for accuracy and precision.
Linearity and working range*	Analyse a blank sample and six independently prepared blanks containing the drug(s) of interest at six different concentrations evenly spaced over the range of interest.	Working range should be fit for purpose
Accuracy	Blank specimens spiked with the drug(s) of interest at three different levels (high, medium and low) should be analysed in replicate on three consecutive days. The number of replicates per concentration level and day should be at least three. The difference between the mean result and the expected result (see part 2 F) should be expressed as a percentage.	All results should fall within ±20% of the expected value at the lower concentrations and within ±15% at higher concentrations.
Precision	Compare the results obtained for each spiked blank specimen at each concentration level when determining accuracy and express the variation in terms of the RSD for each concentration.	The RSD should be better than 20% at lower concentration and better than 15% at higher concentration.

Recovery (where extraction is required)	Prepare samples of the target analyte at three different concentrations in a typical matrix. Make five replicate extracts of each sample. At the same time analyse standard solutions of the target analyte(s). The recovery is then calculated by comparing analyte spectroscopic responses, e.g. absorptions, with those of the standards Recovery % = [A1/A2] x 100 For the extracted samples: A1 = response of analyte A2 = response of standard	Recovery should be reproducible to within ±15%. Note: As indicated in part II F, the absolute percentage recovery is not critical as long as it is reproducible and provides adequate LLOQ
Uncertainty	Estimate the errors at each stage of the analytical process using the data from the validation where available and calculate the overall uncertainty (see 2.6.)	As a general guideline, uncertainty should be in the region of ±15% at LOQ; ±10% mid-range or higher.
(b) GC/HPLC/Capillary electrophoresis		
Specificity/ Selectivity	As for qualitative analysis	As for qualitative analysis
LOQ	Analyse, once each, 10 sample blanks extracted from a typical drug matrix each containing the drug at concentrations near to the minimum level (close to LOD) at which a signal is just observed to indicate its presence. Express the LOQ as ±3 or ±10s.d. of the sample blank value at the drug position.	The LOQ should be fit for purpose (i.e. if participating in an external quality control, it should fit quality objectives applied)

2.9.2 Seized materials—quantitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(b) GC/HPLC/Capillary electrophoresis *(Continued)*		
Linearity and working range*	Analyse a blank sample and six independently prepared blanks containing the drug(s) of interest at different six concentrations evenly spaced covering the range of interest.	Working range should be fit for purpose.
Precision under repeatability and reproducibility conditions	Analyse 10 independently prepared blank samples spiked with the drug(s) of interest at each of six concentrations across the working range and express the variation in terms of the standard deviation at each concentration.	RSD for the lower controls must be lower than 20% and for the other control levels better than 15%.
Accuracy	Analyse 10 independently prepared blank samples spiked with the drug(s) of interest at three different levels (high, medium and low) and express the difference between the mean result and the expected result as a percentage	Errors for the lower controls must be lower than 20% and for the other control levels better than 15%.
Recovery (where extraction is required)	Prepare samples of the target analyte at three different concentrations in a typical matrix. Make five replicate extracts of each sample. To each extract add a known amount of internal standard. At the same time analyse standard solutions of the target analyte(s) containing the same amount of internal standard. The recovery is then calculated by comparing the ratio(s) of analyte peak areas to internal standard peak area (s) for the extracted and unextracted samples. Recovery % = ([A1/A2]/[A3/A4]) x 100	Recovery should be reproducible to within ±15%. Note: As indicated in part 2 F, the absolute percentage recovery is not critical as long as it is reproducible and provides adequate LLOQ

	For the extracted samples: A1 = peak area of analyte A2 = peak area of internal standard For the standard solutions: A3 = peak area of analyte A4 = peak area of internal standard	
Uncertainty	Estimate the errors at each stage of the analytical process using the data from the validation where available and calculate the overall uncertainty (see 2.6).	As a general guideline, uncertainty should be in the region of ±15% at LOQ; ±10% mid-range or higher.

Note: In GC-MS, the mass spectrometer can be operated in either the repetitive full scan mode or the Selected Ion Monitoring (SIM) mode. In LC-MS-MS, the mass spectrometer can be operated in either the repetitive full scan mode or in the multiple reaction monitoring (MRM)mode.

In the repetitive full scan mode, the mass spectrum is used in conjunction with the retention time or relative retention time to identify and confirm drug(s) present. For identification purposes (relates to specificity and selectivity, accuracy) the retention time and mass spectrum of each drug in the sample are compared with those of an authentic standard analysed under identical conditions, usually in the same batch or within the same day. Alternatively, the mass spectrum can be compared with a library spectrum of the standard using a library search routine. The retention time of the suspected drug should agree closely with that of the standard if available (within a window of approximately ±2% of the retention time) and the mass spectrum should have a good visual match to that of the standard or should achieve a fit factor of 900 or more in the library search (on a scale in which a perfect fit achieves a fit factor of 1,000).

When used in the SIM/MRM mode, at least three ions/transitions are selected for each target analyte, typically including the base peak and molecular ion plus one other diagnostic ion. For identification purposes, the areas of peaks in the selected ion chromatograms at the analyte retention time should have relative intensities which match those of a standard analysed in the same batch under identical conditions, with an allowable error of approximately ±20%. Similar criteria may be used for computer-generated mass chromatograms obtained from data acquired in the repetitive full scan mode, if this mode provides adequate sensitivity, and a sufficient number (more than twelve) of mass spectra across the chromatographic peak to permit peak areas to be determined with reasonable accuracy. If the mass spectrometer is operated in the chemical ionisation mode, there may only be one ion present and identification of the drug will have to be on the basis of the retention time and fact that the ion is present. For quantification purposes, one of the selected ions/transitions is selected as the quantification ion/transition and the other ions/transitions serve as qualifier ions to confirm the identity of the suspected drug. The chromatograms produced for the quantification ion/transition are used in method validation in a similar way to those obtained with other GC detectors such as the flame ionization detector.

2.9.3 Biological specimens—qualitative analysis

Parameter	Validation requirements	Acceptance criteria
(a) Thin layer chromatography		
Specificity/ Selectivity	Analyse under the specified test conditions and note the Rf values for: • Standard solutions of drugs and/or metabolites in the group(s) of interest; • Blank specimens spiked with drugs and/or metabolites in the group(s) of interest; • Standard solutions of drugs from other groups. Analyse a blank matrix from at least five different sources and verify the absence of interfering substances at the Rf values of the analyte(s) of interest. If drugs or other substances have similar Rf values to any of the target analytes, analyse a mixture of them to check whether they can be resolved from the target analyte(s).	Verify that the analytical method in terms of specificity and selectivity is fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances)
LOD and cut-off value	Analyse 10 independent randomized replicates of blank extracts (from a typical drug matrix) spiked with the drug of interest at a range of concentration levels. Establish the minimum level at which the drug is consistently detected. If there is a defined threshold concentration (cut-off), the performance of the TLC method is verified by running control samples spiked with a concentration which is about 25% higher than the threshold value.	LOD should be sufficiently low for the purpose of the analysis. The method should be able to detect all target analytes at the cut-off values.

2.9.3 Biological specimens—qualitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(a) Thin layer chromatography *(Continued)*		
Precision under repeatability and reproducibility conditions	Determine the variability (RSDi) in Rf values obtained from control samples. If there is a threshold value, the RSD should be determined from samples spiked with a concentration which is about 25% higher than the threshold value.	The RSD should be better than 20%
(b) GC/HPLC/Capillary electrophoresis		
Specificity/ Selectivity	Analyse under the specified test conditions and note the retention times for: • Standard solutions of drugs and/or metabolites in the group(s) of interest; • Blank specimens spiked with drugs and/or metabolites in the group(s) of interest; • Standard solutions of drugs from other groups. Analyse a blank matrix from at least five different sources and verify the absence of interfering substances at the retention times of the analyte(s) of interest. If drugs or other substances have similar retention times to any of the target analytes, analyse a mixture of them to check whether they can be resolved from the target analyte(s).	Verify the absence of interfering substances at the retention time of the analyte (s) of interest and of the internal standard (IS). Verify that the analytical method in terms of specificity and selectivity is fit for purpose (i.e. minimal false positive rates with different matrices for the screening of controlled substances).

LOD	Analyse samples of each specific drug in the class(es) of interest in a variety of commonly occurring matrices at a series of dilutions to establish the minimum concentration at which the drug can still be detected with confidence (signal to noise ratio of at least 3:1).	LOD should be sufficiently low for the purpose of the analysis. Typically, a test might be required to detect the analyte at the minimum concentration likely to be found in specimens encountered by the laboratory.
Precision under repeatability and reproducibility conditions	Analyse at least 10 replicate samples of known composition at an amount between 1.25 x LOD and 2 x LOD. Determine the variation (RSD) in retention times relative to the internal standard.	No more than one sample in five (20%) should give a false negative result. RSD should be better than ±2%

2.9.3 Biological specimens—qualitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(c) Immunoassay (semi-quantitative)		
Specificity/ Selectivity	Immunoassays from commercial sources already provide information about the specificity and the selectivity of the method. This information does not need to be verified if the immunoassay is used only for its intended purpose. The utilization of the immunoassay with biological matrices other than those already validated by the manufacturer e.g. blood rather than urine, will require its validation, especially for matrix effects. To validate an immunoassay, analyse under the specified test conditions: • Samples spiked with controlled drugs and/or metabolites in the group(s) of interest; • Examples of controlled drugs from other classes; • Substances commonly found in the matrix in which the drug is analyzed; • At least 20 specimens known to be positive; • At least 20 specimens known to be negative from different individuals	There should be no significant cross-reactivity with other drugs or substances.
LOD	Immunoassays from commercial sources usually provide information about the LOD. If validation is necessary, e.g. to analyse for a new amphetamine derivative in the same matrix or to test for drugs in a different matrix, analyse 10 independent randomised replicates of blank specimens spiked with the drug of interest at a range of concentration levels to determine the minimum concentration which can be reliably detected	Should be substantially lower than the cut-off concentration so that both negative and positive samples can be reliably classified.

Precision under repeatability and reproducibility conditions.	For most immunoassays the decision to report the presence or absence of a given substance should not be based on LOD but on a cut-off concentration applied as the criterion for a positive. The performance of an immunoassay with a defined threshold concentration (cut-off) should be verified by running control samples in analytical batches spiked with concentrations close to the cut-off (±25% of the cut-off concentration). Intra-day precision (repeatability) can be determined by running control samples which have been spiked with a concentration of analyte which is about 25% higher than the threshold value. The RSD of these data should be calculated. Inter-day precision (intermediate precision) can be obtained by accumulating data of control samples, spiked with a concentration of analyte which is about 25% higher than the threshold value, which should be run routinely in every analytical batch. The RSD of these data should be calculated.	For repeatability (intra-day precision) and intermediate (inter-day precision) the RSD vallues should be better than ±20%. Spiked control samples must be correctly classified by the immunoassay as being either higher or lower than the cut-off level.

Note: In GC-MS, the mass spectrometer can be operated in either the repetitive full scan mode or the Selected Ion Monitoring (SIM) mode. In LC-MS-MS, the mass spectrometer can be operated in either the repetitive full scan mode or in the multiple reaction monitoring (MRM) mode.

In the repetitive full scan mode, the mass spectrum is used in conjunction with the retention time or relative retention time to identify and confirm drug(s) present. For identification purposes (relates to specificity and selectivity, accuracy) the retention time and mass spectrum of each drug in the sample are compared with those of an authentic standard analysed under identical conditions, usually in the same batch or within the same day. Alternatively, the mass spectrum can be compared with a library spectrum of the standard using a library search routine. The retention time of the suspected drug should agree closely with that of the standard if available (within a window of approximately ±2% of the retention time) and the mass spectrum should have a good visual match to that of the standard or should achieve a fit factor of 900 or more in the library search (on a scale in which a perfect fit achieves a fit factor of 1,000).

When used in the SIM/MRM mode, at least three ions/transitions are selected for each target analyte, typically including the base peak and molecular ion plus one other diagnostic ion. For identification purposes, the areas of peaks in the selected ion chromatograms at the analyte retention time should have relative intensities which match those of a standard analysed in the same batch under identical conditions, with an allowable error of approximately ±20%.

Similar criteria may be used for computer-generated mass chromatograms obtained from data acquired in the repetitive full scan mode, if this mode provides adequate sensitivity, and a sufficient number (more than twelve) of mass spectra across the chromatographic peak to permit peak areas to be determined with reasonable accuracy. If the mass spectrometer is operated in the chemical ionization mode, there may only be one ion present and identification of the drug will have to be on the basis of the retention time and fact that the ion is present.

2.9.4 Biological specimens—quantitative analysis

Parameter	Validation requirements	Acceptance criteria
(a) GC/HPLC/Capillary electrophoresis		
Specificity/ Selectivity	Analyse under the specified test conditions and note the retention times for: • Standard solutions of drugs and/or metabolites in the group(s) of interest; • Blank specimens spiked with drugs and/or metabolites in the group(s) of interest; • Standard solutions of drugs from other groups; Analyse a blank matrix from at least 10 different sources and verify the absence of interfering substances at the retention times of the analyte(s) of interest. If drugs or other substances have similar retention times to any of the target analytes, analyse a mixture of them to check whether they can be resolved from the target analyte(s).	Verify the absence of any significant level of interfering substances at the retention time of the analyte (s) of interest and of the internal standard (IS). A significant level would be equivalent to or greater than the LOQ.

2.9.4 Biological specimens—quantitative analysis *(Continued)*

Parameter	Validation requirements	Acceptance criteria
(a) GC/HPLC/Capillary electrophoresis (*Continued*)		
Linearity and working range*	Analyse a blank sample containing the drug(s) of interest at five different concentrations covering the range of interest. Concentrations must be evenly spaced.	The linearity check should confirm that the method is linear e.g. the regression coefficient should be better than 0.99 over the working range and the working range should be fit for purpose.
	At least six replicate samples should be analysed at each concentration level in order to identify and exclude any outlying values. The Grub's or the Dixon's tests are suitable for that purpose.	LOD and LOQ should be well below the lowest calibration point.
	Plot a calibration curve using the mean values for each concentration and check for linearity e.g. using linear regression analysis, to obtain the regression coefficient r^2.	
	Parameters like the limit of detection (LOD) and the limit of quantification (LOQ) may be estimated from linearity studies. Estimation can be performed by multiplying three (for LOD) or 10 times (LOQ) the ratio between the standard deviation observed for the lowest calibration level and the slope of the lineal regression.	
Accuracy	Blank specimens spiked with the drug(s) of interest at three different levels (high, medium and low) should be analysed in replicate on three consecutive days. The number of replicates per concentration level and day should be at least three. The difference between the mean result and the expected result (see part 2 F) should be expressed as a percentage.	All results should fall within ±20% of the expected value at the lower concentrations and within ±15% at higher concentrations.

Precision under reproducibility conditions	Compare the results obtained for each spiked blank specimen at each concentration level when determining accuracy and express the variation in terms of the RSD for each concentration.	The RSD should be better than 20% at lower concentrations and better than 15% at higher concentrations.
Recovery (where extraction is required)	Prepare specimens of the target analyte at three different concentrations in a blank matrix. Make five replicate extracts of each specimen. To each extract add a known amount of internal standard. At the same time analyse standard solutions of the target analyte(s) containing the same amount of internal standard. (Note: if there are known matrix effects, then extracts of the equivalent amount of a blank matrix should be made and added to the standard solutions along with the internal standard). The recovery is then calculated by comparing the ratio(s) of analyte peak areas to internal standard peak area (s) for the extracted and unextracted samples. Recovery % = ([A1/A2]/[A3/A4]) x 100 For the extracted samples: A1 = peak area of analyte A2 = peak area of internal standard For the standard solutions: A3 = peak area of analyte A4 = peak area of internal standard	Recovery should be reproducible to within ±15%. *Note*: As indicated in part 2 F, the absolute percentage recovery is not critical as long as it is reproducible and provides adequate LLOQ.
Uncertainty	Estimate the errors at each stage of the analytical process using the data from the validation where available and calculate the overall uncertainty (see part 2.6.)	As a general guideline, uncertainty should be in the region of ±25% at LOQ; ±20% mid-range or higher.

Note: In GC-MS, the mass spectrometer can be operated in either the repetitive full scan mode or the Selected Ion Monitoring (SIM) mode. In LC-MS-MS, the mass spectrometer can be operated in either the repetitive full scan mode or in the multiple reaction monitoring (MRM) mode.

In the repetitive full scan mode, the mass spectrum is used in conjunction with the retention time or relative retention time to identify and confirm drug(s) present. For identification purposes (relates to specificity and selectivity, accuracy) the retention time and mass spectrum of each drug in the sample are compared with those of an authentic standard analysed under identical conditions, usually in the same batch or within the same day. Alternatively, the mass spectrum can be compared with a library spectrum of the standard using a library search routine. The retention time of the suspected drug should agree closely with that of the standard if available (within a window of approximately ±2% of the retention time) and the mass spectrum should have a good visual match to that of the standard or should achieve a fit factor of 900 or more in the library search (on a scale in which a perfect fit achieves a fit factor of 1,000).

When used in the SIM/MRM mode, at least three ions/transitions are selected for each target analyte, typically including the base peak and molecular ion plus one other diagnostic ion. For identification purposes, the areas of peaks in the selected ion chromatograms at the analyte retention time should have relative intensities which match those of a standard analysed in the same batch under identical conditions, with an allowable error of approximately ±20%. Similar criteria may be used for computer-generated mass chromatograms obtained from data acquired in the repetitive full scan mode, if this mode provides adequate sensitivity, and a sufficient number (more than twelve) of mass spectra across the chromatographic peak to permit peak areas to be determined with reasonable accuracy. If the mass spectrometer is operated in the chemical ionization mode, there may only be one ion present and identification of the drug will have to be on the basis of the retention time and fact that the ion is present.

For quantification purposes, one of the selected ions/transitions is selected as the quantification ion/transition and the other ions/transitions serve as qualifier ions to confirm the identity of the suspected drug. The chromatograms produced for the quantification ion/transition are used in method validation in a similar way to those obtained with other GC detectors such as the flame ionization detector.

3. Calibration/performance verification of instruments and equipment

3.1 Introduction

The performance of laboratory instruments and equipment may change with time, either in the short term owing to fluctuations in the environment or, in the long term, owing to ageing of the mechanical, optical or electronic components. Slow changes may not be obvious and can lead to errors in the results obtained. In addition, performance can be affected by repairs or replacement of modules or components. It is also possible that new equipment has not been tested or checked against specifications before delivery.

Within a laboratory which maintains a comprehensive quality system, all aspects of analytical work are controlled, and these potential instrumental errors are controlled by carrying out regular preventative maintenance and calibration procedures. The way in which the performance of instruments and equipment is to be monitored (the terms performance verification [4] or performance qualification [16] are used to denote this), and the frequency of the calibration checks (calibration interval), should be stipulated in Standard Operating Procedures (SOPs).

Performance verification should be based on tests which are not specific to particular methods and which use traceable calibrators and standards, thus allowing equipment to be compared between laboratories. Performance verification is not specifically related to either screening or confirmatory methods. The calibration of instruments and equipment (e.g., wavelength calibration of an IR spectrometer, mass calibration of a GCMS) is independent of the type of sample.

Two conceptual approaches to the calibration process exist:

- The traditional approach, where all instruments and equipment are calibrated and
- The approach where calibration applies only to instruments providing physical measurements, and where the result is a direct measurement of a traceable physical parameter. For example, balances, spectrometers, thermometers, centrifuges and chronometers may be calibrated because there are traceable

standards to determine the uncertainty of the measurements. In all other cases, only performance verification of the equipment/instruments can be done in the laboratory; without an estimation of uncertainty there is no calibration.

It is up to the laboratory to decide which approach to follow.

3.2 Metrological requirements

The laboratory should be furnished with all items of sampling measurement and test equipment devices for the correct performance of tests and calibrations. Prior to use, equipment should also be checked and calibrated that meets laboratories' requirements and complies with standard specifications. The laboratory should have an established programme and procedure of the calibration of its equipment [17].

Some instrument and equipment suppliers can provide calibration certificates as part of a routine maintenance contract. Current requirements of quality assurance and good laboratory practice are that records should be kept in the log-book of each instrument of all calibration procedures and checks, and of remedial action if a check indicates that an instrument is out of calibration, as summarized in the table below.

Checklist of information to be kept in the instrument maintenance log-book [4]

Name of the equipment
Name of the manufacturer, model and/or type
Serial number
Date of receipt of equipment in laboratory
Condition when received (new, used)
Details of checks made for compliance with relevant calibration or test standard specification
Date equipment was placed in service by the laboratory
Current location in the laboratory, if appropriate
Copy of the manufacturer's operating instruction(s)
Performance criteria defined according to the requirements of the type of analyses to be carried out with this instrument
Details of maintenance carried out and records of the subsequent performance check
History of any damage, malfunction, modification or repair and records of the subsequent performance check
Frequency of checking the performance criteria

3.3 Procedures for calibration/performance verification of instruments and equipment

Calibration procedures for apparatus used in analytical chemistry are often supplied by the manufacturer, together with information on routine maintenance and the frequency at which these are carried out. The following paragraphs provide guidelines for writing and performing standard calibration procedures for commonly used instruments and equipment [13,15].

Autopipettes

Apart from calibration, routine maintenance requires regular checks on the syringe assembly, by disassembly and cleaning if necessary.

Parameter to be calibrated: volume delivered.

Method: for fixed-volume pipettes, distilled water is pipetted into a weighed container to check the volume actually delivered. Increased accuracy is obtained if the balance used to check the weight has a weighing chamber saturated with water vapour, often provided as an accessory on modern electronic balances. Variable volume pipettes should be calibrated at least at four settings: the maximum volume setting, the minimum volume setting designated by the officer in charge of the autopipette, and two or more intermediate volume settings, one of which should be below the mid-point of the range. A variable volume pipette used to dispense a fixed volume only can be calibrated at that fixed volume. Adjustments to the volume setting mechanism, if required, should be made according to the manufacturer's instructions.

Calibration interval: the rate of drift from calibration should be determined by carrying out frequent calibration checks (daily). The calibration interval can then be lengthened to a time period (normally three month intervals) appropriate for the conditions of the laboratory.

Melting point apparatus

Parameter to be calibrated: Accuracy of thermometer.

Method: The melting points of reference substances are measured at least twice.

Calibration interval: Half-yearly.

pH meters

Parameter to be calibrated: pH accuracy and linearity.

Method: Commercially-prepared buffers or standard buffers (as specified in a pharmacopoeia) are used according to the manufacturer's instructions.

Calibration interval: Daily when in use.

Ovens and heating blocks

Parameter to be calibrated: Temperature.

Method: Checked with a portable reference pyrometer or precision thermometer, which should be placed as close as possible to the oven temperature sensor.

Calibration interval: Annually, and after repairs which may affect the performance of the oven/heating block.

Water baths

Parameter to be calibrated: Temperature.

Method: Precision/reference thermometer.

Calibration interval: Quarterly, when water bath thermometer has been replaced, or when the water bath has not been used for a prolonged period (weeks or months).

Balances

Before use, balances should be checked to ensure that they are clean and level on the bench. An annual service visit by a qualified maintenance engineer is essential. At a minimum, balances used for critical weighing (i.e. where the combined uncertainties in the weighing process contribute significantly, say 10 % of the total error, to the accuracy of the overall result) should have calibration certificates. These certificates should be issued either by an external accredited body or by properly trained laboratory personnel. The certificates should be renewed yearly.

Parameter to be calibrated: Accuracy.

Method: Reference weights are used according to the manufacturer's recommendations. The user may decide as per fitness for purpose, to use weight standards prepared to stricter standards than those stated by manufacturer. A typical sequence is to check and set the zero point with nothing on the balance pan then to place a reference weight on the pan and adjust the reading to give the correct value. Please note that reference weights must be handled with great care using forceps with smooth tips, as serrated tips can result in damage to the weights. Modern electronic balances frequently have internal calibration weights and the calibration check is carried out automatically, according to a pre-set sequence created by the manufacturer or, on demand, by the user.

Calibration interval: Microbalances used for preparing reference standards should be checked daily or each time they are used, if they are not used every day. Top pan balances for reagents and less critical weights can be checked less frequently, for example weekly or monthly, but it is important to monitor the rate of drift in the first instance to determine the correct calibration interval. Calibration checks should also be carried out whenever the balance has been moved.

Refrigerators and freezers

Parameter to be calibrated: Temperature.

Method: Precision thermometer. The temperature should be maintained within a maximum of ±5 degrees of the required temperature.

Calibration interval: continuously.

Instruments for immunological methods

Many different immunological methods are available and most depend on a direct comparison with calibration standards included in each batch of test samples. Particularly important is the cut-off concentration of analyte used, and the analyst should be aware of which cut-off concentrations have been adopted by the manufacturer of the immunoassay kits in use. Calibration procedures to be followed are those specified by the manufacturer.

Notable exceptions to this general procedure are the single-test immunoassay kits (sometimes referred to as "dip-stick" tests), which may have in-built controls, but not always. In principle, these produce a "positive" or "negative" result if the concentration of the target analyte is above or below the cut-off concentration. It should be noted that there is often some degree of interpretation on the part of the operator and that, as always, a trained and experienced operator will produce more accurate and consistent results than one who has little experience of the method.

UV-Visible spectrometers

Parameter to be calibrated: Wavelength accuracy and repeatability, photometric accuracy.

Method: UV absorption wavelengths are checked with holmium and didymium filters, which should be supplied by the manufacturer. The wavelength accuracy and repeatability are checked over the entire UV-visible range. At least two spectra are run. The maximum deviation is ±1.0 nm.

Calibration interval: Annually.

Infrared spectrometers:

The procedures below are for stand-alone spectrometers. Combined instruments such as GC-FTIR should be calibrated according to the manufacturer's instructions.

Parameter to be calibrated: Resolution.

Method: The total range of the instrument is scanned using a polystyrene film. The absorption peak at 3095 nm should be resolved from that at 3080 nm and the absorption at 3020 nm should be resolved from that at 3015 nm.

Calibration interval: Quarterly.

Parameter to be calibrated: Wavelength accuracy.

Method: A polystyrene film is scanned and the accuracy of the peaks at 2852, 1602 and 1028 nm is checked [18]. The accuracy should be within ±3-5 nm in the range 4000-2000 nm and ± 1.5-2.5 nm in the range below 2000 nm.

Calibration interval: Quarterly.

Gas chromatographs

Routine maintenance operations used include checks on the septum, injector liner, gas pressures and inlet filters (e.g. oxygen scrubber, moisture trap and charcoal trap), baseline signal level and background noise. Depending on the degree of usage of the instrument, it is sensible to have a routine maintenance programme involving weekly change of the septum and injector liner (more often if large numbers of samples are analysed).

Parameter to be calibrated: Oven temperature.

Method: Checked with a portable reference pyrometer or precision thermometer, which should be placed as close as possible to the oven temperature sensor.

Calibration interval: Annually.

Parameters to be verified: Column performance (efficiency, resolution, peak shape, retention times).

Method: A set of regularly-used standard(s) is analysed. The precision of retention time(s) can be measured by injecting the standard three times or more. Peak areas can also be measured (see below under integrators). It is useful to plot parameters such as retention times/indices on a control chart.

Verification interval: Monthly.

Parameters to be verified: Detector sensitivity, baseline signal and background noise.

Method: A set of regularly-used standard(s) is analysed and compared with previous runs.

Verification interval: Monthly.

Parameters to be calibrated: Flow rates of detector gases.

Method: A bubble flow meter or calibrated electronic flow meter is used according to the manufacturer's instructions.

Calibration interval: When detector is cleaned or serviced, or the analytical column is changed, or when performance has deteriorated. Difficulty in lighting a flame ionization detector often indicates that the flow rates are incorrect.

High performance liquid chromatographs

Routine maintenance of HPLC systems includes regular changing of inlet and in-line filters and guard columns, which are usually changed when the back-pressure increases beyond acceptable limits (i.e. above the maximum column pressure). If methods are transferred between different instruments, it may be necessary to check the accuracy of some parameters such as flow rate, column temperature and gradient composition, which will affect retention times and relative retention times.

Parameter to be calibrated: Flow accuracy.

Method: The column effluent is collected in a measuring cylinder or volumetric flask over an appropriate interval.

Calibration interval: The absolute flow rate is often less important than its variation during a set of analyses, but it should be checked if a standardized, official or recommended method is being implemented.

Parameters to be calibrated: Flow repeatability and precision of injector volume.

Method: A set of regularly-used standards is injected three times or more and the precision of retention times and peak areas is measured.

Calibration interval: Monthly, or this test may form part of a daily system suitability test.

Parameter to be calibrated: Detector signal: noise ratio.

Method: A set of regularly used standards is analysed and compared with previous runs. Baseline noise is measured in intervals of 0.5-1 minute and the average is calculated. The noise is calculated using a computer programme (if supplied by the manufacturer) or graphically, by drawing two horizontal lines which enclose all observed variations and measuring the vertical distance between them. The noise level can be measured with and without solvent flow, to establish the contribution made by the solvent delivery system.

Calibration interval: Monthly.

Parameter to be calibrated: Accuracy of detector wavelength(s) (UV-visible and fluorescence detectors).

Method: UV absorption wavelengths are checked with a holmium oxide filter, supplied by the manufacturer, (and traceable to a primary standard) which has a characteristic absorption wavelength maximum at 361 nm. The wavelength accuracy and reproducibility are checked over the entire UV-visible range. The maximum (allowable) deviation is ±1.0 nm. Fluorescence emission wavelengths are usually checked using a standard, for example, quinine sulphate which has excitation peaks at 255 and 355 nm and emission peak at 455 nm.

Calibration interval: Annually.

Mass spectrometers

Mass spectrometers are tuned and calibrated in a similar manner whether they are stand-alone instruments or combined with chromatographic interfaces (GC-MS and LC-MS and their multi-sector derivations). Differences arise between quadrupole and magnetic sector instruments, especially if the latter are capable of high resolution. Most bench-top instruments are controlled directly by a computer data system, and tuning and calibration are carried out automatically. Warnings are generated by the data system if the instrument fails to achieve the pre-set performance characteristics, often mandating operator intervention, for example to clean the source.

Parameters to be calibrated: Source tuning and mass calibration.

Method: A calibration compound such as perfluorokerosene (PFK) or heptacosafluorotributylamine (perfluorotributylamine) is introduced to the spectrometer using a direct inlet device. The source is tuned using selected fragment ions to give optimum sensitivity and peak shape, and obtain peak ratios (for example, of m/z 69, 219 and 264 and 502 in the perfluorotributylamine spectrum) usually determined by the manufacturer. Spectra are recorded and compared with the reference spectrum with respect to mass assignments and relative peak intensities.

Calibration interval: Daily or immediately prior to use

Chromatographic integrators and data systems

The validation of computer systems and software is a particularly important exercise which should be carried out by the manufacturer. However, it remains the responsibility of the user to ensure that the software has been validated Formal validation of software can be carried out by the supplier on behalf of the user, but the user should carry out a formal acceptance test based on acceptance criteria for software.

Manufacturers now routinely include test and diagnostic functions in their products for system validation.

Parameter to be calibrated: Accuracy of integrated peak areas.

Method: Either an in-built test function in the chromatograph is used or else a routinely-run standard is used, and compared with previous runs.

Calibration interval: Routine standards are usually run on a daily basis. Tests of hardware function can be carried out at longer intervals, for example monthly.

4. Model standard operating procedure for validation of a new analytical method

A clearly written standard operating procedure (SOP) is required to perform the validation of a method. Several examples of SOP's for chromatographic methods have been published [19]. The following model is not universally applicable since it is not possible to create a single protocol or SOP to cover all situations. The guidelines given here are for the most commonly encountered situations.

Laboratory name			*Revision*	*Page* 1/x
Author	*Reviewer*	*Acceptor*	*Previous revision*	*Date*
File			*Code*	

Title of the SOP

e.g. Validation of a Gas Chromatographic Methods

Purpose of the SOP

This section should contain be a brief instruction of the method to be validated, including planning, performance and documentation.

Performance of validation

Insert detailed instructions of the work to be carried out to determine the validation parameters.

This could be in the form of a work plan, table, etc. as follows:

Batch	*Linearity, Accuracy*
0	36-48 samples (6 concentration levels, 6-8 replicate analysis in each level, in addtion 20 specimens for inteference analyses)
Total	36-48

This table represents a pre-validation study to determine the usefulness of the method.

Batch	*Calibration standards*	*Validation samples*						
		LLOQ		*Intermediate level*		*ULOQ*		
	(6 levels + blank)	*Accuracy*	*Accuracy*	*Recovery*	*Stability*	*Accuracy*	*Selectivity*	*Samples total*
1	6-8 replicates	3	3			3	12	27-29
2	6-8 replicates	3	3	12		3		27-29
3	6-8 replicates	3	3		8	3		23-25
4	6-8 replicates	3	3			3		15-17
5	6-8 replicates	3	3			3		15-17
Total				**139-175 samples**				

Calculation of results and interpretation

Describe the procedures how to calculate the parameters using the experimental results and acceptance criteria—see part 2.9. of this manual.

Reporting the results

Report the results of validation. The method should be described, the results for each validation parameter should be documented and conclusions should be drawn as to whether the method is fit for purpose.

Archiving the validation study data

The validation report (signed, dated and authorized) should be retained along with the validation plan and all the experimental validation data under secure storage and be readily retrievable.

References

These are the ones referred to in this SOP e.g. references describing the theory of validation.

References

1. UNODC publications: Brochure on the International Quality Assurance Programme pamphlet and Protocol for the International Collaborative Exercises. These can be downloaded from UNODC at: www.unodc.org/unodc/en/scientists/publications.html

2. D.R. Jenke, "Chromatographic Method Validation: A review of Current Practices and Procedures. I. General Concepts and Guidelines", *J. Liq. Chrom. and Rel. Technol.*, vol. 19 (1996), pp. 719-736.

3. United Nations Office on Drugs and Crime, *Guidance for the Implementation of a Quality Management System in Drug Testing Laboratories*, ST/NAR/37, 2009.

4. General criteria of competence for calibration and testing laboratories, UKAS, Teddington, UK.

5. Scientific Working Group for the Analysis of Seized Drugs, *Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) Recommendations*, 2008.

6. International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 17025:2005 *General Requirements for Competence of Testing and Calibration Laboratories.*

7. E. Prichard (ed.), Trace Analysis: A structural approach to obtaining reliable results. (Royal Society of Chemistry, Cambridge, 1996), pp 32/39.

8. Eurachem/ Cooperation on International Traceability in Analytical Chemistry (CITAC), *EURACHEM/CITAC Guide: Expression of uncertainty in qualitative testing*, 2003.

9. Eurachem/Cooperation on International Traceability in Analytical Chemistry (CITAC), *EURACHEM/CITAC Guide: Measurement uncertainty arising from sampling: A guide to methods and approaches*, 2007.

10. Eurachem/ Cooperation on International Traceability in Analytical Chemistry (CITAC), *EURACHEM/CITAC Guide CG4: Quantifying Uncertainty in Analytical Measurement*, 2nd Edition, 2000.

11. SWGDRG, *Quality Assurance/General Practices Recommendations*, 2008.

12. A.G. Rowley, *Evaluating Uncertainty for Laboratories, A Practical Handbook* (version 1.1, 2001).

13. Eurachem/Cooperation on International Traceability in Analytical Chemistry (CITAC), *EURACHEM/CITAC Guide: Guide to Quality in Analytical Chemistry*, 2002.

14. L. Huber, *Good Laboratory Practice: A primer for HPLC, CE and UV-visible spectroscopy* (Hewlett-Packard Co., publication No. 12-5091-6259E, 1993).

15. International Organization for Standardization, ISO 9000:2000 *Quality management systems—Fundamentals and vocabulary.*

16. International Organization for Standardization, ISO 9001: 2008. Quality management systems—Requirements.

17. International Organization for Standardization/International Electrotechnical Commission, ISO/IEC 17025:2005 *General Requirements for Competence of Testing and Calibration Laboratories*, paragraphs 5.5-5.6.

18. World Health Organization, *The International Pharmacopoeia: General Methods of Analysis*, vol. 1, 3rd Edition, 1979.

19. David M. Bliesner, *Validating Chromatographic Methods*, (John Wiley and Sons, 2006, p. 72).

Further information about reference documents can be obtained at (16 June 2009):

UNODC	www.unodc.org/
SWGDRUG	www.swgdrug.org/
ISO	www.iso.org/iso/home.htm
EURACHEM	www.eurachem.org/

Annex. A glossary of terms used in the validation and calibration manual

Definitions are taken from the UNODC Glossary (ST/NAR/26) plus additional terms or additional definitions (marked with an asterisk). References to the sources of definitions can be found in ST/NAR/26, if not given below.

Acceptance criteria: Conditions which must be fulfilled before an operation, process or item, such as a piece of equipment, is considered to be satisfactory or to have been completed in a satisfactory way. Specific examples are given below.

> **Acceptance criteria (for software):*** The criteria a software product must meet to complete successfully a test phase or to achieve delivery requirements.
>
> **Acceptance criteria for specimens:** Procedures for acceptance or rejection of specimens arriving at the analytical laboratory. Such procedures are focused on assessing the adequacy of chain of custody.

Accuracy (bias, trueness): Ability to get the true result [1]. For quantitative tests the accuracy expresses the closeness of agreement between the true value and the value obtained by applying the test procedure a number of times. It is affected by systematic and random errors.

> **Accuracy (of a measuring instrument):**
> "Ability of a measuring instrument to give responses close to a true value."
> Note: In this context accuracy is a qualitative concept [2].

Analyte (or target analyte): Substance to be identified or measured.

> **Surrogate analyte:** A well-characterized substance which is taken as representative of the analyte [3].

Analytical run or batch:* A complete set of analytical samples with appropriate number of standards and quality control samples for their validation. Several runs (or batches) may be completed in one day, or one run (or batch) may take several days to complete.

Analytical system (measurement system):* A complete set of measuring instruments and other equipment assembled to carry out a specified measurement task [4]. In the context of analysing controlled drugs in seized materials or biological specimens, the analytical system consists of the laboratory balance(s), pH meter, chromatograph, thin layer chromatography equipment etc. which are used by the analyst to carry out the analysis.

Arithmetic mean or average: Sum of the individual values in a set divided by the number of values.

Average: See arithmetic mean.

Batch (or analytical batch): A group of one or more samples that are analysed under conditions approaching repeatability. Usually it should contain calibrators and quality control samples in addition to the real samples to be analysed.

Biological matrix:* A discrete material of biological origin that can be sampled and processed in a reproducible manner. Examples are blood, serum, plasma, urine, faeces, saliva, sputum, and various discrete tissues

Blank: Specimen not containing the analyte.

Calibration: The set of operations which establish, under specified conditions, the relationship between values indicated by a measuring instrument or measuring system, or values represented by a material measure, and the corresponding known values of a measurand

Calibration curve: The relationship between the signal response of the instrument and various concentrations of analyte in a suitable solvent or matrix.

Calibration interval:* The frequency of specific performance tests that are made on each instrument or item of equipment as part of the laboratory's preventive maintenance programme.

Calibration range:* See range.

Calibration standard:* a biological matrix to which a known amount of analyte has been added or spiked. Calibration standards are used to construct calibration curves from which the concentrations of analytes in quality control and unknown samples are determined.

Calibrator: Pure analyte in a suitable solvent or matrix used to prepare the calibration curve. Calibrators are similar in composition to controls but must be prepared separately from them, since controls are used to check on the accuracy of the calibration curve.

Candidate method:* An analytical method which has been selected and developed for a particular analytical problem and which must be validated to show that it is fit for the intended analytical purpose before being used.

Certification: Procedure by which a certifying body gives formal recognition that the body, person or product complies with given specifications.

Certified reference material (CRM): A reference material one or more of whose property values are certified by a technical procedure, accompanied by or traceable to a certificate or other documentation which is issued by a certifying body.

Certifying body: Independent science-based organization which has the competence to grant certifications. The certifying body may or may not be accredited.

Co-chromatography:* This is a procedure in which the purified test solution prior to the chromatographic step(s) is divided into two parts and:

- one part is chromatographed as such;
- the standard analyte that is to be identified is added to the other part and this mixed solution of test solution and standard analyte is chromatographed. The amount of standard analyte has to be similar to the estimated amount of the analyte in the test solution [5].

Coefficient of variation (or relative standard deviation): A measure used to compare the dispersion or variation in groups of measurements. It is the ratio of the standard deviation to the mean, multiplied by 100 to convert it to a percentage of the average.

Collaborative studies or interlaboratory test comparisons: Organization, performance and evaluation of tests on the same or similar items or materials by two or more different laboratories in accordance with predetermined conditions. The main purpose is validation of analytical methods or establishment of reference methods.

Concentration: Amount of a substance, expressed in mass or molar units, in a unit volume of fluid or mass of solid.

Confidence level (or confidence coefficient): The measure of probability associated with a confidence interval expressing the probability of the truth of a statement that the interval will include the parameter value.

Confidence interval: A range of values which contains the true value at a given level of probability. This level of probability is called the confidence level.

Contamination:* Gain of analyte during the extraction process, in contrast to the losses usually incurred which are assessed by the recovery.

Control chart: Graphical plot of test results with respect to time or sequence of measurements, with limits drawn within which results are expected to lie when the analytical scheme is in a state of statistical control [6].

Controls: Specimens used to determine the validity of the calibration curve, that is, the linearity and stability over time of a quantitative test or determination. Controls are either prepared from the reference material (separately from the calibrators, that is, weighed or measured separately), purchased, or obtained from a pool of previously analysed specimens. Where possible, controls should be matrix-matched to specimens and calibrators.

Correction for recovery:* The recovery of analytes in a method is frequently less than 100%. If there is no internal standard (which automatically compensates for incomplete recovery) then the results of analysis must be multiplied by a correction factor to obtain the values which would have been produced if the recovery had been 100%. This implies that the recovery of the method is known, which will be true if the method has been validated, as recovery is one of the performance characteristics which is measured.

Correlation coefficient: A number showing the degree to which two variables are related. Correlation coefficients range from 0 (no correlation) to –1 or +1 (perfect correlation).

Cut-off concentration (or threshold): The concentration of a drug in a specimen used to determine whether the specimen is considered positive or negative. In some circumstances it is recommended that the cut-off concentration should be set equal to the limit of detection. See also threshold.

End determination (end-step determination):* The final step in a sequence of stages comprising an analytical method, usually involving the application of a technique to an extract or other sample preparation to produce data on the composition of the sample.

Equipment:* In general, the apparatus required for any operation [7]. More specifically, the analytical measurement hardware, for example a gas chromatograph.

Error: Something done which is considered to be incorrect or wrong.

> **Random error:** A component of the total error of a measurement which varies in an unpredictable way. This causes the individual results to fall on both sides of the average value.
>
> **Systematic error:** A component of the total error of a measurement which varies in a constant way. This causes all the results to be in error in the same sense.
>
> **Total error:** The sum of random and systematic errors.

External standard:* One prepared directly from a reference substance, for example as a stock solution or serial dilutions of the stock solution. It is not prepared in the same type of matrix as the specimens or samples for analysis and therefore there is no requirement for an extraction step prior to analysis.

False negative: A test result which states that no drug is present when, in fact, a tested drug or metabolite is present in an amount greater than a threshold or a designated cut-off value

Goodness of fit: How well a model, a theoretical distribution, or an equation matches actual data.

Instrument (instrumentation, measuring instrument):* A device intended to make a measurement, alone or in conjunction with other equipment.

Interference study:* A study to check the selectivity (or specificity) of a method by adding materials which might be encountered in specimens and which it is suspected might cause interference.

Interlaboratory studies (or interlaboratory tests comparisons): See collaborative studies.

Internal standard: The addition of a fixed amount of a known substance which is not already present as a constituent of the specimen in order to identify or quantify other components. The physico-chemical characteristics of the internal standard should be as close as possible those of the analyte. Test compound(s) (e.g. structurally similar analog, stable labelled compound) added to both calibration standards and samples at known and constant concentration to facilitate quantification of the target analyte(s).

International standard:* A standard recognized by an international agreement to serve internationally as the basis for fixing the value of all other standards of the quantity concerned.

Laboratory: Facilities where analyses are performed by qualified personal and with adequate equipment.

Laboratory manager:* The qualified individual who assumes professional, organizational, educational and administrative responsibility for the laboratory's drug testing.

Limit of detection (LOD): The smallest measured content from which it is possible to deduce the presence of the analyte with reasonable statistical certainty. The lowest concentration of an analyte that the analytical procedure can reliably differentiate from background noise. The lowest content that can be measured with reasonable statistical certainty

Limit of quantitation (LOQ)/lower limit of quantification (LLOQ): The smallest measured content from which it is possible to quantify the analyte with an acceptable level of accuracy and precision. In some laboratories the LLOQ is termed the lowest calibration concentration of the working range, as accuracy and precision of this concentration is verified in every analytical run/batch. The content equal to greater than the lowest concentration point on the calibration curve.

Linear regression: A method of describing the relationship between two or more variables by calculating a "best fitting" straight line or graph.

Linearity:* The linearity of an analytical method is its ability to elicit test results that are directly, or by means of well-defined mathematical transformations, proportional to the concentration of analytes in samples within a given range. (See also linear regression). Linearity defines the ability of the method to obtain test results proportional to the concentration of the analyte.

Matrix: Material that contains the analyte, e.g. urine or blood.

Matrix effect:* The direct or indirect alteration or interference in response of an instrument such as LC-MS/MS due to the presence of unintended analytes (for analysis) or other interfering substances in the samples.

Mean: When not otherwise specified, refers to arithmetic mean.

Measurement:* The set of operations having the object of determining a value of a quantity.

Measurement system:* See analytical system.

Measuring instrument:* See instrument.

Method (or analytical method): Detailed (defined) procedure of a technical operation for performing an analysis.

Method authorization form:* A document which certifies that an analytical method has been validated for its intended purpose in the laboratory and has been authorized for that purpose by the laboratory manager, who should sign the form.

Method validation:* Confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use of a method are fulfilled [8]. The United States Pharmacopoeia defines validation of an analytical method as the process by which it is established, by laboratory studies, that the performance characteristics of the method meet the requirements for the intended analytical application. A working definition may include the ideas that a valid method.

- is suitable (reliable) for its purpose;
- provides useful analytical data in a specific situation;

- meets the pre-determined requirements (specifications) of the analytical problem;
- has an established level of performance (accuracy, consistency, reliability);
- does what it is supposed to do.

National standard:* A standard recognized by an official national decision as the basis for fixing the value, in a country, of all other standards of the quantity concerned.

Negative: Indicates that the analyte is absent or below a designated cut-off concentration. "Not detected" is sometimes used as a synonym of negative although this is not recommended.

Organization: Companies, corporations or institutes (or part of one, e.g. laboratory) private or public, that has its own functions and administration. Some of the international organizations dealing with quality assurance are: ISO, IUPAC, IOC, TIAFT, IFCC, IPCS, OECD.

Performance characteristics:* These are key aspects of an analytical method which are evaluated for the purposes of method development and validation, including accuracy (bias), linearity, limit of detection, limit of quantitation, range, recovery, repeatability, reproducibility, ruggedness, and specificity (selectivity).

Performance qualification:* See performance verification.

Performance verification (or performance qualification):* A formal and nationally traceable method of evaluating the performance of an instrument against previously defined procedures and specifications. Performance verification should involve the use of tests which are not method-specific and which use Nationally-traceable calibrators and standards

Positive: Indicates that the analyte is present at a level above a designated cut-of concentration.

Practicability:* The ability to put something into practice. In the laboratory, this means the absence of unnecessarily sophisticated equipment, reagents, instruments, or environmental conditions, so that a method is suitable for routine use [9].

Precision: The closeness of agreement (degree of scatter) between independent test results obtained under prescribed conditions. It is generally dependent on analyte concentration, and this dependence should be determined and documented. The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Higher imprecision is reflected by a larger standard deviation. Independent test results means results obtained in a manner not influenced by any previous results on the same or similar material. Precision covers repeatability and reproducibility.

A measure for the reproducibility of measurements within a set, that is, of the scatter or dispersion of a set about its central value.

Precision (intermediate):* Precision measured in-between repeatability and reproducibility conditions: for example precision measured between different analysts, over extended timescales, within a single laboratory. Expresses within laboratory variation: different days, different analysts, different equipment, etc.

Primary standard:* A standard which has the highest metrological qualities in a specified field.

Probability: A mathematical measurement of how likely it is that something will happen, expressed as a fraction or percentage. Values for statistical probability range from 1 or 100% (always) to 0 or 0% (never). The relative frequency obtained after a long run of measurements or results will give good approximations to the true probabilities. Also, it is a basic concept which may be taken either as indefinable, expressing in some way a "degree of belief", or as the limiting frequency of an occurrence in an infinite random series.

Procedure: A specified way to perform an activity. For quality assurance purposes the procedures should be written. Specified way to carry out an activity or process.

Quality assurance: All the planned and systematic activities implemented within the quality system, to provide adequate confidence that a laboratory will fulfil requirements for quality. Part of quality management focused on providing confidence that quality requirements will be fulfilled.

Quality control: The overall system of laboratory procedures and processes which controls the quality of the laboratory's analytical results.

Quality manual: A document stating the general quality policies, procedures and practices of an organization [10]. Document specifying the quality management system of an organization.

Quantitation (quantification) range:* The range of concentrations, including ULOQ and LLOQ, that can reliably and reproducibly quantified with accuracy and precision through the use of a concentration-response relationship (see also range).

Random error: See error.

Range (working range, calibration range): Concentration interval for which acceptable accuracy and precision can be achieved. Statistically it is the difference between the minimum and the maximum values of a set of measurements.

Recovery: The percentage of the drug, metabolite, or internal standard originally in the specimen that reaches the end of the procedure.

Term used in analytical and preparative chemistry to denote the fraction of the total quantity of a substance recoverable following a chemical procedure. It is measured by adding a known amount of analyte to a blank matrix and comparing this with the amount measured as present by analysis.

Reference method: One which is developed by organizations or groups that use collaborative studies or similar approaches to validate them. Its value depends on the authority of the organizations which sponsor it.

Reference standard: A standard, generally of the highest quality available at a given location, from which measurements made at that location are derived.

Reliability: The extent to which an experiment, test, or measuring procedure yields accurate results on repeated trials.

Repeatability (or repeatable): The closeness of the agreement between the results of successive measurements of the same analyte made under repeatable conditions, e.g. same method, same material, same operator, same laboratory, narrow time period. Results should be expressed in terms of the repeatability standard deviation, repeatability coefficient of variation, or the confidence interval of the mean value. Closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement.

Replicability:* See replicate analysis.

Replicate analysis: The multiple analysis of separate portions of a test material using the same test method under the same conditions e.g. same operator, same equipment, same laboratory.

Reproducibility (within laboratory): The closeness of the agreement between the results of successive measurements of the same analyte in identical material made by the same method under different conditions, e.g. different operators, different laboratories, long time period. Results should be expressed in terms of the reproducibility standard deviation; reproducibility coefficient of variation, or the confidence interval of the mean value. It also represents precision of the method under the same operating conditions over a short period of time. Closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement.

Robustness:* The ability of a method to remain unaffected by small but deliberate variations in the main parameters of the method. The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small but deliberate variations in method parameters and provides an indication of its reliability during normal usage.

Ruggedness:* The ability of a measurement process to withstand small uncontrolled or unintentional changes in its operating conditions. The ruggedness of an analytical

method is the degree of reproducibility of test results obtained by the analysis of the same samples under a variety of conditions, such as different laboratories, analysts, instruments, lots of reagents, elapsed assay times, assay temperatures or days. Ruggedness is normally expressed as the lack of influence on the test results of operational and environmental variables of the analytical method. Ruggedness is a measure of reproducibility of test results under the variation in conditions normally expected from laboratory to laboratory and from analyst to analyst.

Ruggedness test:* An intralaboratory experimental plan, used before undertaking an interlaboratory study, to examine the behaviour of an analytical process when small changes in the environmental and/or operating conditions are made, akin to those likely to arise in different laboratories.

Selectivity (or specificity): Refers to the extent to which a method can determine particular analyte(s) in a complex mixture without interference from the other components in the mixture. A method which is perfectly selective for an analyte or group of analytes is said to be specific. The term specific (in analysis) is considered as the ultimate of selectivity.

Qualitative: The extent to which other substances interfere with the determination of a substance according to a given procedure; quantitative: a term use din conjunction with another substantive (e.g. constant coefficient, index, factor, number) for the quantitative characterization of interferences.

Sensitivity: *(a)* The difference in analyte concentration corresponding to the smallest difference in the response of the method that can be detected. It is represented by the slope of the calibration curve. It is also equivalent to three time the average background reading produced by blank samples from as many different sources as possible (5 minimum, but 20 different sources are ideal). Sometimes sensitivity is erroneously used instead of limit of detection.

(b) The incidence of true positive results obtained when a test is applied to samples known to contain the analyte [11].

(c) The change in the response of a measuring instrument divided by the corresponding change in the stimulus

Specification: Statement of requirements, usually in written form.

Specificity: *(a)* See selectivity.

(b) Incidence of true negative results obtained when a test is applied to samples known not to contain the analyte.

(c) The ability of a method to measure only what is applied to samples. “Specificity is the ability to assess unequivocally the analyte in the presence of components which may be expected to be present. Typically these might include impurities, degradants, matrix etc.”

Specimen: Any material for examination, study or analysis.

Spiked sample: A test material containing a known addition of analyte.

Stability (of sample during the analysis): Resistance to decomposition or other chemical changes, or to physical disintegration.

Standard analyte:* A well-defined substance in its highest available purity to be used as a reference in the analysis.

Standard deviation: A statistic that shows the spread or dispersion of results in a distribution of results. It is calculated by taking the square root of the variance. It is applicable to all kinds of repeated measurements, e.g. between batch, within batch, repeatability, reproducibility, etc.

Standard operating procedure (SOP): Written procedures which describe how to perform certain laboratory activities.

System suitability test:* Validation of an analytical system (system suitability testing) tests a system against documented performance specifications, for a specific analytical method [12].

Systematic error: See error.

Technique:* A technique is a scientific principle, for example gas chromatography or ultraviolet spectrometry, that can be used to provide data on the composition of a material. It is unusual to apply a technique directly to a test sample, as extraction and other steps are frequently required. A technique is therefore used in the last step of an analytical method which is usually the end-determination or end-step determination.

Test: A technical operation to determine one or more characteristics of or to evaluate the performance of a given product, material, equipment, organism, physical phenomenon, process or service according to a specified procedure.

Theoretical probability distribution: The number of times it can be expected to get a particular number of successes in a large number of trials. Important theoretical probability distributions are the normal, t-, chi-square and F-distributions.

Threshold: A particular, significant amount, level, or limit, at which something begins to happen or take effect. See also Cut-off concentration.

Traceable: See traceability.

Traceable standard:* A reference standard which also has the property of traceability. It will usually have a certificate of analysis giving details of the national or international standards used to determine its composition.

Traceability: The ability to trace the history, application or location of an entity by means of recorded identification. The property of a result of a measurement whereby it can be related to appropriate standards, generally international or national standards, through an unbroken chain of comparisons. Property of the result of a measurement or the value of a standard whereby it can be related with a stated uncertainty, to stated references, usually national or international standards (i.e. through an unbroken chain of comparisons. Ability to trace the history, application or location of that which is under consideration.

True value: See value.

Uncertainty: A parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the analyte. An estimate attached to a test result which characterizes the range of values within which the true value is asserted to lie.

Upper limit of quantification (ULOQ): the highest amount of an analyte in a sample that can be quantitatively determined with precision and accuracy.

Validation: Confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled. See also method validation.

Value: The expression of a quantity in terms of a number and an appropriate unit of measurement.

True value: The value which characterizes a quantity perfectly defined in the conditions which exist when that quantity is considered. The true value of a quantity is an ideal concept and, in general, cannot be known exactly. Value consistent with the definition of a given particular quantity.

Verification: Confirmation by examination and provision of objective evidence that specified requirements have been fulfilled. Confirmation, through the provision of objective evidence, that specified requirements have been fulfilled.

Working range:* See range.

Bibliography

National Institute for Drug Abuse, *Urine Testing for Drugs of Abuse*, Research Monograph 73 (Rockville, Maryland, Department of Health and Human Services, 1986).

IUPAC, *Compendium of Analytical Nomenclature, The Orange Book*—3rd Edition, J. Inczedy, T. Lengyel, and A.M. Ure, Blackwell Science, 1998 [ISBN 0-632-05127-2], available on-line at http://old.iupac.org/publications/analytical_compendium/

IOC/Reference Materials Committee of ISO, "Quality control of analytical data produced in chemical laboratories", Publication 271, draft protocol presented to the Fifth International Symposium on the Harmonization of Internal Quality Assurance Schemes for Analytical Laboratories, Washington, D.C., 23 July 1993.

International Organisation for Standardisation, *International Vocabulary of Basic and General Terms Used in Metrology* (Geneva, 1984).

European Community, Guideline Criteria for Reference Methods, BNL SP/Lab/div (92) 5 (1992), p. 27.

IOC/REMCO N 271: Quality Control of Analytical Data Produced in Chemical Laboratories, presented at the Fifth International Symposium on the Harmonization of Internal Quality Assurance Schemes for Analytical Laboratories, 23 July, 1993, Washington D.C.

Chambers English Dictionary, W and R Chambers Ltd., Edinburgh (1990).

International Organization for Standardization/Development Information System 8402, Quality Management and Quality Assurance Vocabulary (Geneva, 1991).

G.T. Wernimont in W. Spendley (Ed.), *Use of Statistics to Develop and Evaluate Methods, Association of Official Analytical Chemists*, Arlington, VA, p. 78-82 (1985).

Quality Assurance: the Route to Efficiency and Competitiveness, 3rd Edition, L. Stebbing, Ellis Horwood (1993).

R.S. Galen and S.R. Gambino, *Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses*, John Wiley and Sons, New York, 1975.

L. Huber, *Validation of Computerized Analytical Systems*, Interpharm Press Inc., Buffalo Grove, IL, 1996.

UNODC
United Nations Office on Drugs and Crime

Vienna International Centre, PO Box 500, 1400 Vienna, Austria
Tel.: (+43-1) 26060-0, Fax: (+43-1) 26060-5866, www.unodc.org

United Nations publication
Printed in Austria

Sales No. E.09.XI.16
ST/NAR/41

V.09-84578—October 2009—350

USD 15
ISBN 978-92-1-148243-0