DRAFT FOR DEVELOPMENT

DD ENV ISO 13530:1999

Water quality — Guide to analytical quality control for water analysis

ICS 13.060.01

National foreword

This Draft for Development is the English language version of ENV ISO 13530:1998.

This publication is not to be regarded as a British Standard.

It is being issued in the Draft for Development series of publications and is of a provisional nature. It should be applied on this provisional basis, so that information and experience of its practical application may be obtained.

Comments arising from the use of this Draft for Development are requested so that UK experience can be reported to the European organization responsible for its conversion into a European Standard. A review of this publication will be initiated 2 years after its publication by the European organization so that a decision can be taken on its status at the end of its three-year life. The commencement of the review period will be notified by an announcement in *Update Standards*.

According to the replies received by the end of the review period, the responsible BSI Committee will decide whether to support the conversion into a European Standard, to extend the life of the prestandard or to withdraw it. Comments should be sent in writing to the Secretary of BSI Subcommittee EH/3/-/1, Precision and accuracy, at 389 Chiswick High Road, London W4 4AL, giving the document reference and clause number and proposing, where possible, an appropriate revision of the text.

A list of organizations represented on this subcommittee can be obtained on request to its secretary.

Cross-references

The British Standards which implement international or European publications referred to in this document may be found in the BSI Standards Catalogue under the section entitled "International Standards Correspondence Index", or by using the "Find" facility of the BSI Standards Electronic Catalogue.

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, the ENV ISO title page, page 2, the ISO/TR title page, pages ii to iv, pages 1 to 54 and a back cover.

This standard has been updated (see copyright date) and may have had amendments incorporated. This will be indicated in the amendment table on the inside front cover.

This Draft for Development, having been prepared under the direction of the Health and Environment Sector Committee, was published under the authority of the Standards Committee and comes into effect on 15 February 1999

ISBN 0 580 30967 3

Amendments issued since publication

Amd. No.	Date	Comments

Contents

blank

EUROPEAN PRESTANDARD
PRÉNORME EUROPÉENNE
EUROPÄISCHE VORNORM

ENV ISO 13530

October 1998

ICS 13.060.01

Descriptors: Water, quality, quality control, water tests

English version

Water quality — Guide to analytical quality control for water analysis

(ISO/TR 13530:1997)

Qualité de l'eau — Guide de contrôle qualité analytique pour l'analyse de l'eau (ISO/TR 13530:1997)

Wasserbeschaffenheit — Richtlinie zur analytischen Qualitätssicherung in der Wasseranalytik (ISO/TR 13530:1997)

This European Prestandard (ENV) was approved by CEN on 5 September 1998 as a prospective standard for provisional application.

The period of validity of this ENV is limited initially to three years. After two years the members of CEN will be requested to submit their comments, particularly on the question whether the ENV can be converted into a European Standard.

CEN members are required to announce the existence of this ENV in the same way as for an EN and to make the ENV available promptly at national level in an appropriate form. It is permissible to keep conflicting national standards in force (in parallel to the ENV) until the final decision about the possible conversion of the ENV into an EN is reached.

CEN members are the national standards bodies of Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

CEN

European Committee for Standardization
Comité Européen de Normalisation
Europäisches Komitee für Normung

Central Secretariat: rue de Stassart 36, B-1050 Brussels

Ref. No. ENV ISO 13530:1998 E

Foreword

The text of the International Technical Report from Technical Committee ISO/TC 147 "*Water quality*" of the International Organization for Standardization (ISO) has been taken over as a European Prestandard by Technical Committee CEN/TC 230 "*Water analysis*", the secretariat of which is held by DIN.

According to the CEN/CENELEC Internal Regulations, the national standards organizations of the following countries are bound to implement this European Standard: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom.

Endorsement notice

The text of the International Technical Report ISO TR 13530:1997 has been approved by CEN as a European Prestandard without any modification.

TECHNICAL REPORT

ISO/TR 13530

First edition
1997-09-01

Water quality — Guide to analytical quality control for water analysis

Qualité de l'eau — Guide de contrôle qualité analytique pour l'analyse de l'eau

Reference number
ISO/TR 13530:1997(E)

Contents

Descriptors: Water, quality, water pollution, tests, water tests, chemicals analysis, quality control, rules (instructions).

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The main task of technical committees is to prepare International Standards, but in exceptional circumstances a technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/TR 13530, which is a Technical Report of type 2,was prepared by Technical Committee ISO/TC 147, *Water quality*, Subcommittee SC 7, *Precision and accuracy*.

This document is being issued in the Technical Report (type 2) series of publications (according to subclause **G.6.2.2** of part 1 of the ISO/IEC Directives, 1995) as a "prospective standard for provisional application" in the field of water quality because there is an urgent need for guidance on how standards in this field should be used to meet an identified need.

This document is not to be regarded as an "International Standard". It is proposed for provisional application so that information and experience of its use in practice may be gathered. Comments on the content of this document should be sent to the ISO Central Secretariat.

A review of this Technical Report (type 2) will be carried out not later than three years after its publication with the options of: extension for another three years; conversion into an International Standard; or withdrawal.

Annex A to Annex E of this Technical Report are for information only.

blank

1 Scope

This Technical Report (type 2) is a guide with the objective of providing detailed and comprehensive guidance on a coordinated programme of within-laboratory and between-laboratory quality control for ensuring the achievement of results of adequate and specified accuracy in the analysis of waters and associated materials.

This Technical Report and its annexes are applicable to the chemical and physicochemical analysis of natural waters (including sea water), waste water, raw water intended for the production of potable water, and potable water. It is not intended for application to the analysis of sludges and sediments (although many of its general principles are applicable to such analysis) and it does not address the biological or microbiological examination of water. Whilst sampling is an important aspect, this is only briefly considered.

Analytical quality control as described in this Technical Report is intended for application to water analysis carried out within a quality assurance programme. This Technical Report does not address the detailed requirements of quality assurance for water analysis.

The recommendations of this Technical Report are in agreement with the recommendations of established quality assurance documentation (for example ISO Guide 25 and EN 45001). A discussion of quality systems in water analysis is provided in clause **4** to set in context the recommendations on quality control.

This Technical Report is applicable to the use of all analytical methods within its field of application, although its detailed recommendations may require interpretation and adaptation to deal with certain types of determinand (for example non-specific determinands such as suspended solids or biochemical oxygen demand). In the event of any disparity between the recommendations of this Technical Report and the requirements of a standard method of analysis, the requirements of the method should prevail.

The basis of the Technical Report is to ensure the achievement of results of adequate accuracy by adherence to the sequential stages of analytical quality control shown in Figure 1.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this Technical Report. At the time of publication, the edition indicated was valid. All standards are subject to revision, and parties to agreements based on this Technical Report are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 5667-1:1980, *Water quality — Sampling — Part 1: Guidance on the design of sampling programmes.*

ISO 5667-2:1991, *Water quality — Sampling — Part 2: Guidance on sampling techniques.*

ISO 5667-3:1994, *Water quality — Sampling — Part 3: Guidance on the preservation and handling of samples.*

ISO 8466-1:1990, *Water quality — Calibration and evaluation of analytical methods and estimation of performance characteristics — Part 1: Statistical evaluation of the linear calibration function.*

ISO 8466-2:1993, *Water quality — Calibration and evaluation of analytical methods and estimation of performance characteristics — Part 2: Calibration strategy for non-linear second order calibration.*

ISO Guide 25:1990, *General requirements for the competence of calibration and testing laboratories.*

EN 45001:1989, *General criteria for the operation of testing laboratories.*

	Activity	Purpose
1.	Establish working group	To plan and coordinate subsequent activity.
2.	Define analytical objectives	To ensure clear specification of analytical requirements.
3.	Choose analytical methods/systems[a]	To select methods/systems capable of the required accuracy.
4.	Ensure unambiguous description of methods	To ensure that the chosen methods are followed properly.
5.	Estimate within-laboratory precision and spiking recovery	To ensure that each laboratory achieves adequate precision and to check certain sources of bias.
6.	Ensure accuracy of standard solutions. Preliminary check on interlaboratory bias	To eliminate this source of bias in each laboratory and to prepare full, more detailed bias checks.
7.	Set up quality control charts	To maintain a continuing check on analytical performance in each laboratory.
8.	Undertake tests of interlaboratory checks	To ensure that each laboratory achieves adequately small errors.
9.	Maintenance of accuracy using control charts and regular follow-up interlaboratory tests	To ensure long-term control of the accuracy and comparability of analytical results.

[a] The analytical method is the set of written instructions followed by the analyst. The analytical system includes all aspects of producing results, i.e. method, equipment, analyst, laboratory environment, etc.

Figure 1 — Sequence of activity for analytical quality control

3 The nature and sources of analytical errors

3.1 General

The following clauses provide a succinct discussion of the nature and origin of errors in analytical results for waters and effluents. Further information on many of the topics covered is given elsewhere in this Technical Report, and the subject is also discussed extensively in [18].

3.2 Nature of errors

The results of chemical analysis of waters and effluents (like those of all measurement processes) are subject to error, i.e. the measured concentrations differ from the true concentrations.

3.2.1 *Total error*

The total error, E, of an analytical result, R, is defined as the difference between that result and the true value, T; i.e.

$$E = R - T$$

As the total error decreases, the accuracy of the result is said to increase.

In general, the total error represents the sum of random error and systematic error.

3.2.2 *Random error*

Repeated analysis of identical portions of the same, homogeneous sample does not, in general, lead to a series of identical results[1)]. Rather, the results are scattered about some central value. The scatter is attributed to random error, so called because the sign and magnitude of the error of any particular result vary at random and cannot be predicted exactly. Precision is said to improve as the scatter becomes smaller — i.e. as random error decreases — and imprecision is therefore a synonym for random error.

Because random errors are always present in analytical results, statistical techniques are necessary if correct inferences regarding true values are to be made from the results.

Terms such as “repeatability” and “reproducibilty” have specialized meanings in the context of interlaboratory collaborative trials. In this Technical Report, random error is quantified in terms of the standard deviation, σ. Since exact measurement of the standard deviation generally requires an infinite number of repeated results, only estimates, s, of σ will usually be obtainable. The number of degrees of freedom (DF) of the estimate provides an indication of its worth; as the number of degrees of freedom increases, the random error of the estimate itself, s, decreases.

3.2.3 *Systematic error*

Systematic error (or bias) is present when there is a persistent tendency for results to be greater, or smaller, than the true value. The mean of n analytical results for identical portions of a stable, homogeneous sample approaches a definite, limiting value, μ, as n is increased indefinitely. When μ differs from the true value, T, results are said to be subject to systematic error or bias, β, where:

$$\beta = \mu - T$$

Because an indefinitely large number of determinations cannot be made on a single sample, the effect of random error prevents exact determination of μ, and hence also of β. Only an estimate, $\bar{x}$, of μ will generally be available, so that only an estimate, b, of β can be obtained.

As the systematic error or bias of results decreases, trueness is said to increase.

3.3 Sources of error

The distinction between random and systematic errors is important for two reasons: first, because they have different effects on the use to be made of analytical results, and second, because they usually have different origins.

3.3.1 *Causes of random error*

Random errors arise from uncontrolled variations in the conditions of the analytical system[2)] during different analyses. The nature of such variations include, for example, differences in the volume of sample or reagent taken on different occasions, fluctuations in temperature — either in time, or across the different sample positions in a heating bath, block or oven, fluctuations in instrumental conditions (for example in temperatures, fluid flowrates, voltages and wavelengths) and operator-induced variations in reading scales. Variations from batch to batch, in the extent to which the calibration function represents the true calibration for that batch, also give rise to between-batch random errors, whereas a consistent calibration error across many batches gives rise to systematic error — see below.

Whilst many of these factors causing random errors can be more closely controlled to achieve better precision, they can never be totally eliminated, so that all results are subject to some degree of random error.

[1)] This may not be true when the discrimination of the analytical system is coarse. However, the apparent perfect concordance of repeated results in such a situation is illusory, because samples differing in concentration will also give the same results.

[2)] The analytical system is the combination of all factors — analyst, equipment, method, reagents, etc. involved in producing analytical results from samples.

3.3.2 *Causes of systematic error*

There are five general sources of systematic error (if clear blunders by the analyst in carrying out the written method, and bias introduced by the sample collection itself are both excluded).

These are:

a) Instability of samples between sample collection and analysis

This is a potentially important source of error in many cases, and evidence should always be obtained — either from the literature or by direct test — to ensure that unacceptable bias is not introduced by this factor. Effective sample stabilization procedures are available for many determinands, but they should be compatible with the analytical system being employed, and with the particular sample type being analysed.

b) Inability to determine all relevant forms of the determinand

Many substances exist in water in a variety of physical and/or chemical forms (or "species"). For example, iron can exist in both dissolved and particulate forms, and within each of those physical categories a variety of chemical species may be present — for example free ions and complexes, including those of different oxidation states, in the dissolved phase. An inability of the analytical system to determine some of the forms of interest will give rise to a bias when those forms are present in samples.

Some determinands are overall properties of a sample, rather than a particular substance — for example biochemical oxygen demand (BOD). Such determinands are called "non-specific" and have to be carefully defined by specifying the use of a particular analytical method. The so-called "dissolved" fractions of, for example trace metals, are also non-specific in the sense that the type and pore-size of filter to be used in their determination should be clearly specified.

c) Interferences

Few analytical methods are completely specific for the determinand. Response to another substance (for example, response to iron by a spectrophotometric procedure for manganese based on formaldoxime) will give rise to biased results when that substance is present in samples, and it is important that the effects of all such interferents likely to be present in samples are known before a new method is applied routinely.

In some cases, the effect of another substance is to alter the chemical state of the determinand such that it is not measured by the method being used — for example, the presence of fluoride will cause aluminium complexes to form, which may not be measured by an ion-selective electrode. Such an effect can be regarded as an interference upon the determination of total dissolved aluminium, or as a failure to recover all forms of dissolved aluminium. Although it more strictly falls into the latter category, the effect — and others like it — may be most conveniently treated as an interference when data on performance characteristics are being obtained or reported (see clause **5** and Annex A).

d) Biased calibration

Most methods require the use of a calibration function (explicit or implicit) to convert the primary analytical response for a sample to the corresponding determinand concentration. If the samples and calibration standards are treated in exactly the same manner (and provided that the materials used to prepare the calibration standards are of adequate purity) no systematic error should arise from the calibration. (It has been noted in **3.3.1** that any variations in the correctness of the calibration from batch to batch will be manifested as between-batch random errors).

If, however, samples and calibration standards are treated differently, this can represent a potentially serious source of error. Thus, for example, a method prescribing some form of pre-concentration of the determinand from samples, but employing direct calibration with standards not taken through the pre-concentration step, will give rise to negative bias if the pre-concentration recovery is less than 100 %. In such cases, evidence should be obtained on the accuracy of the prescribed calibration, or the difference in treatment of samples and standards eliminated.

Impurity of the material used to prepare calibration standard is, of course, another potential cause of biased results.

e) Biased blank

The same considerations as in d) above apply to blanks. There is, however, another source of bias arising from blank correction. If the water used for the blank contains the determinand, results for samples will be biased low by an equivalent amount unless a correction for the determinand content of the blank water is applied. Ideally, however, a source of blank water should be obtained, such that the determinand content is negligible in comparison with the concentration in samples.

4 The quality system in water analysis

4.1 General

The quality system is the term used to describe the aspects which are intended to meet the clients' requirements. The control of analytical errors, usually termed analytical quality control (AQC) is an important component of the quality system. This clause summarizes the key issues which should be addressed in designing a quality system.

For more detail and an authoritative account of quality assurance, readers should consult the standards listed in clause **2**, together with documentation provided by the various national accreditation bodies.

4.2 Quality system

4.2.1 *Aims and form of quality system*

The laboratory should operate a quality system appropriate to the type, range, precision and volume of tests that it undertakes. The quality system should be such as to ensure that the requirements of this Technical Report are fully met on a continuing basis. All staff should be made fully aware of, and be required to comply with, the documented quality system.

The laboratory should possess a statement of the aims and general form of the laboratory's quality system, including the purpose of the quality manual and associated documentation.

4.2.2 *Quality manual*

The quality system should be formalized in a quality manual which should be maintained and kept up-to-date.

The person responsible for authorization and compilation of the quality manual should be identified. A distribution list of the quality manual and identification of holders of controlled copies of the quality manual should be included.

The quality manual should contain, for example the following items or equivalent:

1 Scope.
2 References.
3 Definitions.
4 Organization and management.
5 Quality system, audit and review.
6 Personnel.
7 Accommodation and environment.
8 Equipment and reference material.
9 Measurement, traceability and calibration.
10 Test methods.
11 Handling of calibration and test items.
12 Records.
13 Certificates and reports.
14 Subcontracting of calibration or testing.
15 Outside support services and supplies.
16 Complaints.

4.2.3 *Quality management*

The quality system should include a statement of the responsibilities and authority of the technical manager and quality manager, and any appointed deputies.

The quality system should include a statement of the general arrangements for implementing each of the quality manager's and deputy's responsibilities and the specific procedures for implementing these responsibilities, or identification of laboratory documents containing such procedures.

4.2.4 *Documentation*

The quality system should include a statement of the quality manager's responsibility in relation to control and maintenance of documentation, including the quality manual, and of the specific procedures for control, distribution, amendment, updating, retrieval, review and approval of all documentation relating to the testing work of the laboratory.

It should be made clear that laboratory staff should have ready access to all documentation, including the quality manual test procedures, and all relevant standard specifications.

There should be an instruction forbidding alteration of laboratory documentation, except under conditions specified by management. Instructions should also require adherence by all staff to the laboratory's laid-down policies and procedures, except under clearly specified conditions.

The instruction should indicate clearly the circumstances under which departures from documented policies and procedures or standard specifications are permitted. It should also indicate that departures should be endorsed by management and their justification included in relevant records.

The quality system should include a statement of the specific procedures for dealing with situations where staff have not followed documented policies, procedures and standard specifications as required.

4.3 Quality policy

4.3.1 *Management statement*

The quality system should include a statement by senior management of the policy of the laboratory as regards quality in all aspects of its work.

4.3.2 *Quality systems*

The quality system should include a statement of the intentions of the laboratory management in relation to quality of service.

The quality system should include a statement that it is the responsibility of all staff to familiarize themselves with the content of the quality manual and to comply with the policies and procedures laid down in the quality manual and associated documentation.

These statements should be made on the authority of a senior executive, who has direct management responsibility for the laboratory and who is at the highest level of management on which decisions are taken on laboratory policy and resources. They should be authenticated by the signature, and legible name and position, of the person concerned.

The policy statements should indicate:

— title of the person responsible for implementing the quality policy in the laboratory,

— title of the person having overall responsibility for control of quality and who advises on and monitors all aspects of quality in the laboratory.

4.4 Organization and management

4.4.1 *Organization*

The quality system should include a statement to the identity and legal status of the laboratory (including ownership and corporate position in relation to any parent organization or grouping of companies) and a statement of the technical rôle of the laboratory (for example independent, commercial, calibration/testing, or product quality control in support of a particular manufacturer).

A brief historical background should be included which is relevant to the standing of the laboratory, together with a summary of the scope of operation and range of testing performed by the laboratory and an inventory of tests performed by the laboratory.

4.4.2 *Organization charts*

Organization charts should be used showing:

— technical manager, quality manager, and any deputies;

— general lines of responsibility within the laboratory (including the relationship between management, technical operations, quality control and support services);

— the lines of responsibility within individual sections of the laboratory;

— the relationship between the laboratory and any parent or sister organizations.

The appropriate chart should show that, for matters related to quality, the quality manager has direct access to the highest level of management at which decisions are taken on laboratory policy or resources, and to the technical manager.

4.4.3 *Management*

Details of job descriptions, qualifications, training and experience should be provided for:

— technical manager;

— quality manager;

— other key laboratory managerial and technical posts.

Job descriptions should include:

— title of job and brief summary of function;

— person or functions to whom jobholder reports;

— person or functions that report to jobholder;

— key tasks that jobholder performs in the laboratory;

— limits of authority and responsibility.

Technical manager — the quality system should include a statement that the post-holder has overall responsibility for the technical operation of the laboratory and for ensuring that the quality system requirements are met.

Quality manager — the quality system should include a statement that the post-holder has responsibility for ensuring that the requirements for the quality system are met on a day-to-day basis and that the post-holder has direct access to the highest level of management at which decisions are taken on laboratory policy or resources, and to the technical manager.

The quality system should include a statement of:

— any special arrangements for management and other functions, in the event of the absence of key staff;

— method whereby the laboratory makes all staff aware of the extent and limitations of their responsibility;

— laboratory policy with respect to influences or inducements that might adversely affect the judgement of staff or the result of their work;

— actions to be taken by staff on encountering such influences or inducements;

— laboratory policy and general arrangements for ensuring the protection of proprietary rights and information, and for other aspects of site security.

4.4.4 *Staff qualifications and training*

The quality system should include a statement of the laboratory's general criteria for engagement of staff and their assignment to respective duties and list the relevant academic and/or professional qualifications, training and experience of key managerial, technical and other staff.

The quality system should include a statement of the laboratory's procedures for defining in writing which members of staff are authorized to use equipment or perform specific calibrations or tests.

The quality system should include a statement of the laboratory's policy for the use of staff undergoing training including their supervision.

The quality system should include a statement as to the maintenance, accessibility and location of all records relating to staff competence, qualifications, training and experience.

4.5 Testing environment

4.5.1 *Laboratory testing environment*

Samples, reagents and standards should be stored so as to ensure their integrity. The laboratory should guard against deterioration, contamination and loss of identity.

Special care may need to be taken in some laboratories, for example those involved in trace analysis. In such cases, it would be expected to see physical separation of high level and low level work. Where special areas are set aside for trace analysis, access to these areas should be restricted, and the type of work undertaken carefully controlled.

Staff should be aware of:

— the intended use of a particular area;

— the restrictions imposed on working within such areas;

— the reasons for imposing such restrictions.

Specific procedures for environmental conditions, for example the need for air-conditioning, required by particular categories of tests, and the laboratory's general arrangements for providing such conditions and for dealing with disturbances of environmental laboratory conditions should be established.

4.5.2 *Laboratory housekeeping*

The quality system should include a general statement containing an instruction to staff to maintain good housekeeping throughout the laboratory and where appropriate, to adopt special procedures.

4.5.3 *Staff accommodation and conditions*

The quality system should include a statement of laboratory's intention to provide accommodation and conditions for all staff, conducive to the proper performance of their respective duties.

4.6 Equipment and reference material

4.6.1 *General*

The quality system should include an operating programme for the maintenance and calibration of equipment used in the laboratory.

The quality system should include a statement of the use of reference material (a stable and homogeneous material or substance, one or more properties of which are sufficiently well established to be used for the calibration of apparatus, the assessment of a measurement method or for assigning values to materials).

4.6.2 *General service equipment*

General service equipment should only be maintained by cleaning and safety checks as necessary. Calibrations or performance checks will be necessary where the setting can significantly affect the test or analytical result (for example, the temperature of a muffle furnace or constant temperature bath.)

4.6.3 *Volumetric equipment*

The correct use of volumetric equipment is critical to analytical measurements and it should be maintained, calibrated and used in a manner consistent with the accuracy required of data. For the highest accuracy, measurements can often be made by mass rather than by volume. The type used (glass, polytetrafluoroethylene, etc.) cleaning, storage and segregation of volumetric equipment is critical for the avoidance of contamination, particularly for trace analyses when leaching and adsorption can be significant.

4.6.4 *Balances and other measuring instruments*

Periodic performance checks should be carried out at specified intervals on balances. Other instruments may be checked (for example for response, stability and linearity of sources, sensors and detectors, the separating efficiency of chromatographic systems, the resolution, alignment and wavelength accuracy of spectrometers, etc.).

These checks are particularly important if errors are not easily detected as part of the routine quality control system.

4.6.5 *Computers and data processors*

The computer systems should periodically be checked and maintained. The computer software should be fully documented and validated before use.

4.6.6 *Defective equipment*

Defective equipment should immediately be withdrawn from analytical work and not be returned until re-calibrated, re-validated and re-authorized for use.

4.7 Receipt and handling of test samples

4.7.1 *General*

General arrangements are required for:

— receipt, recording and handling of test items;

— transfer of test items to and from the laboratory;

— transfer of test items between different sections of the laboratory;

— sampling, transport, storage, bonding and protection of items, as appropriate.

4.7.2 *Identification of test samples*

There should be a system for identification of test items and any accessories.

4.7.3 *Condition on receipt of test samples*

The quality system should include a statement of the procedure for recording the condition of an item on receipt and of the steps to be taken in conjunction with the client if the suitability of the item for testing is in doubt.

4.7.4 *Disposal of test samples*

The quality system should include a statement of the procedures to be adopted after completion of the test for the retention and/or safe disposal of test items.

4.8 Test reports

4.8.1 *Control of test reports*

The quality system should include a statement of laboratory policy and procedure for the content, format, issue and transmission of certificates and test reports. The statement should cover the procedures for the use of electronic data transfer as well as hard copy; instruction on minimum information content of certificates and test reports including condensed or coded examples; instruction on layout of test reports; and sample test report forms used by the laboratory.

4.8.2 *Authorization of test reports*

Instruction should be provided on checking and signature of test reports; and identification of approved signatories (by list of named individuals or categories of staff) authorized to take technical responsibility and legal responsibility for test reports on the laboratory's behalf.

4.8.3 *Validity of test reports*

Instruction should be provided on action to be taken, in the event of doubts arising as to the accuracy or validity of results in any test reports issued.

4.8.4 *Transmission of test results*

Instruction should be provided in the procedures to be adopted when clients request transmission of results by means other than as written test reports.

4.9 Records

4.9.1 *Records system*

The quality system should include a statement that the aim of the laboratory is to maintain records systems such that all information of practical relevance to tests performed is available for a defined time period.

The laboratory should ensure that this statement defines records systems that will enable them to identify sources of any error and where appropriate, repeat the test under conditions close to the original.

Arrangements for coordinating the system through the laboratory and specific procedures for maintaining records should be included.

4.9.2 *Information recorded*

The quality system should include a list of types of information held in the records and identification and location of laboratory documents, or computerized records, holding each type of information.

The quality system should also include a statement on the use of worksheets or workbooks to record various types of observation, calculation or other relevant information. Any changes to data should be recorded, authorized and the original results retained.

This statement should indicate that provision should be made in all worksheets and workbooks to record all the information required by the test method in order that the test can be repeated under similar conditions if necessary.

The quality system should include a statement of the procedure to be followed for recording, checking, correcting, signing and countersigning observations and calculations.

4.9.3 *Protection of records*

The quality system should include instruction on protection of records, including those held on computer.

4.9.4 *Procedure for records*

The quality system should include a statement of the procedure for retention of records including observations and calculations, calibration certificates, test reports and test certificates. The period of retention should be included.

4.10 Purchase of outside supplies and services

The quality system should include a statement of laboratory policy in relation to using outside supplies and services of adequate quality with particular reference to goods certified by an accredited certification system.

The quality system should also include a statement of general arrangements for ensuring that outside supplies and services are of adequate quality for the laboratory's purposes and that the necessary checks, calibrations or other actions are made on purchased goods before use.

4.11 Subcontracting of tests

4.11.1 *Subcontracting policy*

The quality system should include a statement of laboratory policy and general arrangements on the subcontracting of tests.

This should emphasise that the laboratory's policy is to carry out all tests and only to subcontract in special circumstances. In the event that an unaccredited laboratory has to be used, the quality manual should indicate that the subcontractor will be assessed against the requirements of the quality system before use by the laboratory.

4.11.2 *Subcontracting register*

Identification of laboratory documents containing a register of all subcontractors used, records of assessments performed where appropriate, and a record of all subcontracted work should be maintained.

4.12 Diagnostics and corrective actions

The quality system should include a statement of laboratory policy on handling of complaints and anomalies. The action to be followed, and the records to be made, by staff who are authorized and in the first instance receive complaints or identify anomalies should be included.

Laboratory documents containing comprehensive records of complaints and anomalies and their handling should be identified.

The quality system should include a statement of the procedures for analysing the records and proposing improvements to the quality system by the quality manager.

4.13 Quality audit and quality system review

4.13.1 *Purpose*

The quality system should include a statement of the laboratory's aims in relation to quality audit and review.

Audits (periodic checks carried out by or on behalf of management) should be ensured so that the laboratory's policies and procedures as set out in the quality manual are being followed. All activities and responsibilities covered by the quality system, including the performance of tests, should be audited.

Audits should be performed as frequently as is necessary but the programme should ensure that each aspect of the quality system is examined.

Review (examination of the quality system by or on behalf of management) should be ensured so that the quality system complies with the quality standards on a continuing basis.

Reviews should be conducted as frequently as is necessary.

4.13.2 ***Responsibility***

The quality system should include a statement of the responsibilities of the quality manager and any appointed audit officers or deputies in relation to the planning, organization and conduct of quality audits.

The laboratory should indicate who has been appointed as audit officer and that it is the policy to choose an officer who is independent of the specific activities being audited. Laboratory staff are not permitted to audit their own activities.

The quality system should include a statement indicating that it is the responsibility of the quality manager to verify that all corrective actions required by audits have been completed and that reviews are recorded in the quality system.

4.13.3 ***Implementation***

The quality system should include a statement of general arrangements for quality audit and quality system review.

Included should also be a summary of any internal data-monitoring carried out by the laboratory, as part of its quality system and a summary of the scope of any external organization performing quality audit or assessment of the laboratory.

4.13.4 ***Planning and documentation***

Laboratory documentation containing records of the planning, implementation and outcome of the activities referred to in **4.13.3** should be identified.

The quality system should include samples of forms used by the laboratory's programme for internal quality audit, for quality system review, and where appropriate, for external quality audits and assessments.

4.13.5 ***Quality control systems***

The quality system should include a statement of general arrangements for implementing checks on the quality of results provided to clients by the use of internal quality control schemes, proficiency testing, interlaboratory comparisons, reference materials, replicate testing and retesting.

5 Performance characteristics of analytical systems

5.1 General

After the analytical requirements (see clause **6**) have been established, the establishment of an appropriate analytical system (see clause **7**) is made easier if adequate performance data are available for the analytical system of interest.

The description of each of the characteristics listed below should be provided whenever possible. The first three are intended to give a qualitative impression of the procedure and the applications for which it is known to be suitable and can be of value in deciding whether it is worth consideration. The remaining items are all quantitative performance characteristics.

5.2 Substance determined

A clear definition should be given of the forms of the substance that are determined by the procedure and also, when necessary to avoid ambiguity, those forms that are not capable of determination.

5.3 Types of sample

The types of sample for which the system has been evaluated and is known to be suitable should be stated. If necessary they may be distinguished from types of samples for which the procedure may be adequate, but for which the analyst should carry out appropriate tests before deciding on its suitability.

5.4 Basis of system

A brief and clear statement of the system should be given that summarizes the principles of the procedure and describes any instrumental techniques utilized.

5.5 Range of application

The range of application corresponds to the lowest and highest concentrations for which tests of precision and bias have been carried out using the system without modification. Where an extension can be used to enable the examination of samples containing concentrations greater than the upper limit, such as by analysis after dilution, then it should be regarded as a different procedure but whose performance characteristics can be inferred from the values quoted for the original.

5.6 Calibration function

The nature of the calibration function (the relationship between measured response and the quantity of interest) obtained by the application of the system should be indicated and, if possible, an equation should be quoted along with any concentration limits for its validity. An advantage of this is that information is also given on the sensitivity of the procedure. This topic is discussed further in Annex C.

5.7 Total standard deviation of analytical results

The total standard deviation of results is an aggregate of short-term (within-batch) and longer-term (between-batch) sources of random error. It is a measure of the uncertainty which affects individual data points produced by a given analytical system.

In general, the total standard deviation is of greatest interest in routine analysis and therefore is normally quoted in the quantification of random error and should be accompanied by the relevant number of degrees of freedom. Between-batch and within-batch standard deviations may also be of interest and should be quoted if available. When estimates for a given concentration are available from more than one source, then the range of these estimates should be described to indicate the expectation that the quoted precision will be achieved by another analyst. Since the total standard deviation typically increases with the concentration of the determinand, values at zero concentration and the upper limit of the range of application should be quoted. (Further information is found in Annex C and clause **8**).

5.8 Limit of detection

There has been much diversity in the way in which the limit of detection of an analytical system is defined. Most approaches are based on multiplication of the within-batch standard deviation of results by a factor (usually between 2 and 10, depending on the degree of confidence required for detection).

The limit of detection may thus be defined as that concentration of the determinand for which there is 95 % probability of detection when a single analytical result is obtained, detection being defined as obtaining a result which is significantly greater (p = 0,05) than zero.

The magnitude of the limit of detection can be determined from the within-batch standard deviation, s_w, of results for a solution, such as a blank, containing a very small (preferably zero) concentration of the determinand. s_w is expressed in concentration terms so that the effects of calibration procedures on the variability of results for determinations on low concentration samples are accounted for. The limit of detection is given by $2 \times \sqrt{2} \times t_{0.05} \times s_w$ where $t_{0.05}$ is the tabulated value of Student's t (single-sided) at the 95 % probability level and for the relevant number of degrees of freedom (which should also be stated).

When a number of estimates of limit of detection is available from more than one source, the range may be of interest and could be quoted (together with the number of degrees of freedom in each case). Other indicators of measurement capability have been proposed, for example "limit of quantification" at 10 s_w.

5.9 Sensitivity

The sensitivity of the system is defined as the change in analytical response with changing concentration of the determinand (at stated concentrations),

$$\text{sensitivy} = \frac{\delta(\text{response})}{\delta(\text{concentration})}$$

5.10 Bias

The nature of any known form of bias should be summarized. Comparison of results obtained using the system under consideration with those using reference procedures, and also the results obtained by the analysis of certified materials using the system under consideration are all relevant — but are often not available.

The concentration of many determinands may change between sampling and analysis, and large systematic errors may result. Sample stability is more fully discussed in **3.3.2** and clause **7**.

Evidence concerning the magnitude of these systematic errors and the efficiency of measures to eliminate them is required. Possible errors due to the inability of the system to measure all forms of the substance as defined above, and any bias attributable to the methods of calibration and blank correction, should be known and reported.

5.11 Interferences and matrix effects

An important source of systematic error in results is the presence of constituents of a sample other than the determinand that cause an enhancement or a suppression of the analytical response. The evaluation of interference effects on analytical systems is described in **3.3.2** and Annex A, and the results of such evaluation should provide estimates of error at or near the lower and upper concentration limits of the system. These estimates should be available for each substance of interest (interferents) at a concentration slightly higher than the greatest value expected in samples.

6 Specifying analytical requirements for water analysis

6.1 General

Careful specification of analytical requirements is a vital feature in the design of programmes of sampling and analysis to assess water quality. A failure to pay proper attention to this topic can jeopardize the validity of such a programme, since analytical results of inadequate accuracy can result in false conclusions being drawn.

Particular attention should be paid to the following: Unambiguous definition of the determinand, description of the sample, the concentration range of interest, the accuracy (trueness and precision) required of results, and the expression of results.

6.2 Definition of the determinand

It is an obvious point, but worth emphasizing, that the determinand of interest should be defined unambiguously. If this is not done, the analytical method employed may be inappropriate.

Many substances exist in water in a variety of forms or "species", and most analytical systems provide a differential response to the various forms. For example, when a separation of "dissolved" and "particulate" material is required, special care is necessary to define precisely the nature and pore-size of the filter to be used.

Particular care should also be exercised with non-specific determinands such as BOD or total organic carbon (TOC), which are defined solely by the method used to determine them. Comparability of results for such determinands from different laboratories demands that all use the same experimental conditions of measurement.

6.3 Description of the sample

A precise description of the type and nature of sample is important before the analytical system can be chosen. The precautions to be taken when a sample is analysed will depend to a high degree on the sample.

6.4 Specification of concentration range of interest

The concentration range of interest can have a marked effect on the choice of analytical technique; of primary concern is the smallest concentration of interest.

As an example, consider the determination of a substance subject to a water quality limit of 100 $\mu g\ l^{-1}$ the concentration of primary interest. The lowest concentration of interest will clearly be less than 100 $\mu g\ l^{-1}$, but to set it at 0,1 $\mu g\ l^{-1}$ would normally be unnecessary and cause needless analytical effort to be expended. It is normal practice that the lowest concentration of interest be set at one-tenth of the relevant limit — i.e. at 10 $\mu g\ l^{-1}$ in the above example — although other considerations could be applied, if appropriate.

It is also prudent to define the upper concentration limit of interest, but this is normally of limited importance in selecting an appropriate analytical system, because recourse can always be made to dilution if the concentration in a sample lies above, the upper limit of the system's range. The effect of dilution on any matrix effect should also be taken into account.

6.5 Specification of maximum tolerable random and systematic errors of analytical results

The following paragraphs illustrate a logical general approach to be adopted for specifying the maximum tolerable random and systematic errors of individual analytical results:

"The systematic error of individual analytical results should not exceed c concentration units (for example $\mu g\ l^{-1}$) or p % of the result, whichever is greater."

"The random error of individual analytical results should not exceed c concentration units (for example μg l^{-1}) or p % of the result, whichever is greater."

"The total error of individual analytical results should therefore not exceed $2c$ concentration units (for example μg l^{-1}) or $2p$ % of the result, whichever is greater."

The values of c and p should be chosen for each application, but a value of $p = 10$ may often be suitable. It may at first sight appear large, but experience shows how difficult it is to achieve, for determinands at trace concentrations in particular. Near the detection limit, p will increase to a value of 100 (depending on the definition of the detection limit).

The random error (95 % confidence limits) is equal to twice the total standard deviation of analytical results. (The factor is, strictly, 1,96 rather than 2, but the difference is so small as to be unimportant in setting analytical requirements). Thus if $p = 10$, it follows that the maximum tolerable total standard deviation, s, is $0{,}5p = 5$ %.

Very often the lowest concentration of interest, CL, is set to one-tenth of the concentration of primary interest (for example water quality limits). The value of c may usefully be set at one-half of CL. This has the effect of making the tolerable total error at the lowest concentration equal to that concentration. The maximum tolerable total standard deviation, s, is then $0{,}5c = 0{,}25CL$ (or 5 % of the result, whichever is the greater).

Table 1 shows the effect of these recommendations, for a range of values of the concentration of primary interest (for example water quality limits).

Table 1 — Examples of the calculation of maximum tolerable errors of individual analytical results

Maximum tolerable errors			
Water quality limit	**Total standard deviation**	**Systematic error**	**Total error**
100	2,5 or 5 % of result	5,0 or 10 % of result	10,0 or 20 % of result
50	1,25 or 5 % of result	2,5 or 10 % of result	5,0 or 20 % of result
10	0,25 or 5 % of result	0,5 or 10 % of result	1,0 or 20 % of result

NOTE 1 All concentrations across the line are in the same units.
NOTE 2 Of the two values tabulated for each error, the larger applies for any given concentration.

The above recommendations regarding the setting of the lowest concentration of interest and the maximum tolerable random and systematic errors are regarded as suitable for general water quality monitoring to check compliance with most types of water quality standards.

The concept of "smallest concentration of interest" may not be applicable to some determinands — for example pH — for which the expressions for the tolerable errors will usually need to specify only fixed proportional or absolute errors.

A clear statement of the random and systematic errors of the analytical system should be available to accompany each analytical result, for example expressed as standard deviation and recovery. The number of significant figures should normally be the same for the results as for the analytical responses from which they were derived. For example, a useful rule is to make the last reported digit the first digit of the standard deviation. As always, the uses of the data should be borne in mind. The units in which results are expressed should be stated unambiguously.

7 Choosing analytical systems

7.1 General considerations

Sampling of waters and effluents is carried out in order to provide information on their qualities. This information may be used for different reasons, for example

— legal control of discharges;

— environmental monitoring;

— process-control of treatment plant performance;

— evaluation of taxes and charges based on actual emissions.

The user's needs are of primary importance. The user of the information has the important responsibility to define precisely the objectives of the measurement programme and to help to choose the measurement techniques to be used. The following topics shall be defined in the measuring programme:

a) definition of the quantitative information required;

b) definition of the determinands;

c) location, time and frequency of sampling;

d) requirements for analytical results;

e) use of data and data handling routines, including statistical calculations;

f) introduction of a quality assurance programme.

In this clause it is emphasized that all analytical work should be based on a sound and precisely defined measurement programme, providing the analyst with representative and stable samples. The inclusion of a quality assurance system means that procedures are undertaken to produce data of stated quality. This is partially attained by analytical quality control activities which keep random and systematic errors within prescribed limits.

The following approach is purely practical, intended for routine use in laboratories.

7.2 Important factors in selecting analytical systems

Careful specification of analytical requirements is necessary to produce data of stated quality. Particular attention should be paid to the following items, important in the subsequent selection of an appropriate analytical system.

7.2.1 *Definition of determinand*

This is discussed in **6.2**. At this point, it is worth emphasizing that the analyst's selection of an analytical method should meet the user's definition of the determinand. Non-specific determinands need the use of rigorously standardized analytical methods in order to obtain reliable and comparable results.

7.2.2 *The concentration range*

This issue is discussed in **6.4**.

7.2.3 *Calibration and sensitivity*

In order to convert analytical responses obtained for samples to concentrations of the determinand, a calibration procedure is needed. The analytical procedure used for calibration should be identical to that used for real samples, and the calibration procedure should prescribe exactly the standard solution, number of concentrations, number of replicates, etc. The calibration curve indicates the range of concentrations over which the procedure can be used (see ISO 8466-1 and ISO 8466-2). The calibration normally includes the use of an analytical blank; any blank correction implied in the technique should be stated.

The quality of water used in preparation of standard and blank solutions should be examined carefully. The standard addition technique is used to overcome sample effects on the calibration curve. Methods for the determination of non-specific determinands are calibrated by the use of arbitrarily chosen standard solutions, prescribed and exactly defined in the standardized method. Detailed instructions on calibration procedures are given in [12], [18], ISO 8466-1 and ISO 8466-2.

Sensitivity may be defined as the change in analytical response for a given change in concentration. It does not indicate the ability to detect small concentrations.

7.2.4 *Types of samples and possible interferences*

Most analytical techniques produce accurate results with standard solutions at the optimal concentration. Relevant information should include the types of samples (fresh water, sea water, waste water, etc.) for which the method is suitable. Samples high in particles and suspended solids often cause problems, especially in automated analytical systems using extremely small sample volumes. Coloured or turbid samples offer problems in photometric determinations of COD and TOC. The potentially important rôle of sample types in water analysis should be recognized. The analyst needs information as complete as possible on sample types, concentration levels and possible interferences.

A measurement programme, for example a river survey, may often include a high number of very different types of samples. For this reason, routine analytical laboratories often prefer robust, multipurpose analytical techniques applicable to a broad range of samples.

The relation between the concentration of the determinand and the interfering substances needs to be considered. For example, it is extremely difficult to determine low COD values in the presence of chlorides or to determine small concentrations of organically bound nitrogen in samples with high levels of organic substances.

7.2.5 ***Accuracy (trueness and precision) required of results***

The general term accuracy is used to refer to trueness and precision combined. Accuracy is a measure of the total displacement of a result from the true value, due to both random and systematic errors. See **6.5**.

Results from interlaboratory tests may be useful in choosing analytical systems. Consistently accurate results in such tests often indicate an analytical approach which is robust. It is often important for a laboratory to keep a high level of quality in day-to-day routine operation. A strict adherence to a quality control programme is necessary, using reference substances where possible to check trueness and control charts in order to keep precision under control.

7.2.6 ***Practical considerations***

The agreed measurement programme, including the quality assurance programme, should be able to produce data that are documented thoroughly, scientifically sound and of known quality. As noted above, in selection of analytical systems the emphasis should be put on accuracy.

When discussing the requirements with the user and selecting suitable analytical systems to fit the measuring programme, the following practical points should be considered:

— the frequency of sampling and the total number of samples on each occasion;

— the maximum period between sampling and analysis, in relation to sample stability;

— the maximum period between sampling and the user's need for the results;

— the volume of sample available;

— automatic or manual techniques;

— equivalent analytical methods;

— robustness and description of the proposed method;

— applicability of the proposed method in the laboratory concerned with respect to cost, speed, etc.

Regarding these practical considerations, factors such as convenience, speed and cost may have a great influence on the final selection of analytical systems. When analysis is required infrequently, it may be necessary to adopt a different approach from that used for regular, frequent determinations. It is still essential that the most appropriate action is taken to ensure control of the measurement process and to provide an estimate of analytical accuracy. See clause **12**.

7.2.6.1 *Sampling and sample stability*

General guidelines on these matters are given in ISO 5667-1, ISO 5667-2 and ISO 5667-3.

7.2.6.2 *Automatic or manual techniques*

The analyst has also to decide which techniques are most suitable: automatic or manual analysis. Generally, both techniques have about the same trueness but the automatic variant normally shows a better precision. Automated systems or other advanced instruments often use very small sample volumes and need special techniques in order to get representative subsamples. Advanced instrumentation, for example atomic absorption techniques, may be sensitive to sample matrix effects.

7.2.6.3 *Alternative analytical methods — Methods of equivalent performance*

In many cases, the use of a particular analytical method is prescribed, normally referring to a manual technique in a national standard method. This may lead to practical difficulties; advanced laboratories prefer the use of automated systems and a small laboratory at a treatment plant prefers a simplified technique for daily control of discharges. The concept of "equivalent analytical methods" means that laboratories are allowed to use different methods provided that it is shown that adequate performance is obtained. The precision and trueness for both methods on identical samples are determined and compared, these tests should be performed on relevant concentrations and types of samples.

7.2.6.4 *Robustness and description of method*

A "robust" or "rugged" analytical procedure means a procedure which is so designed that the accuracy of analytical results is not appreciably affected by small deviations from the experimental, design prescribed by the analytical method. The use of robust procedures is of great help to achieve reliable results in routine laboratories. The most robust procedure is the preferable choice, if the procedure meets the user's requirements. There is no simple numerical value indicating the robustness, but results from intercalibrations may help to illustrate the robustness of different procedures. Special responsibility falls on those who improve or standardize methods to produce robust techniques.

The need for complete and clear specification of analytical procedures should be stressed. The method should specify all details regarding analysis, equipment, calibration, calculation of results, etc., and also include any details on sampling, sample handling and preservation, any digestion step or other specific pretreatment of samples. Any optional operations should be specifically noted.

In addition to the description of the procedure any published or "standardized" method should preferably contain (or give references to) detailed and complete information on all factors mentioned above, such as determinand, concentration range, limit of detection, trueness and precision, calibration, interferences, robustness, etc.

7.2.6.5 *Applicability of methods*

Most routine laboratories have a set of "standardized" methods in day-to-day operation. As noted above, these methods may be characterized as robust, multipurpose analytical techniques, applicable to a broad range of sample types. In using these methods the laboratory should have experience, tested equipment and well-trained staff. The system incorporating these methods have well known performance characteristics, for different concentrations, sample types, interferences, etc. There is often information from intercalibrations and control charts. Such methods are to be preferred, provided that they meet requirements. It is also impractical for a routine laboratory to have too many methods for the determination of identical determinands.

If the routine methods do not meet requirements, the introduction of improved procedures and new instrumentation should be discussed. But it should be emphasized that introduction of new techniques, new instrumentation and training of staff are highly time consuming activities. Any discussions on the application of improved procedures should be initiated long before the start of the corresponding routine measurement programme. In order to reach a stated level of analytical quality, the new procedures should be in routine operation for some time before start of the real measurement programme.

7.3 Stepwise procedure to select analytical techniques to be used in a measurement programme

In establishing a measurement programme, including decisions on sampling, analysis and treatment of data, all aspects and parts of work are inter-connected. All important factors in selecting analytical techniques can be discussed in a stepwise procedure. A check-list composed of sequential stages may be useful. However, it is important to note that some decisions may require a change in this process at a later stage. The final programme often is a compromise between what is desirable and what is practicable. As noted above, all analytical work should be based on a defined measurement programme, providing the analyst with stable and representative samples. It should be emphasized that all analytical quality control activities should be performed in connection with an overall quality assurance programme.

Sequential stages in the decision process are as follows:

a) Define use to be made of data, determinands of interest.

b) Establish type, nature and number of samples.

c) Set accuracy requirements.

d) Choose analytical methods, considering determinands, accuracy, concentrations, types of samples and interferences.

e) Choose sampling techniques and sample pretreatment procedures.

f) Define AQC activities such as use of reference substances, control charts and participation in interlaboratory tests.

g) Define data handling and data reporting.

h) Define the analytical systems, considering all points mentioned in clause **7**.

i) Define and introduce a quality assurance system, including laboratory and data audits, in order to maintain a stated level of quality throughout the programme period.

j) Use audit results for corrective actions and make any necessary changes in the analytical systems, return to the previous stage.

8 Initial tests to establish performance of analytical system

8.1 General

Once a method has been chosen for a particular application, it is necessary to test the performance of its routine operation. The emphasis should be placed on an examination of the performance of the whole analytical system, of which the method is only a part. All the components of the analytical system — instrumentation, analysts, laboratory facilities, etc. — should be critically examined before routine analysis is started.

This clause describes the approach recommended for the experimental estimation, and when necessary reduction, of errors; this stage should be completed before any samples are analysed and may be called "preliminary error estimation". Subsequent quality control activity may be termed "routine quality control" and is dealt with separately.

8.2 Preliminary estimation of errors

8.2.1 *Systematic error*

The estimation of likely bias should already have been made in the initial evaluation of the technique. It will usually be impossible to check many of the most important sources of bias when a method is used routinely for the first time. A check on some sources of bias, by means of a spiking recovery, is included at this stage.

8.2.2 *Random error*

The estimation (and, when necessary, control) of random error is an essential precursor to routine analysis. Preliminary tests provide the necessary evidence that the precision of routine data is adequate, and form the basis for routine quality control.

8.3 Precision tests — General experimental design

Precision is likely to deteriorate as the experimental conditions pertaining to successive replicate determinations become more and more different. Thus, the precision obtained in one batch of analyses is often better than that of results spread over a longer period of time. Estimates of precision from one batch of analyses may, therefore, give an overoptimistic idea of the precision of results produced during routine analysis.

For this reason, precision should be estimated from analyses taken from separate batches, spread over a suitable period. The duration of this period is a matter of choice and depends on which sources of random variation are to be assessed. Testing to give at least 10 batches of analysis is recommended if a reliable estimate is to be obtained.

The approach described in this clause allows the total random error to be separated into random error arising from variations within and between batches of analysis. This information is of value in indicating the dominant sources of random error. Estimates of within-batch standard deviation are pooled from all batches and so provide an indication of what is achievable on a regular basis.

The basic approach is to make n determinations on a representative group of samples in each of m batches of analysis. In deciding on suitable values for n and m, care is required for two reasons.

— Too few analyses will not provide a worthwhile estimate of standard deviation. The uncertainty on an estimate of standard deviation depends on the number of associated degrees of freedom. Designs of test which are likely to provide estimates of standard deviation with fewer than 10 degrees of freedom may prove uninformative.

— It is desirable to design the test so that a satisfactory estimate of the dominant source of error is obtained. For example, if between-batch error is likely to be dominant, a design where $n = 10$ and $m = 2$ will give a relatively precise estimate of the less important source of error, but will estimate the dominant source of error very imprecisely. A more appropriate design would be for n to be made small and m large.

The experimental design recommended for general use is to make $n = 2$ and $m = 8$ to 10. Such a design provides estimates of within- and between-batch standard deviations with approximately equal numbers of degrees of freedom. This design should be modified as indicated by knowledge of the analytical technique. In particular, when within-batch errors are assumed to be dominant, values such as $n = 4$ and $m = 5$ could be chosen. This has the merit of reducing the number of batches of analysis which need to be conducted, whilst (given the assumption is correct) providing a reasonably good estimate of total standard deviation. The product $m\,n$ should not be less than 10 and should be preferably 20 or greater. Analysis of 11 batches of duplicate samples will guarantee that the estimates of total standard deviation will have at least 10 degrees of freedom.

8.3.1 ***Samples to be analysed***

8.3.1.1 *Design of test*

It is clearly essential that the solutions used for tests of precision show no appreciable changes in concentration during the period in which portions of them are taken for analysis. The solutions should also be sufficiently homogeneous that the concentration of the determinand is essentially the same in each portion of a solution. Water samples may sometimes be inadequately stable to allow tests over several days (adequate sample stability can sometimes be achieved by suitable preservation techniques, but these should be used only if specified in the analytical methods of interest).

It is convenient to use standard solutions when estimating precision. Standards of any desired concentration can be obtained, so that a range of concentrations is available for the estimates of precision; real samples of the desired concentration may not be available. However, the analyst should also have estimates of the precision for water samples as it should not, in general, be assumed that standard solutions and water samples can be analysed with the same precision. Therefore, precision estimates for samples and standards should normally be obtained.

For these tests, standard solutions and samples should be analysed, measured and evaluated in exactly the same way as normal routine samples.

When the limit of detection is of interest, a solution containing essentially none of the determinand should also be included. In many cases, the blank determination is suitable.

These various requirements may seem rather complex, but the worked example shows that they may be simply resolved. Clearly, the greater the number of different solutions included in the tests, the greater the information obtained on precision, but a compromise with the effort required will often be necessary. As a guide to the minimum number of solutions, it is suggested that the following test samples should be included in each batch of analysis.

8.3.1.1.1 Two standard solutions or samples at concentrations near the upper and lower concentrations of interest.

8.3.1.1.2 When standard solutions are used for **8.3.1.1.1**, one water sample near the average concentration in samples.

8.3.1.1.3 A spiked water sample; the sample analysed in **8.3.1.1.2** with the addition of a known quantity of determinand. If estimation of precision at a variety of different concentrations is of key interest, the level of the spike should be chosen so that the final concentration differs from those of the other solutions. Otherwise, it is advisable to make as large an addition as the initial sample concentration and the range of the method will allow.

8.3.1.1.4 Whenever an estimate of limit of detection is required, n replicate blank samples should be analysed. A blank sample, as distinct from an analytical reagent blank, is a natural sample which contains no determinand. It may not be possible to obtain a sample which approaches this ideal. For most determinands, under these circumstances, an analytical blank solution may be used as a substitute. If precision at the blank level is known to be dependent on the sample matrix, it will be necessary either to use a blank sample which contains the determinand (and risk a likely overestimation of limit of detection) or to take steps to remove the determinand from a sample so that it may be used as a blank. When, as with some chromatographic techniques, no response is obtained for a blank, it is recommended that a blank is spiked with enough determinand to produce a measurable response. This can form the basis for an estimate of limit of detection. (The measured values should not, of course, be used for blank correction.)

In any event, at least a single analytical blank per batch should be included as one of the solutions analysed to allow calibration.

8.3.1.2 *Precision tests for unstable samples*

The simplest approach to the design of precision tests is to prepare all samples for analysis at the start of the tests and use these without preparing fresh aliquots for each batch of analysis.

This is satisfactory provided there is no sample instability. The possibility that sample instability may be present rules out the direct estimation of between-batch standard deviation and may call into question the assessment of within-batch standard deviation. Further discussion is reported in [18].

8.3.2 ***Statistical techniques and calculations***

8.3.2.1 *Randomization*

Randomization of the order of analyses should normally be used to eliminate the effects of any systematic changes in factors that cannot be controlled, and which might otherwise cause false conclusions to be drawn.

8.3.2.2 *Rounding of data*

It is usually desirable to record the primary experimental results with the greatest discrimination possible; three significant figures should be aimed for, except when near the limit of detection when only one or two will often be possible.

8.3.2.3 *Calculating analytical results*

Standard deviations should be calculated from the set of results for each sample. Thus, for each solution analysed, two results are available from each batch, corresponding to the first and second portions of the sample to be analysed. These results should, if necessary, be blank corrected using the analytical blank for the appropriate batch. The within-batch standard deviation of the blanks is used to calculate the limit of detection.

8.3.2.4 *Estimating precision*

It is useful to obtain estimates of the within-batch and between-batch standard deviations, s_w and s_b, respectively. These two estimates are needed to allow an estimate of the total standard deviation, s_t, to be obtained. A statistical technique known as “analysis of variance” is used. The theoretical basis of the technique is described in statistical texts, but in the present context it may be taken simply as a convenient means of calculating $s_{w'}$, s_t, and s_t.

8.3.2.5 *Testing standard deviations*

If the total standard deviation should not be significantly greater than some target value, Z, a variance ratio test, at a specified significance level, is used. However, an estimated number of degrees of freedom, DF, for the total standard deviation should first be calculated. For n results in each of m batches:

$$DF = \frac{m(m-1)\left[M_1 + (n-1)M_0\right]^2}{mM_1^2 + (m-1)(n-1)M_0^2}$$

where M_0 and M_1 are the within-batch and between-batch mean squares, respectively, obtained from the analysis of variance (see **8.3.3**). The calculated value for DF may be nonintegral; if this is so, the nearest whole number should be used for DF in the following test.

The variance ratio $F = s_t^2/Z^2$ is calculated, (clearly, this calculation is not needed if $s_t \leq Z$) and compared with the tabulated value $F_{a'}$ using DF and ∞ degrees of freedom for the numerator and denominator, respectively. The value of α appropriate to each situation should be used; a value of 0,05 should usually be suitable for analytical applications, s_t is taken to be significantly greater than Z if the calculated value for F is larger than the tabulated value. If it is found that s_t is significantly greater than Z, steps should be taken to identify and eliminate the important sources of error. If s_t is appreciably greater than Z but not significantly so, it is desirable either to carry out more tests to obtain a better estimate of s_t or to attempt to reduce the most important sources of error.

Variance ratio tests may also be used to test whether within-batch standard deviations are significantly greater than given target values. The general procedure is as for the total standard deviation, but the experimental estimates of the within-batch variances have $m(n\text{-}1)$ degrees of freedom.

8.3.2.6 *Evaluation of "spiking recovery"*

Consider the example where n determinations of both "spiked" and "unspiked" samples are made in each of m batches. To show that the mean recovery does not differ significantly (at the significance level α) from (100 ± D) % (where D is the accepted limit for recovery) of the amount added, the following procedure is used.

The recovery should be calculated from the difference between the results for the n pairs of "spiked" and "unspiked" samples in each batch. The mean recovery, *Rec*, should be calculated from the m n results. Also the standard deviation, s, of the m mean recoveries should be calculated for the m batches. Let the amount added for "spiking" be d (in the same units as *Rec*). The mean recovery is then acceptable if:

— Rec > (1,00–0,01 D)d – $s\,t_{2\alpha}/m^{½}$ (for *Rec* < 100 %)

— Rec < (1,00+0,01 D)d + $s\,t_{2\alpha}/m^{½}$ (for *Rec* > 100 %)

where $t_{2\alpha}$ = Student's t with (m-1) degrees of freedom, at the 2α probability level.

(Further discussion of recovery testing is given in **8.4**.)

8.3.3 ***Analysis of variance***

This calculation identifies the different sources of variation and allows the estimation of total standard deviation. It is a standard statistical operation and is most conveniently performed by computer. Details of manual calculation are given in most statistical textbooks.

As stated above, analysis of variance is used to give two statistical parameters, the within- and between-batch mean squares, M_0 and M_1, respectively. The mean squares are then compared to determine whether M_1 is significantly greater than M_0, that is, whether there is a statistically significant between-batch source of error. The results of an analysis of variance are usually presented in the form of a table, a general example of which is given in Table 2.

Table 2 — General analysis of variance

Source of variability	Sums of squares	Mean squares	Degrees of freedom	Quantity estimated by mean square
Between batches	$n\sum_{i=1}^{m}\left(\bar{x}_i - \bar{\bar{x}}\right)^2 = S_1$	$S_1/(m-1) = M_1$	$m-1$	$n\sigma_b^2 + \sigma_w^2$
Within batches	$\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2 = S_0$	$S_0/[m(n-1)] = M_0$	$m(n-1)$	σ_w^2
Total	$\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - \bar{\bar{x}}\right)^2$		$nm-1$	

Where:

n	=	number of replicate analyses within a batch
m	=	number of batches of analysis
$\bar{\bar{x}}$	=	overall mean value
$\bar{x}i$	=	mean of ith batch
$\bar{x}ij$	=	jth replicate analysis in ith batch
σ_w	=	within-batch standard deviation
σ_b	=	between-batch standard deviation
M_1	=	between-batch mean square
M_0	=	within-batch mean square

8.3.3.1 *Examples of calculations*

In order to illustrate the principle of this relatively simple application of analysis of variance, details of calculations using a calculator with standard deviation function are given. It assumes that the calculator is used to compute mean results and standard deviations only. The remainder of the calculation is also shown.

A practical way of carrying out the analysis of variance is presented below.

m = Number of batches = 10, indices for batches, $i = 1...m$

n = Number of repeats within each batch = 2, indices for repeats, $j = 1...n$

— Calculate the batch means ($\bar{x}i$) and standard deviations (s_{wi}) from the n repeats for each of the m batches.

— Calculate the pooled estimate of the within-batch variance s_w^2

$s_w^2 = (\Sigma_i\, s_{wi}^2)/n$ is the best estimate of the true within-batch variance, σ_w^2 , with $m(n-1)$ degrees of freedom.

— Calculate the variance of the batch means, s_{bm}^2

s_{bm}^2 is the estimate of $\sigma_b^2 + (\sigma_w^2/n)$ with $(m-1)$ degrees of freedom;

σb^2 is the variance of the between-batch random error;

— Test with F-test to see if the between-batch variance is significant, i.e. if it is significantly larger than within-batch variance.

$F = sbm^2/(sw^2/n)$, which is the estimate of $[s_b^2 + (s_w^2/n)]/(s_w^2/n)$ with $(m-1)$ and $m(n-1)$ degrees of freedom for the numerator and denominator respectively.

— If the test is significant, variance of the between-batch error can be estimated as

$s_b^2 = s_{bm}^2 - (s_w^2/n)$

— The estimate of the total variance of a single determination, if s_b is significant, is

$s_t^2 = s_b^2 + s_w^2 = s_{bm}^2 + (n-1)\, s_w^2/n$

otherwise $s_t^2 = s_w^2$ if $s_b = 0$.

NOTE M_1, the between batch mean square produced by computer analysis of variance is equivalent to $n.s_{bm}^2$; M_0 is equivalent to s_w^2.

EXAMPLE 1

Batch	x_{i1}	x_{i2}	$\bar{x}_i$	S_{wi}^2
1	223	256	239,5	544,5
2	214	210	212,0	8,0
3	199	208	203,5	40,5
4	206	232	219,0	338,0
5	228	226	227,0	2,0
6	231	227	229,0	8,0
7	226	226	226,0	0,0
8	227	226	226,5	0,5
9	229	202	215,5	364,5
10	215	225	220,0	250,0

Pooled variance $s_w^2 = 135{,}6$ is the estimate of σ_w^2, the within-batch variance, with degrees of freedom, $DF = 10$.

Variance of means $s_{bm}^2 = 101{,}5$ is the estimate of $\sigma_b^2 + (\sigma_w^2/2)$

The means are affected by the between-batch and within-batch error. In estimating the variance, the latter has to be divided by the number replicate analyses, 2 in this case.

The significance of σ_b can be tested with an F-test.

$$F = \frac{s_{bm}^2}{s_w^2}$$

$\sim [\sigma_b^2 + (\sigma_w^2/2)] / (\sigma_w^2/2)$
$= 1{,}497 < F\,(95\ \%) = 3{,}02$ (9 and 10 degrees of freedom)

This is not significant with 9 and 10 degrees of freedom, and it can be concluded that the estimate of the between-batch variance

$$s_b^2 = s_{bm}^2 - (s_w^2/2) = 101{,}5 - 135{,}6 / 2 = 33{,}7 \quad \text{and } DF = 0{,}7, \text{ i.e. } \sim 1$$

is not significantly different from zero and it is accepted that the hypothesis $\sigma_b = 0$ and the total standard deviation $s_t = s_w = 11{,}6$.

EXAMPLE 2

Batch	x_{i1}	x_{i2}	$\bar{x}_i$	s_{wi}^2
1	293	296	294,5	4,5
2	293	291	292,0	2,0
3	282	281	281,5	0,5
4	293	292	292,5	0,5
5	296	293	294,5	4,5
6	301	299	300,0	2,0
7	291	295	293,0	8,0
8	286	263	274,5	264
9	291	291	291,0	0
10	295	291	293,0	8,0

Pooled within-batch variance estimate $s_w^2 = 29{,}45$, $DF = 10$.

Variance of the batch means $s_{bm}^2 = 53{,}11$, $DF = 9$.

F-test

$F = s_{bm}^2/(s_w^2) = 3{,}61$

which is significant with $m-1 = 9$ and $m(n-1) = 10(2-1) = 10$ degrees of freedom. Since the test exceeds the F-table value of 3,02 at 95 % confidence level, the between-batch standard deviation is significant compared with the within-batch standard deviation and can be estimated as

$s_b^2 = s_{bm}^2 - (s_w^2/2) = 38{,}39$ $DF = 4{,}4 \sim 4$

The total variance of a single determination is

$s_t^2 = s_b^2 + s_w^2 = 67{,}84$ $DF = 13{,}7 \sim 14$

Thus the results are:

Error estimate	**Case A**	**Case B**
S_w	11,6	5,43
S_b	0	6,20
S_t	11,6	8,24

8.3.4 *Calculation of degrees of freedom by the Satterthwaite formula*

The number of degrees of freedom associated with the "total standard deviation" estimate is

$$DF_{st} = \frac{\left(s_t^2\right)^2}{\frac{\left(s_{bm}^2\right)^2}{m-1} + \frac{(n-1)\left(s_w^2\right)^2}{mn^2}}$$

8.4 Recovery tests

8.4.1 *Principle of test*

The object of the test is to identify bias from certain sources occurring in the analysis of real samples. A known quantity of the determinand is added to a real sample, forming the spiked solution, and the two are analysed, the difference in concentrations found being used to calculate the recovery. This is repeated n times and the mean differences compared statistically with the theoretically expected recovery.

8.4.2 *Analysis of results*

Since the spiked solution is made up by adding a fixed quantity of standard solution to a fixed quantity of real sample, the calculation of its recovery can be made as follows:

Standard solution:	volume v, concentration c (known)
Real sample:	volume V, concentration u (measured)
Spiked sample:	volume $(v + V)$, concentration s (measured)

Recovery, *Rec*, is the percentage of the added determinand which is determined.

$$Rec = [s(v + V) - uV] \times 100\ \%/(cv)$$

It should be emphasized that *Rec*, calculated from m n values (i.e. m batches, n replicate analyses in each batch), is only an estimate of the true mean recovery.

The standard deviation, s, is calculated from the m daily mean recoveries (each calculated from two replicates); s refers therefore to the standard deviation of m daily mean recoveries.

The standard error, S_{Rec}, of *Rec* is calculated from

$$S_{\mathrm{Rec}} = s/\sqrt{m}$$

where m is the number of values on which s is based. The standard error is, in fact, the standard deviation of an estimate of the mean (as opposed to the standard deviation of a single observation). The true mean can be expected to lie within $\pm\, t_{0.05}S_{\mathrm{Rec}}$ of the estimated *Rec* with 95 % confidence.

In the test of acceptability, the normal requirement is that *Rec* is not outside the range of 95 % to 105 %, and that a recovery is unsatisfactory when it is 95 % certain that it does not comply with that condition.

8.4.3 *Interpretation*

It should be noted that the spiking recovery test is fairly limited in the information it yields. For example if bias is found in the standard solution results, it is quite probable that it will also occur in the spiking recovery results and yield no additional information. It only assumes importance when significant bias does not occur elsewhere. In this case, the implication is a cause of bias in the real sample only, and this usually implies interference proportional to the concentration of the determinand. (Clearly an interference effect of absolute magnitude would not affect the difference between spiked and real samples). In the case of unsatisfactory spiking recovery, it is advisable to check the precision of the real and spiked results, particularly if the spiked solution has not been freshly prepared for each analysis. If either of the two solutions shows signs of deterioration, this could easily produce an unsatisfactory spiking recovery.

8.4.4 *Worked example*

A water sample was analysed in duplicate before and after spiking on each of 10 days. The results are given in Table 3.

Table 3 — Example of recovery with spiked sample

Day	Real sample mg/l	Spiked sample mg/l	Recovery %	Batch mean recovery %
1	7,5	15,0	82,5	
	7,5	17,0	102,5	92,50
2	6,5	15,5	96,5	
	8,0	16,0	88,0	92,25
3	9,0	17,0	89,0	
	8,5	15,0	73,5	81,25
4	7,5	17,5	107,5	
	8,0	16,5	93,0	100,25
5	8,0	17,0	98,0	
	9,0	17,5	94,0	96,00
6	6,5	16,0	101,5	
	7,0	15,0	87,0	94,25
7	8,5	16,0	83,5	
	7,5	16,0	92,5	88,00
8	9,0	17,0	89,0	
	8,0	17,5	98,5	93,75
9	9,0	16,0	79,0	
	8,0	16,5	93,0	86,00
10	7,5	17,5	107,5	
	9,0	18,0	99,0	103,25
Mean	**8,0**	**16,475**	**92,75**	

The spiked solution was made up of 10 ml of standard solution, of 100 mg/l concentration, made up to 100 ml with real sample.

Therefore (using the equations given above):

Mean recovery, *Rec* = (sum of daily means)/10 = 92,75 %

Standard deviation, *s*, of 10 daily means = 6,506 %

Standard error = $s/\sqrt{10}$ = 6,506/3,1622 = 2,058 %

If the true mean is 92,75 %, then 90 % of estimates of that mean would lie in the range:

92,75 ± (1,833 × 2,058) where 1,833 is from Student's *t* distribution, that is, between 88,98 and 96,52.

The observed recovery is therefore not significantly (α =0,05) outside the range 95 % to 105 %.

9 Intralaboratory quality control

9.1 General

The previous clause deals with the evaluation of the capabilities of an analytical technique in order to judge its likely suitability for a particular application. This clause describes the procedure to be adopted when the system is put into routine use. The progression from choosing a method (whether on the basis of published data or by in-house testing) to its use in routine analysis involves a change in perspective. In the former case, the choice of method is often looked at in isolation; in the latter, the emphasis is usually transferred to an examination of the performance of the whole analytical system, of which the method is only a part. All the components of the analytical system — instrumentation, analysis, laboratory facilities, etc. — should be examined critically before routine analysis is started. The emphasis given to the choice of method is, however, not unwarranted since this is usually the factor in the analytical system which is of crucial importance and, of course, it is the one with which there is likely to be the greatest flexibility.

Having chosen an analytical system capable of being used to produce results of adequate accuracy, the next stage is to establish control over the system and to monitor routine performance. The aim is to achieve a continuing check on the errors in routine analysis. This provides both a means of detecting error as results are produced and a demonstration of the accuracy of data.

Within-laboratory quality control thus involves two stages.

— Experimental estimation, and when necessary reduction, of error; this stage should be completed before any samples are analysed and may be called "preliminary error estimation" (**8.2**).

— When a method has been put into routine use, regular checks should be made to ensure that control over the system is maintained; this stage is termed "routine quality control". The principal tool in routine quality control is the control chart.

9.2 Terms relating to within-laboratory quality control

Control sample:	Sample material whose analytical results are used to construct control charts, for example, standard solutions, real samples, blank samples.
Control value:	Value entered on the control chart.
Response value:	Value obtained by application of the measurement process.
Analytical result:	Value reported as defined in the method. It is derived from the response by application of the calibration.

9.3 Control of accuracy

The total error associated with an analytical result has components of random and systematic errors. Both these sources of error can be controlled on a routine basis. Inaccuracy or analytical error is untrueness and imprecision, i.e. a combination of random and systematic errors.

It is not sufficient for a laboratory to adopt a suitable method, check its performance initially and assume that thereafter the results produced will be of adequate accuracy. The chosen method should be subject to routine tests to ensure that adequate performance is maintained.

The control of accuracy can be carried out using control charts. The simplest form of control chart is one in which the results of the individual measurements made on a control sample are plotted on a chart in a time series.

This type of chart (for example see Figure 2) provides a check on random and systematic error (from the spread of results and their displacement). It is an easy procedure to be used by the analyst because it is simple to plot and no data processing is needed. It is useful when the size of analytical batches is variable or when batches consist of a small number of determinations. Individual result charts are used widely and often form the mainstay of a laboratory's approach to control charting.

However, this type of chart may produce false out-of-control values if random error does not follow the normal distribution. For these reasons, a range of more specialized types of chart has been devised. These are described below.

9.4 Control of trueness

One way of assessing systematic error is to participate regularly in interlaboratory trials.

As a routine procedure for controlling systematic error, the use of Shewhart control charts based on the mean, spiking recovery and analysis of blanks is recommended.

9.4.1 *Mean control chart*

For trueness control, standard solutions, synthetic samples or certified real samples may be analysed using a Shewhart chart of mean values.

The analysis of standard solutions serves only as a check on calibration. If, however, solutions with a synthetic or real matrix are used as control samples, the specificity of the analytical system under examination can be checked, provided an independent estimation of the true value for the determinand is available.

The respective control sample should be analysed a fixed number of times in each batch of samples and the mean result entered in the mean control chart.

It is advisable to analyse certified reference samples (if suitable ones are available and are not too expensive) with routine samples as a check on trueness. A restricted check on systematic error by means of recovery control charts is often made instead (**9.4.2**).

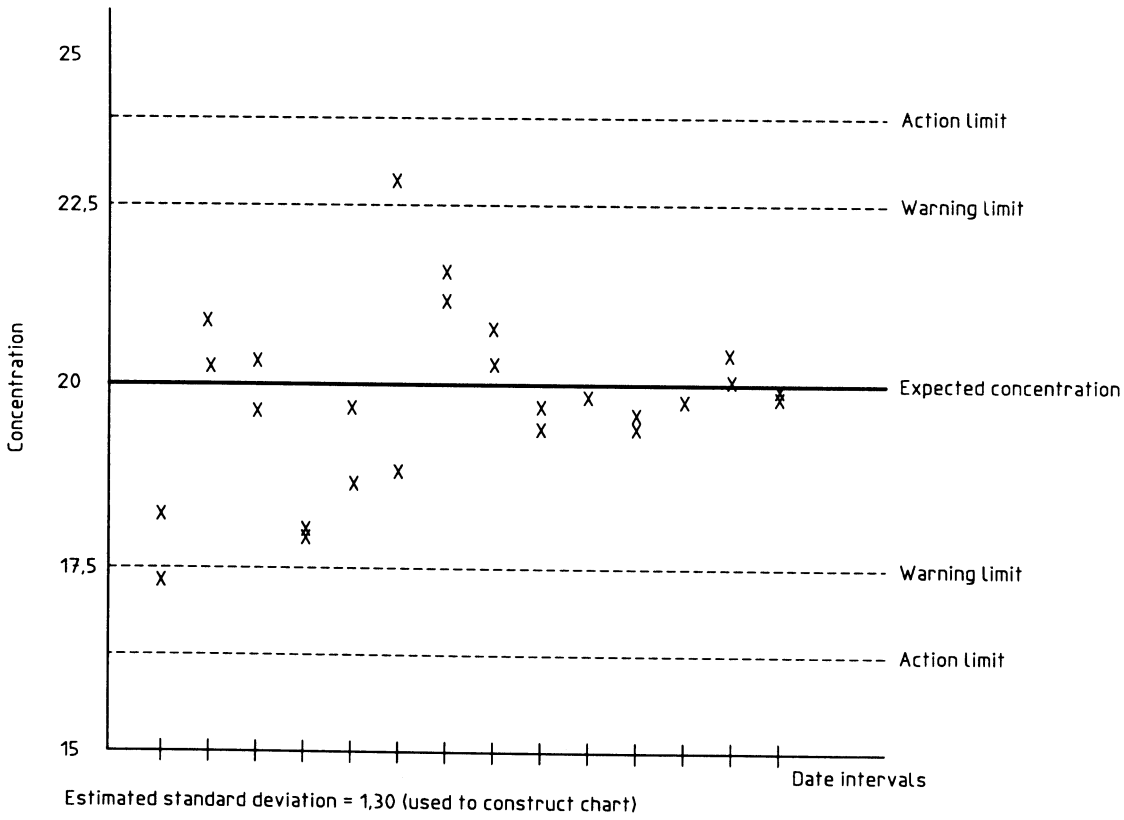

Figure 2 — Example of Shewhart control chart

9.4.2 ***Recovery control chart***

The recovery control chart is used as a check on systematic errors arising from matrix interferences. A separate control chart for each type of matrix is required in water analysis, because samples of strongly varying matrix composition, such as surface water, municipal and industrial waste water, may be subject to errors of differing sizes and natures.

The recovery control chart, however, provides only a limited check on trueness because the recovery tests will identify only systematic errors which are proportional to determinand concentration; bias of constant size may go undetected.

9.4.3 ***Blank control chart***

The blank control chart represents a special application of the mean control chart.

The blank control chart may help to identify the following sources of error:

— contamination of reagents;

— contamination of reaction vessels and of the measurement system;

— instrumental faults (for example baseline drift).

It is appropriate therefore, to analyse a blank solution at the beginning and at the end of each batch of samples. The blank values thus obtained are then entered on the blank control chart.

9.5 Control of precision

9.5.1 ***General***

There are four ways of controlling the precision of analytical results in routine analysis:

— use of a range control chart (**9.5.2**);

— replicate determinations (**9.5.3**);

— standard addition (**9.5.4**);

— use of the mean control chart (**9.4.1**).

9.5.2 *Range control chart*

A range control chart is used to control the precision of an analytical method. In addition, it allows some assessment of errors caused by calibration drift. The standard deviation for a certain analytical result can be estimated from an existing range control chart, provided the matrix of the sample under examination is similar to that of control samples chosen for the range control chart. The range of the sample in question may also be determined and entered on the control chart as well, in order to prove that an out-of-control situation does not exist.

Estimation of the standard deviation, s, from range control charts:

$$s = \frac{\overline{R}}{d_2}$$

where

d_2 is a factor (see Table 4) and

$\overline{R}$ is the mean range

Table 4 — d_2 values

n	d_2
2	1,128
3	1,693
4	2,059
5	2,326
6	2,534
7	2,704
8	2,847
9	2,970
10	3,078
where n is the number of replicate analyses in each batch (reference 5)	

9.5.3 *Estimation of precision with replicate analysis*

It is highly recommended that the analyst perform replicate analyses of the sample in question to obtain higher reliability of the final result, especially in those cases where the contravention of a threshold value is to be proved. From the data obtained, the standard deviation valid for the matrix in question can be estimated.

Additionally, the performance of replicate determinations offers two further advantages: firstly, coarse errors (outliers) can be detected, and secondly, the analytical error can be reduced.

A number where $n \geq 6$ is recommended.

The standard deviation is calculated as:

$$s = \sqrt{\sum_{i=1}^{i=n} \frac{\left(X_i - \overline{X}\right)^2}{(n-1)}}$$

for replicate measurements

9.5.4 *Standard addition*

The estimation of the standard deviation from the range control chart, or with replicate analysis, can help to identify a matrix-dependent imprecision.

The application of the method of standard addition, whilst helping to control untrueness, can tend to degrade precision compared with direct determination. This is the price paid for control over systematic error. The method of standard addition should be applied with caution. It is essential that the linear range of the method be established.

9.6 Principles of applying control charts

9.6.1 ***Choice of control samples***

The choice of control samples depends on the matrix, the analytical method and the accuracy required. Advantages and disadvantages of the several types of control samples are described in [13] and [16].

9.6.2 ***Construction of control charts***

9.6.2.1 *Construction of mean and blank control charts*

9.6.2.1.1 *Preliminary estimation of standard deviation*

At least 20 mean control values, X_i are required for a trial period to estimate the following tentative control parameters. They are obtained by analysing the control sample on at least 10 working days in duplicate (see clause **8**).

From the control values x_i ($n \geq 20$) estimate the statistical parameters:

— x_i (mean of the replicate analyses of the ith batch);

— mean (x);

— standard deviation (s);

— upper warning limit and lower warning limit (UW, LW) = $(\bar{x}) \pm 2s$;

— upper action limit and lower action limit (UA, LA) = $(\bar{x}) \pm 3s$;

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

$$s = \sqrt{\sum_{i=1}^{i=n} \frac{\left(X_i - \bar{X}\right)^2}{(n-1)}}$$

The control chart is constructed in a coordinate system with the ordinate "concentration" and the abscissa "time of analysis" and/or "batch number". The numerical values for mean, warning limit and action limit are plotted on the ordinate and drawn as lines parallel to the abscissa in the control chart.

9.6.2.1.2 *Routine operation*

The control value should be obtained at least once per batch of analyses. The frequency with which control values are obtained within a batch is the responsibility of the laboratory and should be related to the risks of important errors and the seriousness of their likely consequences. At intervals chosen by the analyst, the control chart should be examined for changes in mean and standard deviation.

In the long-term operation of a control chart, the question arises whether or not to update the estimate of mean and standard deviation used to generate the action and warning limits and, if so, how this might best be done. The guiding principle should be that the chart is intended to detect (with known risks of making the wrong decision) departures from the existing state of statistical control. Including the latest data in the overall estimates of mean and standard deviation may not be sufficient to allow this aim to be fulfilled.

There are two approaches to monitoring for changes:

9.6.2.1.2.1 *Review of the last 60 data points*

It is assumed that the last 60 data points are a homogeneous set and that the issue is whether or not these points are of the same precision as that implied by the initial choice of control limits.

It is also assumed that the normal practice is to base the action and warning limits on a mean and standard deviation derived from all available data points (including the latest). Data points corresponding to "out-of-control" situations for which a definite cause has been identified should not, of course, be included in the calculations.

Review the last 60 data points on the chart. If there are between 1 and 6 (inclusive) cases where the 2*s* warning limits have been exceeded, there is no clear evidence that the precision of analysis has changed. No revision of the chart is required except, as usual, the incorporation of new data points into the estimates of s and μ.

If there are either no cases where the warning limits have been exceeded or more than 6 cases, it may be concluded with approximately 90 % confidence that the precision has changed (improved or degraded, respectively) and that a revision of the action and warning limits is needed.

In this case, recalculate the control limits on the basis of the mean and standard deviation of the last 60 points and proceed as usual.

Whenever new control limits are calculated as a result of a change in precision, review the new standard deviation (and where appropriate the bias implied by the new mean) against the accuracy targets which apply to the analyses in question. Take corrective action if necessary.

The above procedure need not be carried out every time a new data point is generated. This check on the validity of the current control limits might be worthwhile after, say, 20 successive points have been plotted – though any obvious changes in the operation of the chart would warrant immediate concern.

9.6.2.1.2.2 *Comparison of consecutive control periods*

This approach involves comparison of mean and standard deviation values from different control periods of predefined length.

Mean-*t*-test between the x values of two consecutive periods

A test of difference of the variances (*F*-test) for the two periods, and before and after updating

Statistically significant differences between two consecutive periods may indicate a change in analytical quality (better or worse). Any change should be reviewed with respect to the accuracy requirements.

9.6.2.2 *Construction of recovery control charts*

The design and the criteria of decision of the recovery control chart are similar to those of the mean control chart.

9.6.2.2.1 *Preliminary tests*

For the construction of a recovery control chart it is recommended to run a trial period of tests, or to obtain an estimate of the relevant parameters from the performance tests described in clause **8**.

The control variable to be plotted is Rec_i

$$Rec_i = (x_a - x_0)\ 100/c_a$$

where

X_a is the analytical result (for example concentration) of the determinand in the spiked sample;

X_0 is the analytical result (for example concentration) of the determinand in the original sample;

C_a is the concentration or mass respectively of the spiked determinand. This assumes negligible dilution of the sample by the spiked addition.

After completion of the trial period the following statistical characteristics are derived from the recoveries Rec_i ($n \geq 20$):

— mean recovery ($\overline{Rec}$);

— standard deviation of the mean recovery (S_{Rec});

— upper warning limit and lower warning limit (UW; LW);

— upper action limit and lower action limit (UA; LA).

Calculation:

$$\overline{Rec} = \frac{1}{n}\sum_{i=1}^{i=n} Rec_i (\%)$$

$$s_{Rec} = \sqrt{\sum_{i=1}^{i=n} \frac{\left(Rec_i - \overline{Rec}\right)^2}{n-1}}\ (\%)$$

$$UW = \overline{Rec} + 2S_{\mathrm{Rec}}$$

$$LW = \overline{Rec} - 2S_{\mathrm{Rec}}$$

$$UA = \overline{Rec} + 3S_{\mathrm{Rec}}$$

$$LA = \overline{Rec} - 3S_{\mathrm{Rec}}$$

9.6.2.2.2 *Further processing*

The recovery chart is constructed and maintained in the same way as described in **9.6.2.1**. For the calculation of the statistical parameters, x and s should be replaced with Rec and s_{Rec}, respectively.

9.6.2.3 *Construction of range control charts*

At least 20 control values ($n \geq 20$) are required for the pre-period. Control value is the relative range $Rrel_j$:

$$Rrel_j = \frac{x_{i\,\max} - x_{i\,\min}}{\bar{x}_i} \bullet 100\,(\%)$$

where j is the jth batch of i replicates

with:

$$\bar{x}_i = \frac{1}{n}\sum_{i=1}^{i=n} x_i$$

where

x_i is the individual analytical result of the respective control sample;

n is the number of replicate determinations of the respective control sample.

After the completion of a preliminary test period, the relative range values $Rrel_j$ ($n \geq 20$) are used to calculate the following statistical parameters:

$$\overline{Rrel} = \frac{1}{n}\sum_{j=1}^{j=n} Rrel_j\,(\%)$$

$$UA = \overline{Rrel} \bullet D_{UA}\,(\%)$$

$$LA = \overline{Rrel} \bullet D_{LA}\,(\%)$$

Several calculation models may be used to estimate the action limits for this type of control chart. For application in routine work it is recommended that only the upper action limits (UA) be calculated. When performing replicate determinations (duplicate to six-fold), the lower action limit (LA) is identical with the abscissa (zero-line).

The numerical values for the factors D_{UA} and D_{LA} are:

	Duplicate determination	**Three-fold determination**	**Four-fold determination**	**Five-fold determination**
D_{UA} (P = 99,7 %)	3,267	2,575	2,282	2,115
D_{LA} (P = 99,7 %)	0	0	0	0

NOTE For further numerical values for the factors D_{UA} and D_{LA} refer to ASTM (1976).

9.6.3 ***Interpretation of control charts, out-of-control situations***

The quality control chart is intended to identify changes in random or systematic error.

The following criteria for out-of-control situations are recommended for use with Shewhart charts:

— 1 control value being outside the action limit; or

— 2 consecutive values outside warning limits; or

— 7 consecutive control values with rising tendency; or

— 7 consecutive control values with falling tendency; or

— 10 out of 11 consecutive control values being on one side of the central line.

The following out-of-control situations apply to the range type of control chart if:

— a range $Rrel_j$ falls outside the upper action limit; or

— a range $Rrel_j$ falls below the lower action limit (valid only for LA > 0); or

— 7 consecutive control values show an ascending/descending tendency; or

— 7 consecutive control values lie above the mean range *Rrel*.

A cyclic variation of ranges may be observed, for example, by a regularly scheduled maintenance of an analytical instrument or by re-preparation of reagents.

9.7 Conclusions

9.7.1 *Measures to be taken in out-of-control situations*

An out-of-control situation occurring on a control chart implies that an important error might apply to the analysis of the routine samples. It is very important to identify and eliminate the cause of the error in order to maintain control over the performance of the analytical system. For fast and effective identification of the source of analytical error, the approach described in the following subclauses is recommended.

9.7.1.1 *Elimination of gross errors*

The analysis of the control sample is repeated, strictly following the analytical method and avoiding possible gross errors. If the new result of the control sample shows that the method is under control again, it may be assumed that the method of analysis had not strictly been observed on the previous batch of analyses or that a gross error had occurred. The entire batch should then be re-analysed.

If, however, the result of the analysis of the control sample is erroneous but reproducible, a systematic error is very likely to exist.

9.7.1.2 *Elimination of systematic errors*

To check for systematic errors, several different trueness control samples are analysed. To detect errors depending on the reagents or the method, control samples should be used whose concentrations cover the entire measuring range. As a minimum, a trueness control sample in the lower and one in the upper part of the working range should be used. In the event of a systematic error with results predominantly being higher or lower than the actual values, a step by step examination should be performed to find the reason for this bias. Exchanging experimental parameters, such as reagents, apparatus or staff, might help to identify quickly this type of error.

9.7.1.3 *Improving precision*

The precision can also be improved by a step-by-step approach to find the causes of random error.

The total precision of an analytical method can be improved by examining its individual procedural steps to find the one which contributes most to the total error.

9.7.2 *Plausibility control*

There could be errors which may not be detected by a statistical approach to quality control. In most cases, this concerns errors influencing individual analyses in a batch, but not ones before or after. This type of error can only be revealed by means of plausibility controls — checks on the observed value in relation to expectations based on previous knowledge. Such knowledge may be based on chemical consideration, for example checks on the equivalence of anions and cations in a sample, or on prior expectation, for example that COD will be greater than BOD.

A successful approach to plausibility control requires that appropriate background information is available. The procedure of plausibility control may be subdivided in two parts:

— Information/harmonization

— Control (details may be found in [13].

Plausibility control may form a worthwhile additional check to supplement routine AQC. A large proportion of failures on the basis of plausibility control (which is not mirrored by routine AQC) suggests an inadequate routine system of quality control or a system which is not stable in its operation.

9.7.3 ***Further corrective action***

In the event of repeatedly occurring out-of-control situations being detected in the control charts, the initial tests for implementation of analytical quality control, as described in clause **8**, should be performed with the matrix in question, if the out-of-control situations cannot be remedied by simpler actions, such as exchange of vessels, apparatus or reagents.

10 Quality control in sampling

10.1 General

Careful attention to the soundness of sampling and sample handling systems is essential if data of adequate accuracy are to be obtained, and it is therefore necessary to ensure that appropriate control tests are applied to these aspects of the overall process, as well as to analysis. Whilst many issues concerning sampling and sample handling lie outside the scope of this Technical Report, it is fitting to complete this Technical Report with an outline discussion of several important aspects of routine control testing of sampling and sample handling.

It is emphasized that, as with the analytical stage, the initial selection of soundly based sampling procedures is of fundamental importance. Indeed, given the difficulty of assessing by practical tests many of the potential errors which may arise during sampling, the need for careful initial selection of equipment and procedures is probably even more crucial than in analysis. Similarly, control tests of sampling and sample handling have the same basic objectives as their counterparts in analysis, namely to ensure that any important deterioration of the accuracy of results arising from these steps is detected as rapidly as possible, so that corrective action can be taken.

In addition to general "good practice" aspects of routine quality control in sampling (for example checks and preventative maintenance on sampling equipment), the specific control tests described below should be considered and put into practice wherever appropriate and practicable. A guiding principle in their application is that a sound approach to quality control should cover as many potential sources of error as possible in sampling and sample handling.

10.2 Routine tests on effectiveness of cleaning sampling vessels and sample containers

Whilst field blanks (see below) give some check that such vessels and containers do not cause important contamination of samples, laboratory tests have the advantage that they can be routinely undertaken before sampling is performed; thus, if contamination problems are revealed, they can be rectified before sampling, thereby saving potentially wasted effort and resources.

10.3 Field blanks to provide routine checks on contamination

Field blanks are samples of (typically) deionized or distilled water which are taken into the field and treated, so far as possible, in exactly the same way as real samples. The exact details of the approach to be followed will, therefore, vary according to the particular system being controlled, but field blanks should generally be subjected to the same preparatory steps (such as filtration and centrifugation) as are applied to real samples, and should subsequently be handled, preserved and stored in the same way.

10.4 Field controls to provide routine checks on sample stability

In situations where, despite careful initial selection and testing of equipment and procedures, the stability of samples is in question, it can be useful to prepare control samples of known determinand concentration and treat them, so far as possible, in exactly the same way as real samples. Such a control sample may be prepared by dividing a typical sample into two and making a known addition to one portion. The recovery of the added determinand verifies that sample preservation, transport and storage are satisfactory and that loss of the determinand — by adsorption or evaporation of volatile components, for example — is adequately controlled.

10.5 Duplicate samples to provide routine checks on sampling precision

Collection and analysis of duplicate samples can provide evidence on the contribution of sample collection and handling to overall random error.

10.6 Control chart of field blanks

A routine chart of field blanks may be a valuable way of monitoring control over sample contamination.

10.7 Conclusions

Control samples of types **10.3**, **10.4** and **10.5** bear similarity to some of the analytical control samples described previously. Indeed, when analysed they will inevitably cover the sources of analytical error controlled by those samples (as well as the potential sources of error in sample collection and handling that they are specifically intended to control). However, their use should not be regarded as a substitute for the use of the relevant analytical controls, because they can only be fully effective in controlling errors in sample collection and handling if the analytical process itself is under separate and effective control.

11 Interlaboratory quality control

11.1 General

The design of an interlaboratory test and the way in which the results are interpreted should take account of the context in which the test is performed; it should reflect the aims of the analytical work to which it relates. It is particularly important, before an interlaboratory test is organized, that the objectives be carefully examined. Such considerations will form the basis of the approach which should be adopted and will provide both the rationale for laboratories' participation and the basis for an acceptable means of interpreting the data produced.

As far as the individual laboratory is concerned, the usual reason for taking part in an interlaboratory test is to supplement its within-laboratory quality control as a means of detecting and guarding against undiscovered sources of error.

The aims of interlaboratory testing may be divided into two groups: collaborative studies and proficiency testing. Within both groups, several subgroups may be identified. The most important of these are discussed below.

11.1.1 *Aims of interlaboratory tests*

11.1.1.1 *Collaborative study to test capabilities of an analytical method*

A collaborative study is an essential feature of the validation of a candidate method for standardization. Such a study will show whether the method allows a suitably chosen cross-section of experienced laboratories to obtain comparable results on the same samples.

A collaborative study as such gives no help to individual laboratories which wish to identify (and where necessary control the size of) the different sources of analytical error affecting their results. For this purpose, other types of interlaboratory test should be applied (see below).

11.1.1.2 *Cooperative study to obtain a general picture of analytical errors existing in a group of laboratories (possibly with several methods)*

It might be necessary to estimate the accuracy of data produced by a group of laboratories which share a common interest. For example, the laboratories concerned might be contributing data to a monitoring programme or they may be the only ones analysing a given type of sample. In this case, the aim may be achieved by the simple approach of circulating samples for analysis. The number of samples distributed need not be great and the extra work requested of participating laboratories is relatively small. An approach which has been described in [24], known as the paired sample technique, provides a valuable means of summarizing the results of an interlaboratory test in graphical form. This has been widely used to give a picture of the size and to some extent the nature of errors between laboratories.

11.1.1.3 *Interlaboratory study to certify a standard reference material*

A special case of the use of interlaboratory tests is to arrive at a consensus (certified) value for the composition of a reference test material. This approach is only of value if a group of laboratories of proven expertise take part.

11.1.1.4 *Interlaboratory study to establish a consensus value to a reference material*

Reference materials, in contrast to certified reference materials (CRM) are not traceable to national or International Standards. Thus the establishment of a consensus value for these materials is of more restricted scope than for the certified reference materials and does not require such rigorous choice of participants.

11.1.1.5 *Proficiency tests*

Proficiency tests are to ensure that each laboratory of a group of laboratories achieves an acceptable standard of analytical accuracy — that analytical errors are controlled within adequately small bounds.

This is the objective towards which most programmes of interlaboratory testing are directed, either explicitly or implicitly. As the complexity of interlaboratory tests — in terms of the numbers and types of samples distributed and the work required of laboratories taking part — is increased, it becomes possible to draw more conclusions concerning the sources and nature of errors which may be present. On the basis of this knowledge it is then possible to direct efforts towards achieving the desired level of accuracy. In achieving this objective, the importance of a sound programme of within-laboratory quality control cannot be over stated. As given above, the interlaboratory test supplements within-laboratory activity, providing a means of detecting and guarding against undiscovered sources of error and acting as a demonstration of the accuracy achieved.

11.2 Between-laboratory quality control

Since laboratories following the recommended approach to quality control will have first controlled their random errors, attention will need to be directed towards systematic errors. Interlaboratory tests are costly to organize and should preferably be undertaken only when the activities described in earlier clauses of this Technical Report have been successfully completed.

It is highly desirable that results from interlaboratory tests organized in different regions or by different organizers are comparable, and it is therefore recommended that international protocols and standards be used.

11.2.1 ***Features of interlaboratory tests***

Whatever the purpose of an interlaboratory test, there will be certain common issues which should be considered. Guidance cannot be given here which is appropriate in all cases, but participants in or the organizers of interlaboratory tests may find it useful to consider the following points.

11.2.1.1 *General considerations*

— organization

— general information to participants

— determinands of interest

— number of participating laboratories and how they are selected

— the way the test is financed

— timetable for analysis and reporting.

11.2.1.2 *The test sample(s)*

— type of sample

— number of samples to be distributed

— range of concentration of determinand(s) to be covered by the samples sent out; range of interest to participants.

11.2.1.3 *Sample preparation*

— how to ensure sample homogeneity and stability

— how to preserve the sample

— whether or not to use sample concentrates

— whether to use split-level samples or uniform-level samples

— whether or not to use reference materials.

11.2.1.4 *Analysis and reporting*

— preparation of written instructions to be followed by participants

— other information required from laboratories

— number of replicate analyses required from each participant

— choice of analytical method.

11.2.1.5 *Evaluation of the test*

— how to determine the nominal or reference concentration

— mathematical/statistical treatment of the data

— assessment of performance

— performance criteria

— form in which the test is reported.

For further details, the following standards and guidelines can be consulted: AOAC (1975), ISO 5725:1990. Furthermore, accreditation bodies using the data from interlaboratory tests have requirements on the organization of such tests (WELAC 1992).

12 Quality control for lengthy analytical procedures

Some multistage analytical procedures, for example the determination of trace organic contaminants, are capable of producing relatively few results at a time. This raises the question of how to implement quality control measures which were initially put into practice with high-throughput techniques. The argument that because organic analyses are time-consuming they should not be subject to performance tests of the same complexity as, for example, nutrient determinations is unsound.

An analytical result which takes hours to produce should be supported by performance and quality control information of at least the same reliability as that associated with "simple" determinations. Indeed, because trace analysis is subject to greater uncertainty and is more costly to repeat, it can be argued that proportionally more effort needs to be directed towards quality control. The maxim that relative few results of known and adequate accuracy are better than many results of unknown and probably inadequate accuracy remains true.

It is essential, therefore, that the level of performance testing carried out for trace analyses is at least equivalent to that recommended in clause **8**. The stated approach to tests of precision and recovery should not be regarded as an ideal only attainable under favourable circumstances. Rather, it is the minimum of testing which will provide a modestly reliable indication of performance. For trace analysis, there is a strong case for expanding the range of samples tested to include checks on precision and recovery from samples of differing matrices. Where limit of detection is of special interest, it is particularly important that an estimate is obtained from multibatch tests, rather than from replicate determinations performed on a single occasion.

Similarly, the approach to routine quality control should follow the recommendations given in clause **9**. Particular attention should be paid to the implementation of recovery control charts.

12.1 Analysis carried out infrequently or on an *ad hoc* basis

The procedures recommended for preliminary performance tests (clause **8**) and routine quality control (clause **9**) are most easily put into practice for analyses which are carried out regularly and often. It is necessary to consider what approach to quality control should be adopted for analyses which may be performed infrequently or which may be undertaken only once. The same considerations apply to analyses carried out over a short period in relatively few batches.

Two main features distinguish this type of analysis from frequent, regular determinations. Firstly, any quality control activity is likely to take up a relatively large proportion of the total analytical effort compared with routine analyses. This is inconvenient and expensive, but it is a consequence of organizing analysis in this way. It should not be used as an excuse to avoid evaluation of the analytical system. Any analytical system used to produce data should be tested to provide an estimate of its performance. Not to test would be to provide data of unknown accuracy. This is unacceptable to users of analytical data. Tests as described in clause **8** are recommended as a means of providing background performance data for all analytical systems.

Secondly, it is not possible to establish and maintain a state of statistical control in relatively few batches of analysis. This is an important drawback of not carrying out frequent, regular batches of analysis. It may be a consideration why analytical work might be subcontracted to laboratories which have reason to perform the determination in question frequently. However, when analyses are carried out on a one-off basis the following approach is recommended. As a minimum, quality control measures should include:

— checks on spiking recovery in the matrix of interest;

— replicate analyses of samples;

— use of field and procedural blanks;

— confirmation of the calibration using material from an independent source;

— use of reference materials (where an appropriate CRM is available) as blind controls.

The proportion of samples analysed more than once should not be less than 20 % but could be as large as 100 % in the case of very small batches or highly important analyses. Single analysis of samples is an acceptable approach only when a state of statistical control can be established and maintained.

Annex A (informative)
Evaluation of interference effects on analytical methods

A.1 General

This guidance is meant to provide bodies organizing interlaboratory collaborative studies with a statistical tool to assess the degree of interference of substances on an analytical method.

One of the most commonly occurring types of bias in an analytical method, in the analysis of water, is the interference produced by substances other than the determinand. The magnitude of the interference depends on the results of effects of all individual substances causing interferences and any other substances that may affect the effects of such interfering substances. Interference may lead to positive or negative bias and the size of the effect may depend on the concentration of the determinand as well as that of the interferent. "Interference" can be best defined as follows:

> "For a given analytical system, a substance is said to cause interference if its presence in the original sample for analysis and/or in the sample during analysis leads to systematic error in the analytical result, whatever the sign and magnitude of the error."

If the magnitude of the interference effects is to be assessed for the effects of individual or combinations of substances, two general and important points follow:

— In general, analytical methods suffering from as few interference effects as possible should be used.

— It is advantageous to use analytical methods for which the principles and mechanisms are well known, so that likely interferences and their magnitude can be predicted.

The final choice of analytical method generally depends on a number of other factors which need to be considered along with the above points. If the main aim is, however, to minimize the bias of analytical results, the importance of the above points cannot be over-emphasized.

A means of continually monitoring for the effect of interference caused by other substances is to carry out spiking recovery tests. This is achieved by adding known amounts of the determinand to the sample under examination and assessing the recovery of that addition by carrying out at least duplicate analyses. It is not unreasonable to expect to achieve 100 % ± 5 % recovery for most determinands.

A.2 Procedure

If the bias B_{jk} due to the presence of a given concentration of an interfering substance (k) for a determinand concentration of c_j is to be assessed, replicate analyses should be carried out on a solution containing only the substance and determinand at the specified concentration. The mean result R_{jk} is calculated and the estimation of bias is then given by:

$$B_{jk} = R_{jk} - c_j$$

If the calibration parameters for the method do not vary appreciably over a number of batches of analyses the above approach can be used. If these conditions do not apply a second approach should be adopted where n portions of a second solution containing only the determinand at a concentration c_j are also analysed at the same time and in the same way as the first solution mentioned above. If the mean analytical results for the solutions with and without the interfering substance are denoted by R_{jk} and R_j, respectively, the estimate of bias B_{jk} is then given by:

$$B_{jk} = R_{jk} - R_j$$

The advantage of using the second case is that the bias, B_{jk}, has been assessed using two results, the difference between which should only be attributable to the interfering substance. Using this approach there is not such a need for an accurate calibration and it is therefore the most appropriate means of testing bias caused by interfering substances.

To assess an interference effect and, as an analyst, be confident that an effect has been detected, it is important to consider the number of analyses that are required to make it both practical and valid. The main considerations are:

a) Since the magnitude of the interference effect may depend on the concentration of the determinand the effect of any substance chosen for the test should be estimated for at least two determinand concentrations. It is suggested that the lower and upper limits of the concentration range of interest are studied if only two concentrations are tested. If there appears to be large discrepancies between these results then additional intermediate concentrations should also be tested.

b) If a substance is present in a sample then the effect of that substance should be considered as a potential interferent. In this situation, the methodology and analytical system should be reviewed to decide if the effect of these substances can be considered as negligible and to be deleted from the list of possible interfering substances. Substances with concentrations less than the determinand may also produce significant interferences and should be tested. It is impractical to test for the effect of all substances present in complicated matrices such as water and numbers of substances can and should be reduced by using literature reviews and consideration of the methodology.

c) The effect of the substances identified above should be estimated experimentally at concentrations slightly greater than the expected maximum level in samples. Substances causing appreciable interference should be tested at other concentrations.

It should be remembered that any effects produced have only been estimated at the concentrations chosen for the test for a particular sample matrix and if other sample types are analysed using that method, the interference information obtained may not be applicable and other tests should be carried out to determine if there is an effect and to what level. A second problem may be that an effect may occur at concentrations less than the concentration level tested. Knowledge of the physical and chemical mechanisms should identify substances which may cause interference at lower concentrations. Of course the major problem is that the larger the number of substances to be tested the greater the analytical effort. The final difficulty to be mentioned is that the magnitude of an effect caused by a substance may depend on the concentration of a second substance. Again, detailed knowledge of the analytical method should provide sufficient information to identify likely effects caused by interaction of those substances. It is useful to test the effects of a few combinations of at least the major components of samples.

A.3 Experimental design

Analytical results are subject to both within- and between-batch random errors. The best experimental design is to estimate interference effects from tests within one batch of analyses. This assumes that interference effects do not vary with time; if they do, other considerations apply. A typical experimental design for one batch of analyses is shown in Table A.1.

Table A.1 — Example of experimental design for one batch of interference tests

Other substance	Concentration of other substance	Concentration of determinand	
		C_o	C_j
– (j = 0)	Nil	S_{00}	S_{10}
w (j = 1)	C_w	S_{01}	S_{11}
x (j = 2)	C_x	S_{02}	S_{12}
y (j = 3)	C_v	S_{03}	S_{13}
z (j = 4)	C_z	S_{04}	S_{14}

In this table S_{jk} denotes a sample with determinand concentration c_j and the kth other substance present at a defined concentration; $k = 0$ corresponds to no other added substance.

Other batches of similar design would be analysed until all the substances at all concentrations and combinations of interest had been tested.

Another way of reducing random error is to carry out at least two analyses for each solution, S_{jk}. The more analyses carried out the greater the reduction in random error. Generally duplicate analyses are acceptable. This provides, in addition to reducing random error, an estimate of those errors from the tests themselves. It is statistically possible to estimate the number of replicate analyses to carry out or obtain statistically acceptable results, but this often makes the test impractical if large numbers of analyses are required.

A.4 Interpretation and reporting of results

The estimation of the effect of the kth substance at the jth concentration of the determinand is given by:

$$B_{jk} = R_{jk} - R_j$$

where R_{jk} is the mean analytical result for the solution containing the other substance, and R_j is the result for a solution not containing the other substance, but containing the same concentration of the determinand (which may be zero).

When B_{jk} is calculated using this equation, blank-correction is not needed. The differences, B_{jk}, are the primary experimental estimates of the interference effects and it is strongly recommended that the individual differences be reported together with their limits at a defined confidence level for each concentration of determinand (i.e. level of j). The precision of analytical results is assumed here to depend on the concentration of determinand but to be unaffected by the presence of the other substances. Table A.2 gives an example of a suitable format for presenting the results.

Table A.2 — General format for presenting results of interference tests

Substance	**Concentration of substance**	**Effect of other substance** (for example as µg/l of determinand) **at a determinand concentration of:**	
		c_0	c_i
w	c_w	B_{01}	B_{11}
x	c_x	B_{02}	B_{12}
y	c_y	B_{03}	B_{13}
z	c_z	B_{04}	B_{14}

The results in the table have 100 (1–α) % confidence limits, L, equal to the result ± L_0 and result ± L_i for the determinand concentrations c_0 and c_i respectively.

The method of presenting the results shown in Table A.2 allows rapid examination to identify those substances causing statistically significant effects — i.e. those for which $B_{0k} \pm L_o$ and/or $B_{1k} \pm L_i$ are greater than zero. The biases observed for such substances are directly recorded, may easily be converted to relative effects if desired, and may also be easily assessed at other confidence levels if required. Further, the table shows the results for any substances whose apparent effects have not achieved statistical significance.

The calculation of the confidence limits L_0 and L_i is performed as follows (assuming that precision is not affected by the other substances, that m replicate analyses are made for each sample, that n other substances are tested and that the means are normally distributed):

A.4.1 Estimate variance, s^2_{jk}, for each solution from its m results

$$s^2_{jk} = \frac{\left(\sum_k R^2_{jkl}\right) - \left(\sum_k R_{jkl}\right)^2 \Big/ m}{m-1}$$

where subscript l refers to the lth replicate result for a sample.

A.4.2 Combine all estimates of variance with a given value of j for each value of j to obtain pooled estimates of variance, s^2_j

$$s^2_j = \frac{\sum_j s^2_{jk}}{n+1}$$

When $m = 2$ the following equation can be used:

$$s^2_j = \frac{\left(\sum_j R_{ik} - R_{2k}\right)^2}{2(n+1)}$$

A.4.3 Form the confidence limits, L_j, for the differences R_{jk} for each value of j

$$L_j = t(2s^2_j/m)^{1/2}$$

where t is the tabulated value of the t-statistic at the desired confidence level (for example $t = 2{,}45$ for $\alpha = 0{,}05$ and $n = 5$, $m = 2$).

A.5 Worked examples of the use of interference testing

A.5.1 *For a constant calibration*

A.5.1.1 *Table of results*

Other substance	**Concentration of other substance** (mg/l)	**Concentration of arsenic** (µg/l)	
Concentrations of arsenic		0 (c_0)	20 (c_i)
Sodium fluoride	50	R_{0il}	R_{1il}
		0,15	19,42
		0,17	19,37
		0,14	19,47
Mean R_{jk}		0,153	19,42

A.5.1.2 *Calculation of interference effect*

$B_{jk} = R_{jk} - c_j$

where B_{jk} = Interference effect of kth substance at jth arsenic level

R_{jk} = Mean result for kth substance at jth arsenic level

c_j = Concentration of arsenic at jth level

For $c_0 = 0$ $B_{01} = 0.153 - 0 = 0{,}153$

For $c_1 = 20$ $B_{11} = 19{,}42 - 20 = -0{,}58$

A.5.1.3 *Calculation of confidence limits*

Estimate variance S^2_{jk} for each solution using:

$$S^2_{jk} = \left[\left(\sum_l R^2_{jkl}\right) - \left(\sum_l R_{jkl}\right)^2 / m\right] / (m-1)$$

For determinand concentration of 0 µg/l (i.e. $j = 0$):

$S^2_{01} = (0{,}071 - 0{,}070\ 53)/2 = 0{,}000\ 23$

For determinand concentration of 20 µg/l (i.e. $j = 1$):

$S^2_{11} = (1\ 131{,}41 - 1\ 131{,}409)/2 = 0{,}002\ 5$

These values of S^2_{jk} can be easily calculated using a scientific calculator.

A.5.1.4 *Reporting of results*

These results should be reported in tabular form:

Other substances	**Concentration of other substances** (mg/l)	**Effect**[a] **of other substances,** in µg As/l, at an arsenic concentration of:	
		0 µg/l	20 µg/l
Sodium fluoride	50	0,153	– 0,58
(etc. for other potential interferents tested)			

[a] If the other substances had no effect, results would be expected to lie (95 % confidence limits) within the following ranges:

0,000 ± 0,02 for 0,000 µg As/l

20,00 ± 0,08 for 20,00 µg As/l

A.5.2 ***For an unstable calibration***

A.5.2.1 *Table of results*

Other substance	**Concentration of other substance** (mg/l)	**Concentration of arsenic** (µg/l)
Unspiked arsenic standard (Mean result R_j)		0 (R_0)20,05 (R_1)
Sodium fluoride	50	$R_{011}R_{111}$
		0,1419,44
		0,16 19,42
		0,15 19,49
Mean R_{ji}		0,15 19,45
Sodium selenite	0,1	$R_{021}R_{121}$
		– 0,0820,18
		– 0,0620,23
		– 0,10 20,19
Mean R_{jk}		– 0,0820,20

A.5.2.2 *Calculation of interference effects*

$B_{jk} = R_{jk} - R_j$

For sodium fluoride (k = 1):

R_{01} = – 0,08, therefore B_{01} = 0,15 – 0,00 = 0,15

R_{11} = 19,45, therefore B_{11} = 19,45 – 20,05 = – 0,60

For sodium selenite (k = 2):

R_{02} = – 0,08, therefore B_{02} = – 0,08 – 0,00 = – 0,08

R_{12} = 20,19, therefore B_{12} = 20,20 – 20,05 = 0,15

A.5.2.3 *Calculation of confidence limits*

Estimate variance S^2_{jk} for each solution using:

$$s^2_{jk} = \frac{\left(\sum_l R^2_{jkl}\right) - \left(\sum R_{jkl}\right)^2 / m}{m-1}$$

For sodium fluoride (k = 1) and for an arsenic concentration of 0 µg As/l (j = 0):

S^2_{01} = (0,067 7 – 0,067 5)/2 = 0.0001

For sodium fluoride (k = 1) and for an arsenic concentration of 20 µg As/l (j = 1):

S^2_{11} = (1 134,910 1 – 1 134,907 5)/2 = 0.0013

Similarly, for sodium selenite (k = 2) and for an arsenic concentration of 0 µg As/l (j = 0):

S^2_{02}= 0,000 4

and for sodium selenite (k = 2) and for an arsenic concentration of 20 µg As/l (j = 1):

S^2_{12} = 0,000 7

The pooled estimate of variance is given by:

$$S^2_j = \frac{\sum_j S^2_j k}{n+1}$$

Thus, for a determinand concentration of 0 µg As/l ($j = 0$):

$S^2_0 = 0{,}000\ 5/3 = 0{,}000\ 166\ 7$

and for a determinand concentration of 20 µg As/l ($j = 1$):

$S^2_i = 0{,}002\ 0/3 = 0{,}000\ 666\ 7$

The confidence limit L_i is given by:

$$L_j = t\left(2S_j^2 / m\right)^{\frac{1}{2}}$$

where $t = 2{,}78$ for $\alpha = 0{,}05$ and $n = 2$, $m = 3$

Thus, for a determinand concentration of 0 µg As/l ($j = 0$):

$L_0 = 2{,}78 \times (2 \times 0{,}000\ 166\ 7/3)^{½} = 0{,}029\ 3$

and, for a determinand concentration of 20 µg As/l ($j = 1$):

$L_1 = 2{,}78 \times (2 \times 0{,}000666\ 7/3)^{½} = 0{,}058\ 6$

A.5.2.4 *Reporting of results*

These results should be reported in tabular form:

Other substances	**Concentration of other substances** (mg/l)	**Effect**[a] **of other substances,** in µg As/l, at an arsenic concentration of:
		0 µg/120 µg/l
Sodium fluoride	50	0,15–0,60
Sodium selenite	0,1	– 0,080,15
(etc. for other potential interferents tested)		

[a] If the other substances had no effect, results would be expected to lie (95 % confidence limits) within the following ranges:

0,000 ± 0,0293 for 0,000 µg As/l

20,00 ± 0,0586 for 20,00 µg As/l

Annex B (informative)
Recovery of all forms of the determinand

Within a sample, the determinand may occur in many physical and/or chemical forms. It is therefore essential that the analytical method defines the determinand unambiguously to ensure that a laboratory does not choose an inappropriate and/or noncomparable method.

It is essential that the analytical response for a given concentration of the determinand is independent of the form of the determinand in the sample. Failure to ensure this may result in a significant bias in the method. Consider the analysis of a sample for arsenic using hydride generation. Arsenic can be present in aerobic conditions as As(V), in anaerobic conditions as As(III), or as organoarsenic compounds. Normal preservation using H_2SO_4 will not convert organic arsenic compounds into the inorganic form. They have to be oxidized using an acid/persulfate digestion stage which converts them into the inorganic form. As(III) can be quantitatively converted to arsine over a wide range of conditions. However, conditions can be chosen under which quantitative generation from As(V) occurs, but it does so at a slower rate. The usual solution is to pre-reduce the As(V) to As(III), commonly using sodium or potassium iodide. Consequently, negatively biased results would be obtained if samples contained organic or As(V) forms of arsenic.

Many common but important determinands in water analysis can occur in different physical and/or chemical forms, for example total metals, where in addition to dissolved or particulate forms they may also be present in different oxidation states, and organometallic compounds. To differentiate between the dissolved and particulate forms, the sample has to be filtered in such a way that the technique is reproducible and precise, and it is therefore essential that precise details of the filtration process is defined unambiguously. Generally, all measurement techniques employed in analytical systems for waters show some degree of dependence on the form in which the determinand is present in the sample.

It is very important that the information available on the performance of an analytical method, and also the form in which the determinand is present in a sample, is carefully considered when assessing whether the degree of bias is important for a particular application. To control this type of bias the following points should be considered:

— Preference should be given to the choice and use of methods giving sufficient evidence on the magnitude of the problem.

— The degree of sample pretreatment should be sufficiently vigorous to convert all forms of the determinand to the form in which it is most easily determined with little bias. It is possible, because of increased sample pretreatment and manipulation, that increased risk of contamination is likely which may result in a decrease in precision of the analytical results. Therefore, a compromise of bias and precision may be necessary.

— Tests may be performed to determine the suitability of an analytical method from the point of view described in the previous paragraph. These tests are often slow and difficult to obtain unambiguous results. Two main approaches to such tests can be identified:

a) Known amounts of pure standard substances corresponding to those likely to be present in the sample are analysed according to the analytical method to assess the efficiency with which they are determined.

In some situations it is not possible to obtain particular substances in the form in which they would be present in a sample, for example particulate forms of metal. It should be remembered that the results of such tests are subject to random errors so that, if these errors are large enough, bias may not be detected even if it is present.

b) Sample pretreatment is often necessary. To test this, treatments of increasing severity can be repeated on a number of samples.

If the results indicate that the increased severity of pretreatment does not increase the result to a significant degree, it can be reasonably assumed that the pretreatment specified in the method is adequate for the forms of the determinand present in that particular sample type.

In practice, time and resources generally prohibit most laboratories from carrying out the tests suggested in a) and b) above. Therefore it is imperative that analysts developing methods and investigating analytical procedures should pay great attention to this problem and any evidence of this type of bias should be published.

It can be seen from the above considerations that it is generally very difficult to estimate the bias due to this potential source of error when it may be present. It is therefore strongly recommended that methods should be chosen for which this bias is essentially zero. If this is not possible there is little alternative but to have discussions between the analyst and the user of the results and either accept the inherent bias or to change the determinand in order to have results of better known accuracy.

A final point to be made is the relationship between the ability of an analytical system to determine all forms of the determinand and the same system's liability to interference effects. Since the chemical and/or physical forms of a determinand may be affected by other substances in a sample, such effects may lead to biased results. For example, aluminium may complex fluoride, which under certain circumstances does not produce a response with a fluoride ion selective electrode. If the determinand quoted is "total fluoride", i.e. the total concentration of fluoride, regardless of the form in the sample, biased results may be quoted because of the presence of aluminium. The question arises whether this bias is caused by the inability to determine all forms of the determinand or by the interference caused by aluminium. In this case, the latter reason is most likely. Accordingly, the concept of the ability to measure all forms of the determinand should be restricted to include only those effects that are not included in the accepted definition of interference. The separation of these effects is essential otherwise there may be the danger that the same effect may be included twice when attempting to assess total bias in an analytical system by the summation of individual effects.

Annex C (informative)
Calibration and blank correction

C.1 General

Most systems require, either explicitly or implicitly, a calibration against some form of standard, usually a standard solution, so that the primary analytical responses can be converted into determinand concentrations. The calibration routinely applied to an analytical system nearly always involves one or more assumptions about the behaviour of the system. Such assumptions are usually made in order to simplify routine operation, and are based either on knowledge gained during the initial development of the system or on a theoretical consideration of the system. If any of these assumptions are not justified, the analytical results will be subject to errors (either systematic or random, or both).

A feature of accurate calibration worth special mention is the use of an appropriate analytical blank. If the procedure applied to blanks is not identical to that used for samples, the blank correction may introduce a bias.

Again, if the water used to prepare blanks contains the determinand, blank correction will cause results to be biased low. It is necessary therefore, if this source of bias is to be eliminated, to check the determinand content of the blank water and, ideally, to ensure that this concentration is negligible in comparison with the concentration in samples.

C.2 Approaches to calibration

Four frequently-used approaches to calibration are considered here:

a) fixed calibration;

b) within-batch calibration;

c) standard additions calibration; and

d) internal standardization.

(It should, however, be noted that variants of these — for example the use of matrix modification or matrix matching — are also used occasionally in practice, according to the nature and needs of the analytical technique.)

The choice between these different methods of calibration should be made on the basis of their influence on the performance of the analytical system. This in turn depends on the nature of the dominant forms of error.

C.2.1 *Fixed calibration versus within-batch calibration*

Fixed calibration involves the use of the same calibration curve (or line) over many batches of analysis. Within-batch calibration means the preparation of a fresh calibration curve for each batch of analysis.

The total variance of analytical results, s_t^2, is given by:

$$s_t^2 = s_b^2 + s_w^2$$

where s_b^2, is the between-batch variance of results and s_w^2 is the within-batch variance (both corrected for the blank).

The above equation reflects the fact that s_t, is affected by both within- and between-batch sources of random error. Between-batch errors are likely to be caused by changes in the slope of the true calibration curve from batch to batch. It is the relative importance of within- and between-batch errors which determines whether or not a method is best calibrated using a fixed or within-batch approach. Figure C.1 a) shows a method with good within-batch precision but large variations in the calibration from batch to batch. For all but the lowest concentrations, σ_b is the dominant source of random error. In order to improve precision it is necessary to reduce σ_b. This can only be done by preparing a separate calibration for each batch of analysis — the extra effort of performing a within-batch calibration may well be justified. Figure C.1 b) shows the reverse situation, where σ_w is the dominant source of random error for all concentrations. Here, there is little to be gained from further control of between- batch error; sources of error within each batch of analysis should be sought and controlled if precision is to be improved. In this case, it may be possible to take advantage of the saving of time and effort associated with fixed calibration. It is worth noting that modern instrumental analytical methods nearly all fall into the category of Figure C.1 a) — hence the current predominance of within-batch calibration. Further discussion of this subject may be found in [18].

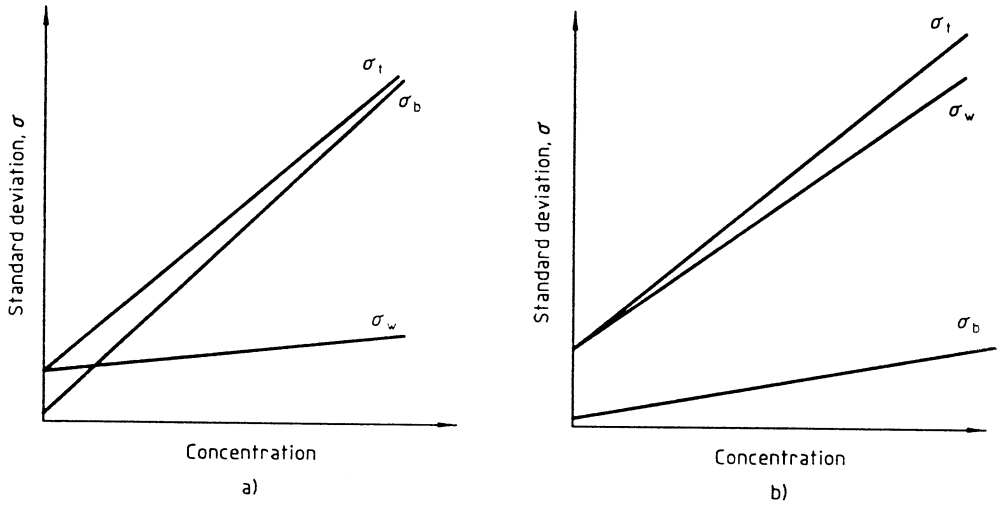

Figure C.1 — Relative importance of within- and between-batch errors

The effect on total standard deviation of calibrations with different numbers of calibration standards is shown in Figure C.2. For a technique with a relative standard deviation of individual responses of 5 %, using two standards (blank and top of range) for calibration contributes to an increase in the total standard deviation of actual individual analytical results to around 7 %. A further four calibration points has the effect of reducing the total relative standard deviation of individual results to 6 %. This is for standards equally spaced up the calibration range. A more efficient design, from the precision standpoint, is to replicate measurements at the top of the range. That gives a relative standard deviation of 5,5 %, but at the cost of not confirming the linearity of calibration.

The conclusion to be drawn from this is that there is little to be gained in precision terms, relative to the extra effort involved, by using more than, say, four calibration points. If improvement of precision is still needed, this is better achieved by replication of the measurements made on samples.

C.2.2 *Standard additions calibrations*

Standard additions calibration is a special case of within-batch calibration in which each sample is separately calibrated. It is used to overcome the effect of sample matrix on the slope of the calibration curve (a form of matrix interference). The concept of a batch of analyses (in the sense of a set of samples analysed under the same calibration) is, therefore, strictly inapplicable to the case of standard additions. However, there are sufficient similarities between determinations made on one occasion (even if each calibration is separate, the instrumental conditions will be the same) to make the idea of a batch still useful. The details of quality control may need small modifications if standard additions is used but the overall approach is unchanged.

The precision of analytical results obtained using standard additions calibration varies in a complex way according to the number of additions made and the relative concentrations of the unspiked and spiked samples. Further details are given in [14] and [15], and by [18]. The consequences of using standard additions are summarized below.

a) Standard additions tend to lead to a degradation of precision compared with direct measurement using within-batch calibration. This degradation is worst when the concentration of added determinand is small compared with that already in the sample (say only once or twice the sample concentration). At best, when the addition is large in relation to the sample concentration, an increase in the standard deviation of results by a factor of 1,4 to 2,0 should be expected when standard additions replaces direct calibration.

b) The need to spike to a concentration well in excess of that originally in the sample means that the use of standard additions tends to restrict the range of concentrations of samples which may be analysed with acceptable precision. Even when the optimum additions of determinand are made, the usage range may be reduced to as little as one third of that achievable by direct methods of calibration. This is a serious limitation if the analytical technique is inherently of narrow range.

c) Special attention needs to be paid to blank correction. Standard additions will not correct for blank errors so that efforts need to be made to determine the true sample blank and make the necessary correction to the observed responses. See also [7] for a discussion of this subject.

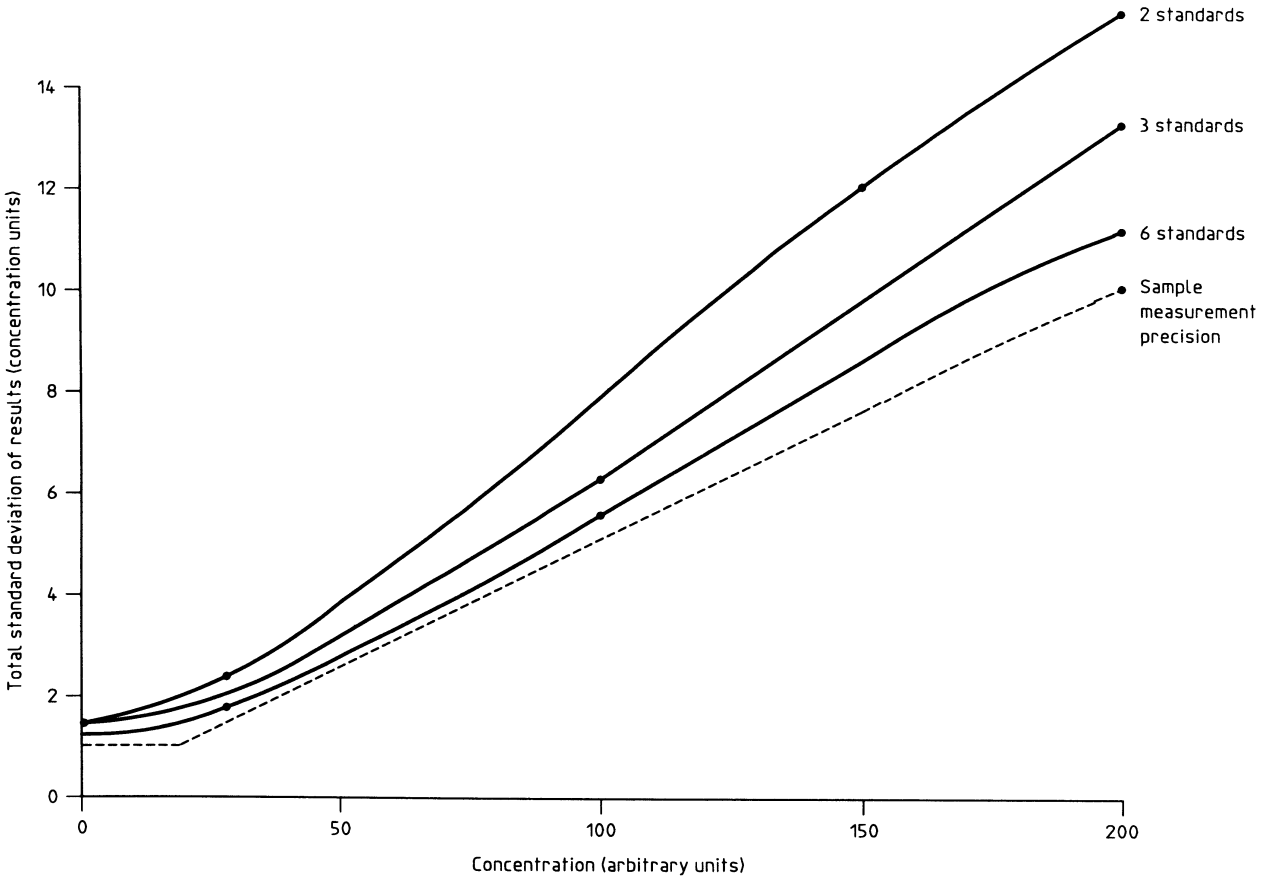

Figure C.2 — Effect of number of calibration standards on total standard deviation of results

C.2.3 *Internal standardization*

The use of internal standards is becoming increasingly popular. The approach involves adding to the sample a known amount of a substance or material. The sample is then analysed and the responses for the determinand and the added (internal) standard are measured. The observation for the internal standard is then used to relate determinand signal to determinand concentration. The effect of this on analytical error and the type of likely error will vary according to the exact approach adopted. Two common examples are discussed below.

C.2.3.1 Use of unusual elements as internal standards for multielement techniques, such as optical emission spectroscopy

Usually there is an initial, conventional calibration which relates the responses for all elements to their concentrations. Each subsequent analysis then relies on the internal standard as a means of adjusting for changes in instrumental sensitivity, perhaps caused by changes in sample uptake or by drift in detector response.

Here, care needs to be taken to eliminate factors (such as the efficiency of excitation) which affect the standard and one or more of the determinands to different extents — since these will lead to systematic error. Unless the size of response for the initial standard is the same as that for all the elements of interest (which is most unlikely), nonlinearity of response can also lead to error. This may well go undetected, since it is rare to make a range of internal standard additions.

The consequences of this approach to calibration will be to increase random error by the random variation associated with the internal standardization. Overall precision may still be better than otherwise, however, since the consequent control over drift etc. may improve the observed total standard deviation.

C.2.3.2 *Use of isotopic standards*

This is particularly useful for analysis of organic substances by mass spectroscopy, where the determinand can be added to the sample in a form that does not occur naturally in significant amounts — for example, with deuterium replacing a hydrogen atom or ^{13}C replacing ^{12}C. The response for a known addition of isotopic standard can be measured and compared with that of the determinand to give results in concentration terms. This method of calibration relies on the assumption that the added standard behaves in exactly the same way as the determinand (as in conventional standard additions).

This assumption may be incorrect if there are different chemical species present in the sample and if there is no rapid equilibrium (so that the added standard does not achieve the same speciation as the determinand). If the different species respond or are recovered to different extents, a systematic calibration error may be introduced. The question of linearity of response, mentioned above, should also be considered here.

Again, the internal calibration will introduce a degree of random error comparable with that associated with the determinand measurement itself.

In both of the above examples, the method of internal standardization shows characteristics similar to that of standard additions:

a) Only certain types of proportional error are calibrated-out.

b) Nonproportional errors are not properly corrected.

c) Linearity of response is important.

d) The behaviour of the added determinand should be the same as that of the determinand.

In summary, because all approaches to calibration have some potential drawbacks, all need to be used with care. The desirability of making an independent check on calibration error using a sample of known concentration, taken through the whole analytical procedure, cannot be overemphasized.

Annex D (informative)
Sample stability

D.1 General

A number of physical, chemical and biological processes may take place in a sample in the time between sampling and analysis. These may have the effect of changing the determinand concentration. Whilst it is often possible to appreciate that sources of sample instability might exist, it is usually difficult to predict the rate at which changes take place and hence the effect upon the concentration of the determinand of interest. For this reason it is important that laboratories check their sample handling and storage procedures to ensure that any changes which occur are not of important size.

The aim of this annex is to identify the key features which need to be considered before tests of sample stability can be undertaken. The following summarizes the recommended approach.

D.2 General approach to the design of tests

In essence, any design will involve the analysis of portions of the sample at two or more different times after sampling. For a given number of analyses there are two main options:

a) to make few replicate determinations on each of few occasions;

b) to make many replicate determinations on each of few occasions.

It is thought that a likely relationship between the determinand concentration and the time after sampling would be very difficult to predict. Accordingly, the use of regression analysis on the results of analyses made at frequent intervals after sampling is not likely to be fruitful. It follows that any design of test should be such that changes in concentration of the magnitude of interest can be detected at each time at which the sample is analysed.

Determinations on each analytical occasion are therefore replicated in order to increase the precision with which the determinand concentration is known, such that it is possible to detect (as statistically significant at a chosen probability level) the smallest degree of instability of interest. In other words the test is designed to have adequate power.

In view of the complexity of processes leading to sample instability, it is probable that samples of different type will be subject to differing degrees of instability. It is therefore highly desirable, in designing tests of sample stability to seek to test a number of samples which cover the range of composition and determinand concentrations likely to be encountered routinely.

D.3 Control of random analytical error in stability tests

The number of replicate determinations necessary to provide adequate power will increase as the standard deviation of an industrial, blank-corrected, analytical result increases. That standard deviation, σ_1, is given by:

$$\sigma_t^2 = \sigma_w^2 + \sigma_b^2$$

where σ_w and σ_b are the within-batch and between-batch standard deviations of blank-corrected results, respectively. The control of various sources of random analytical error, both between batches of results and within each analytical batch is therefore important. The impact of lack of control over random errors on the results of a stability test depends on the nature of the errors in question. Large within- batch random errors will cause a relatively wide spread of results obtained on each analytical occasion. This means that it will not be possible to detect (as statistically significant) small changes in sample concentration from one occasion to another. In other words, the power of the test is adversely affected by increasing within-batch random error. Between-batch random error will be manifest as changes in mean concentration from one analytical occasion to the next. Contribution to differences in observed concentration on different occasions which arise from between-batch analytical sources will thus be indistinguishable from true sample instability.

Replication of determinations in each batch of analyses is one way in which the effect of random error can be reduced. The standard deviation, σ_c, of the mean of n replicate analytical determinations is given by:

$$\sigma_c^2 = (\sigma_w^2/n) + \sigma_b^2$$

Replication on any analytical occasion therefore allows the reduction of variability from within-batch sources. The power of the test is correspondingly increased, but the possibility that between-batch analytical errors might be interpreted as sample instability remains.

D.3.1 ***Control of between-batch analytical random error***

Two approaches to the problem of between-batch error have been used. However, since it is not readily possible to assess the extent to which between-batch variability contributes to the estimate of instability, the choice of either approach is a matter of judgement and relates to the determinand in question, its concentration and the type and performance of the analytical method used.

a) The simple stratagem is to assume that between-batch effects are negligible with respect to the instability of interest. This approach may be adopted on the grounds that between-batch variability is small in relation to the allowable instability. The number, n, of replicate determinations required on each analytical occasion is then calculated from:

$$n = 13/D^2 \qquad \text{(D.1)}$$

where

- D is equal to d/σ;
- d is the smallest change in determinand concentration which is of interest;
- σ_r is the standard error of difference between mean results obtained on different analytical occasions.

If it is assumed that the standard deviation of blank-corrected measurements s_w made on two different occasions is the same (this assumption can be statistically tested), then

$$\sigma_r = (\sigma_w \sqrt{2})/\sqrt{n} \qquad \text{(D.2)}$$

An estimate, s_w, may be used in place of σ_w provided the former has at least 10 degrees of freedom.

The consequence of this approach to the problem of between-batch random error, if such errors are important, is that the results of the stability test may be misleading. The sample under test (in combination with its storage conditions) might be falsely judged to be unstable. Unnecessary (and potentially expensive) changes in sampling and storage procedures might then be made. Alternatively, sample instability of an important magnitude may be missed and the necessary remedial action not taken.

An additional disadvantage of this simple procedure for the design of tests is that any drift of analytical response which occurs within each batch — this may be drift of calibration or of baseline (blank), or a combination of both — will inflate the estimate of within-batch random error, thereby reducing the power of the test.

If the above disadvantages are considered too great, the following alternative approach may be used.

b) The following design of test is of value where there is a possibility of calibration drift within each batch. Such drift is likely to be a reflection of important between-batch variability and hence serves as an indicator that the approach described in a) above is inadequate.

On each analytical occasion, n replicate analyses of a standard solution of concentration closely similar to that of the sample are made in addition to n replicate analyses of the sample and of the blank. This gives estimates c'_{si} and c'_{ci} where c_s and c_c are concentrations (corrected for blank) of the sample and standard, respectively, and the suffix i denotes replicate determinations within a batch. For each analytical occasion, n values of a blank and calibration-corrected sample concentration, c_{si}, are calculated from:

$$c_{si} = \frac{\mu_c \bullet c'_{si}}{c'_{ci}}$$

where μ_c is the true concentration of the standard solution. The concentration of the sample, c_s is taken as:

$$\Sigma\, c_{si} = n\, c_s$$

The number of replicate determinations, n, is calculated using formula (D.2). In this case, σ_w, the standard deviation of blank-corrected results refers to results which have been subjected to the blank and calibration correction described. In other words, σ_w is the standard deviation of results c_{si}. Usually, a value for this standard deviation will not be available in advance of the test and an estimate should be employed to calculate n. Such an estimate may be gained by multiplying the standard deviation of independently blank-corrected results by a factor of $\sqrt{2}$.

The "correction" of the results of determinations on the sample is intended to provide a high degree of consistency, both within each batch and from one batch to another, thus eliminating the effect of calibration and baseline drift. The adverse effect of drift on within-batch precision is also minimized. This can (see below) result in an improvement of the power of the test.

The order of analysis on each occasion should be arranged in such a way that each sample, standard and blank determination should be subject to the same error. Given that some drift in the analytical response to a given concentration may occur in a batch of analyses, the order generally recommended is B, S, C, B, S, C, B, S, C ... etc. where B = blank, S = sample and C = standard. If the analytical system is such that the result of one determination may appreciably be affected by the preceding one (i.e. carry-over effects are present), the order suggested is B_1, B_2 .... $B_{n/2}$, S, C, S, C .... $B_{(n/2+1)}$,.... B_n. In dispensing aliquots from the sample container, all possible precautions should be taken to ensure that each aliquot is identical.

This approach is not without its potential disadvantages, however. Quite apart from the extra effort expended in analysing a separate calibration standard and blank for each replicate sample determination, the correction can itself lead to a degradation in within-batch precision with corresponding reduction in the power of the test. In the complete absence of calibration drift, the effect of the correction procedure described above will be to increase within-batch standard deviation by a factor of approximately $\sqrt{2}$. In order to compensate for this a doubling in the number of replicate determinations required on each analytical occasion would be necessary. Such an increase is often impracticable on the grounds that it is not possible to collect and divide a suitable bulk sample into identical sub-samples for testing. The value of this correction procedure in providing improved power of a stability test will depend on the importance of drift in the analytical method employed. The greater the drift the more likely the benefit gained from its reduction will outweigh the increase in within-batch error caused by the application of an independent correction to individual results in the batch.

It is, therefore, inadvisable to adopt this approach unless it is considered essential to take steps to control drifting calibration and between-batch random error.

As outlined above, it is impossible to prove that the determinand concentration of a sample does not change during a period of storage. By a suitable design of stability test, however, it is possible to have the capability of detecting an important change in concentration, should it occur.

D.3.2 *Blank-correction procedure*

The question of the blank correction to be applied in a test of sample stability is worthy of consideration. There is a risk that random variation of the blank from batch to batch (strictly a within-batch phenomenon, but manifest as a between-batch effect) will contribute to an important extent to apparent changes in concentration which will lead to false conclusions regarding sample stability. Ideally, such an effect, which represents a systematic error in the results of the test, should be eliminated. In practice, it is only possible to take steps to reduce the risk of drawing false conclusions.

This error in the blank correction may be reduced in two ways. One is to perform blank-correction using the mean of several blank determinations made in each batch. The alternative is individually to blank-correct each of the n replicate determinations with a separate blank measurement.

The former procedure will not increase the invariability of results on each occasion. Hence the power of the test will be unimpaired. The number of blank measurements required cannot easily be estimated since it depends on the relative sizes of the standard deviations of blank and sample measurements and on the relative direction of the apparent instability (caused by blank correction) and any true instability which may be present. It is suggested, as a general rule, that n blanks be determined (where n is the number of replicate determinations made on the sample) on any analytical occasion. Fewer blanks may be analysed if it is known that the standard deviation at the sample concentration is considerably larger than that of blank determinations.

The approach of individually correcting each result will cause an increase in within-batch variability but has the benefit of correcting any baseline drift which occurs in each batch. It follows, in this case, that n blank determinations will be performed, whatever the relative magnitude of sample and blank variability.

D.4 Number of occasions on which analysis is performed

The minimum number of analytical occasions, m, is two, i.e. immediately after sampling and at some subsequent time. The choice of $m = 2$ leads to difficulty in the selection of suitable time for the second occasion, t_1 hours after sampling Suppose t_1 is set at a time between sampling and analysis. Then, if an important instability is found, its magnitude after more usual delays will not be known. In addition, if no effect is found, the possibility remains that counterbalancing effects are present, and that the net effect after a shorter delay is greater than that observed. For such reasons, it is better to arrange for m to be 3 or greater, and, in view of the effort involved, $m = 3$ is a reasonable basis from which to begin tests.

If the times after sampling are denoted by t_0, t_1 and t_2, t_0 should ideally be equal to zero — especially since it is likely that some forms of instability will have their greatest effect in the period immediately after sampling. However, some relaxation of this stipulation has to be made. It is suggested that the analysis at t_0, which effectively sets the reference point against which any instability is assessed, should not be performed later than one hour after the sample is taken.

As a protection against failing to detect gradually developing effect, t_2 should be equal to, or preferably a little greater than the longest time between sampling and analysis likely to occur in routine operation. No clear principles governing the choice of t_1 are apparent, but two considerations may be borne in mind. On the one hand, t_1 might be set at a time which could, if storage for as long as t_2 resulted in unacceptable instability, be specified as the maximum storage period for samples. Alternatively, the sample might be stored at room temperature for an interval, t_1, chosen to simulate the period of transport to the laboratory, followed by refrigeration until t_2 (if refrigeration is routinely employed), to simulate storage after receipt at the laboratory.

D.5 Choice of sample containers

There is also the question of the containment of the sample under test. If, for example, one large container is used initially to contain all the sample required for all replicate analyses and analytical occasions, this may, for some determinands, necessitate the use of containers different from those used routinely. In addition, other problems may arise, depending on the determinand, if the analytical method specifies that sample containers be completely filled with sample until analysis is started. On the other hand, if three containers (of the routine) type of sample are collected initially, one for each of the analytical occasions, there is a possibility that the concentrations of the determinand in each bottle will differ and invalidate the tests. On balance, the most important consideration is that the tests be made under conditions identical to those used routinely. It is, therefore, suggested that each of m normal sample containers be used to hold the initially-collected sample, and that all necessary precautions are taken to ensure that the concentrations of determinands (and any other sample constituents that may affect analytical results) in each container are essentially identical.

This may present problems for certain determinands (especially those that are present in particulate form) and preliminary tests may well be necessary to ensure that the proposed method of filling the m sample containers is capable of ensuring the required equality of concentrations.

Notwithstanding such precautions, it is not unknown to find that unexpected problems develop, and much time and effort may be wasted if this happens. It is useful, therefore, in addition to attempt to check, in the stability test itself, whether or not a "between-container" effect is present. This can be done by filling $k\,m$ sample containers initially with the sample. Then n/k aliquots[3)] are analysed from each of k containers at t_0, n/k analyses from each of another k containers at t_1 and similarly for t_2. For simplicity, k=2 seems generally adequate provided all necessary precautions have been taken to eliminate the "between-container" effect.

The procedure in the last paragraph is also necessary if a normal sample container contains insufficient sample to provide n aliquots for analysis.

D.6 Conditions of storage

Whatever the detailed conditions under which sample containers are stored, care should be taken to ensure that no appreciable contamination occurs during storage. Given the type of sample container used routinely, three physical factors other than storage time seem capable of affecting sample stability, i.e. temperature, illumination and agitation. Such factors and the time period during which they may operate will generally vary from one sampling location to another, and from time to time at any one location; these aspects are also affected by the sample collection practices of different laboratories. There is a problem, therefore, in deciding the values of these factors to be adopted in stability tests. In principle, this problem could be simplified if the effects of these factors on all determinands of interests were known.

Unfortunately, this information appears not to be generally available, and as it is impracticable to check all possible combinations of factors at once, some arbitrary, decisions on these factors appear to be necessary and the following suggestions are made.

a) After routine sampling, sample containers should be exposed to negligible illumination, for example by placing them in covered boxes. If this is done, sample containers for stability tests should be stored in the dark.

b) It appears that the time between sampling and analysis, T, can be generally divided into two periods ($T = T_1 + T_2$). The first T_1, is that period immediately after sampling during which the temperatures of samples are not controlled. The second period, T_2, commences when the sampling is placed in refrigerated surroundings and ends when analysis is started. The storage conditions and times adopted should follow this pattern.

c) The samples for stability tests that are to be stored for the longest time should then be stored at the chosen high temperature [see e) below] for a time T_1, and under refrigerated conditions for a time $T - T_1$. If $T_1 > T/3$, samples should be stored only at the chosen high temperature [see e) below].

d) It should be assumed, at least initially, that instability of samples increases with increasing temperature. On this basis, samples should be stored for a period T_1 at the highest temperature that may be experienced routinely before placing them in refrigerated surroundings.

[3)] If this approach is used, n should, if necessary, be adjusted upwards so that it is an integral multiple of k, i.e. so that equal numbers of aliquots are taken from each container.

e) The type of agitation envisaged as of possible significance is that experienced by sample containers during transport in vehicles. It seems difficult to reproduce this agitation in stability tests where sample containers are likely to need to be stored at temperatures greater than ambient. However, agitation should be used if the necessary special facilities are available, for example a mobile laboratory with facilities for storing sample containers at a chosen temperature, or a room or incubator whose temperature can be controlled and in which sample containers can be agitated. If such facilities are not available, no reasonable means of including sample agitation is apparent, and it is simply suggested that samples be stored without agitation until results from initial stability tests become available.

D.7 Choice of sample for stability tests

It seems virtually certain that sample stability depends on the determinand, its concentration, and the natures and amounts of other materials and species in the sample. There is, therefore, a need to decide those locations from which samples are collected for initial stability tests. Again, there seems to be no clear-cut principles for deciding which samples are likely to show the most important instability. Accordingly, it is suggested arbitrarily that, unless a particular laboratory has specific reasons otherwise, initial stability tests should involve at least two sample-types:

a) one sample with relatively high concentrations of the determinands of interest, and

b) one sample with relatively low concentrations (see below) of those determinands. In addition, the locations of the sampling points should be such that analysis of the sample corresponding to "zero" delay time can be started within one hour (and preferably less) of sampling. This last suggestion is based on the idea that any changes occurring in the first hour after sampling are small compared with those occurring subsequent. This idea may be incorrect, but there seems little alternative to accepting it, at least until results from initial stability test are available.

A useful convention in stability tests for many projects is to set the target for maximum possible instability at half the target for maximum possible bias. This corresponds to 5 % of the determinand concentration or one quarter of the required limit of detection (implied by the random error requirement), whichever is the larger.

It is possible to choose a sample for stability testing which has a determinand concentration equal to or less than a quarter of the required limit of detection. This situation is more likely to arise when the limit of detection of the analytical method used in the laboratory is considerably lower than that required. If such samples were used in stability tests, it would be possible for the determinand concentration to fall to zero, yet still not result in the target for stability being exceeded. Indeed, if sample instability tends to cause a decrease in concentration, it is unlikely that the laboratory will identify any problem.

The use, in stability tests, of test samples of a determinand concentration less than the target value for maximum possible instability should be avoided, unless the majority of routine samples are of a similar concentration or lower. In the normal situation where most routine samples are of concentration greater than the stability target, it is recommended that an alternative test sample should be taken if the concentration found on analysis at time $T = 0$ is less than critical value.

Annex E (informative)
Bibliography

[1] ALHA-AWWA-WPCF (1989) Standard Methods for the Examination of Water and Waste, 17th edition, APHA, Washington.

[2] Association of Official Analytical Chemists (AOAC) (1975) Statistical Manual.

[3] ASTM — American Society for Testing and Materials (1963) Conducting an Inter-laboratory Study of a Test Method, *ASTM Special Technical Publication Number*, **335**, Philadelphia.

[4] ASTM (1975) "Standard Methods for the Examination of Water and Waste Water" 14th edition, 104B, 26–33.

[5] ASTM (1976) STP 15D "Manual on Presentation of Data Control Chart Analysis".

[6] British Standards Institution (1982) Data analysis and quality control using cusum **techniques**. British Standard 5703-1 to 5703-**4**.

[7] Cardone, M.J. (1986) New technique in chemical assay calculations. Parts 1 and 2, *Analytical Chemical*, **58**, 433–445.

[8] Davies, O.L. (editor) (1956) Design and Analysis of Industrial Experiments, Oliver and Boyd, Edinburgh, 550.

[9] Davies, O.L. and Goldsmith, P.L. (editors) (1972) Statistical methods in research and production. 4th edition, revised. Edinburgh, Oliver and Boyd for ICI, pp 478.

[10] Eisenhart, C. (1963) Realistic evaluation of the precision and accuracy of instrument calibration systems, *J Res National Bureau of Standards*, **67C**, 161.

[11] Franke, J.P. and de Zeeuw, R.A. (1978) "Evaluation and Optimisation of the Standard Addition Method for Absorption Spectrometry and Anodic Stripping Voltammetry", *Analytical Chemistry*, **50**, 1374–1380.

[12] Funk, W. *et al* (1985) Statistische Methoden in der Wasseranalytik, Begrifffe, Strategien, Anwendungen, GCH, Weinheim.

[13] Funk, W., Donnevert, G., Dammann, V. (1992) "Quality Assurance in Analytical Chemistry", Verlag Chemie, Weinheim.

[14] Gardner, M.J. and Gunn, A.M. (1986) Optimising precision in standard additions determination, *Z Analyt Chem*, **325**, 263–268.

[15] Gardner, M.J. and Gunn, A.M. (1988) Approaches to calibration in GFAAS: direct or standard additions, *Z Anal Chem*, **330**, 103–106.

[16] Gardner, M.J., Wilson, A. L., and Cheesman, R. J. (1989) "A Manual on Analytical Quality Control for the Water Industry", NS30, WRc Medmenham.

[17] Hunt, D.T.E. and Morries, P. (1980) The determinand content of the blank water, *Analyt Proc,* **19**, 407–411.

[18] Hunt, D.T.E. and Wilson, A.L. (1986), *The Chemical Analysis of Water. General Principles and Techniques*, Second edition, The Royal Society of Chemistry, London.

[19] ISO 5725:1994-1 to ISO 5725:1994-4 and ISO 5725:1994-6, *Accuracy (trueness and precision) of measurement methods and results.*

[20] Mancy, K.H. (1974), Design of water quality measurement programs. In: *Design of Environmental Information Systems*, edited by R.A. Deininger. Ann Arbor Science Publishers Inc, Ann Arbor MI, USA, pp. 173–197.

[21] Ratzlaff, K.L. (1979), "Optimising Precision in Standard Addition Measurement", *Anal Chem*. **51**, pp. 232–235.

[22] Shewhart, W. (1931), *The Economic Control of Quality of Manufactured Products*. van Nostrand, New York.

[23] Western-European Laboratory Accreditation Co-operation (1992) *WELAC criteria for proficiency test in accreditation*.

[24] Youden, W.J. (1963) Ranking laboratories by round-robin tests, *Materials Research and Standards*, **3**, pp. 9–13.

[25] ISO Guide 43:—, *Proficiency testing by interlaboratory comparisons — Part 1: Development and operation of proficiency testing schemes*[4)].

[26] EN 45002:1989, *General criteria for the assessment of testing laboratories*.

[27] EN 45003:1989, *General criteria for laboratory accreditation bodies*.

[4)] To be published. (Revision of ISO Guide 43:1984).

blank

BSI — British Standards Institution

BSI is the independent national body responsible for preparing British Standards. It presents the UK view on standards in Europe and at the international level. It is incorporated by Royal Charter.

Revisions

British Standards are updated by amendment or revision. Users of British Standards should make sure that they possess the latest amendments or editions.

It is the constant aim of BSI to improve the quality of our products and services. We would be grateful if anyone finding an inaccuracy or ambiguity while using this British Standard would inform the Secretary of the technical committee responsible, the identity of which can be found on the inside front cover. Tel: 020 8996 9000. Fax: 020 8996 7400.

BSI offers members an individual updating service called PLUS which ensures that subscribers automatically receive the latest editions of standards.

Buying standards

Orders for all BSI, international and foreign standards publications should be addressed to Customer Services. Tel: 020 8996 9001. Fax: 020 8996 7001.

In response to orders for international standards, it is BSI policy to supply the BSI implementation of those that have been published as British Standards, unless otherwise requested.

Information on standards

BSI provides a wide range of information on national, European and international standards through its Library and its Technical Help to Exporters Service. Various BSI electronic information services are also available which give details on all its products and services. Contact the Information Centre. Tel: 020 8996 7111. Fax: 020 8996 7048.

Subscribing members of BSI are kept up to date with standards developments and receive substantial discounts on the purchase price of standards. For details of these and other benefits contact Membership Administration. Tel: 020 8996 7002. Fax: 020 8996 7001.

Copyright

BSI
389 Chiswick High Road
London
W4 4AL