NORDTEST NT TR 569 ed. 6.1 Eng 2026:08

NORDTEST REPORT TR 569

Internal QUALITY CONTROLL

Handbook for Chemical Laboratories

Nordtest
Slettetoften 9,
DK-2630 Taastrup

Telephone +45 21 21 44 44
nordtest@nordtest.info
www.nordtest.info

Handbook of Internal Quality Control

Edition 6.1 2026

Authors
Bertil Magnusson, Trollboken AB
Håvard Hovind, Norway
Mikael Krysell, Qualab, Sweden
Ulla Lund, Denmark
Irma Mäkinen, Finland
Eskil Sahlin, RISE Research Institutes of Sweden, Sweden

Drawings by Petter Wang, Norway and Timo Vänni, Finland

Valuable comments on the content of this edition have been provided by:
Elizabeth Prichard, United Kingdom
Dmytro Leontiev, Ukraine
Elcio Cruz de Oliveira, Brazil
Ivo Leito and Riin Rebane, Estonia
Michael Koch, Germany
Nikola Obradović, Norway
and many, many devoted analytical chemists.

Downloading address:
www.nordtest.info

Recommended citation
This document should be cited* as "B. Magnusson et al, Handbook of Internal Quality Control, Nordtest Report TR 569 (ed. 6.1) 2026. Available from www.nordtest.info."
*Subject to journal requirements

Preface

The aim of the Trollbook is to give practical guidance on internal quality control for analysts in their daily work with analytical methods. The main updates for this 6th edition are:

- only two rules recommended for the daily evaluation of the quality control;
- discussing the frequency and how to reduce the false out-of-control situations;
- more focus on target control limits, when possible, based on uncertainty requirements.
- introducing control charts with only action limit.

The first edition of *Internal Quality Control* [1] – *Handbook of Internal Quality Control in Water Laboratories* was prepared in Nordic cooperation in the 1980s. A major revision was published in 2005, best known under the name *Trollboken* [2]. Later it was translated to several other languages and has been widely used as a tool in chemical laboratories – not only in environmental laboratories.

During the years since the first edition was prepared, there have been many developments in the field of analytical quality. Above all, the requirement for accreditation of analytical laboratories has put pressure on the laboratories to document their analytical quality, and internal quality control is an important part of this documentation.

Since the accreditation standard ISO/IEC 17025 [3] was introduced, there has been an increased focus on the concept of measurement uncertainty for both chemical and microbiological methods. When a laboratory estimates measurement uncertainty for a test method knowledge of the within-laboratory reproducibility obtained from internal quality control (intermediate precision) is essential.

The task of compiling this book has been made possible by financial support from Nordic Innovation Centre/Nordtest, Swedish Environmental Protection Agency and Trollboken AB.

This edition 6.1 of the handbook TR569 can be downloaded from www.nordtest.info. *Frequently Asked Questions* (FAQ) about Nordtest handbooks can be found at www.trollboken.se under menu item *FAQ*.

Information to our readers

The Trollbook starts, after an introduction, with two chapters (Chapters 2 and 3) on general issues of analytical quality, described with specific reference to internal quality control. They are followed by an introduction to control charting (Chapter 4).

The tools of control charting are described in the following chapters: control charts (Chapter 5), control samples (Chapter 6) and control limits (Chapter 7). Chapter 8 summarises the tools in a description of how to start a quality control programme.

How internal quality control data are used is described in the following two chapters. Chapter 9 explains the interpretation of quality control data being performed after every analytical run, whereas Chapter 10 explains how the quality control programme should be reviewed periodically to investigate whether the programme is still optimal to control the quality of analyses.

Quality control data can be used for a number of purposes other than just control of the quality in every run. Chapter 11 gives examples of other uses of quality control data and the principles of control charting.

Chapter 12 gives definitions and useful equations and in Chapter 13 you can find tables for statistical tests and internal quality control.

Chapter 14 contains examples illustrating how control charts can be started as well as practical application of the control rules presented in Chapter 9 and the yearly review described in Chapter 10. In Example 8 we present a detailed review of preliminary control limits and setting new, more reliable control limits based on more data. Example 12 describes pooling of standard deviation to obtain s_r and s_{Rw} from internal control data.

Chapter 15 lists references.

Some common symbols and abbreviations used in this handbook are found below. Further explanation is given in Chapter 12.

n	Number of measurement values	CRM	Certified Reference Material
$\bar{x}$	Mean value	GC	Gas Chromatography
s	Standard deviation	ICP-OES	Inductively Coupled Plasma-Optical Emission Spectrometry
s_r	Standard deviation under repeatability conditions	ICP-MS	Inductively Coupled Plasma-Mass Spectrometry
s_{Rw}	Standard deviation under within-laboratory reproducibility conditions (intermediate precision)	LOQ	Limit of Quantification
s_R	Standard deviation under reproducibility conditions	XRF	X-Ray Fluorescence
r	Repeatability limit		
QC	Quality Control	WL	Warning Limit
CL	Central Line	AL	Action Limit

CONTENTS

1. Introduction 1
2. Measurement uncertainty and within-laboratory reproducibility 3
3. Requirement for analytical quality 9
4. Principles of quality control charting 11
5. Different types of control charts 13
6. Different control samples 15
7. Setting the control limits 17
8. Setting up a quality control programme 21
9. Daily interpretation of quality control 23
10. Long-term evaluation of quality control data 25
11. Other uses of quality control data and control charts 27
12. Terminology and equations 29
13. Tables 33
14. Examples 35
15. References 50

Blank page

1. Introduction

Internal quality control at the analytical laboratory involves a continuous, critical evaluation of the laboratory's own analytical methods and working routines. The control encompasses the analytical process starting with the sample entering the laboratory and ending with the analytical report. The most important tool in this quality control is the use of control charts. The basis is that the laboratory runs control samples together with the test samples. The control values are plotted in a control chart. In this way it is possible to demonstrate that the measurement procedure performs within given limits. If the control value is out of control, no analytical results are reported until the source of the error is identified and removed. *Figure 1* illustrates the most common type of control chart, the X-chart.

According to ISO/IEC 17025 7.7.1 [3]: The *laboratory shall have a procedure for monitoring the validity of results. The resulting data shall be recorded in such a way that trends are detectable and, where practicable, statistical techniques shall be applied to review the results. This monitoring shall be planned and reviewed…* and then referring to both internal QC and proficiency testing in… *Data from monitoring activities shall be analysed, used to control and, if applicable, improve the laboratory's activities. If the results of the analysis of data from monitoring activities are found to be outside pre-defined criteria, appropriate action shall be taken to prevent incorrect results from being reported.*

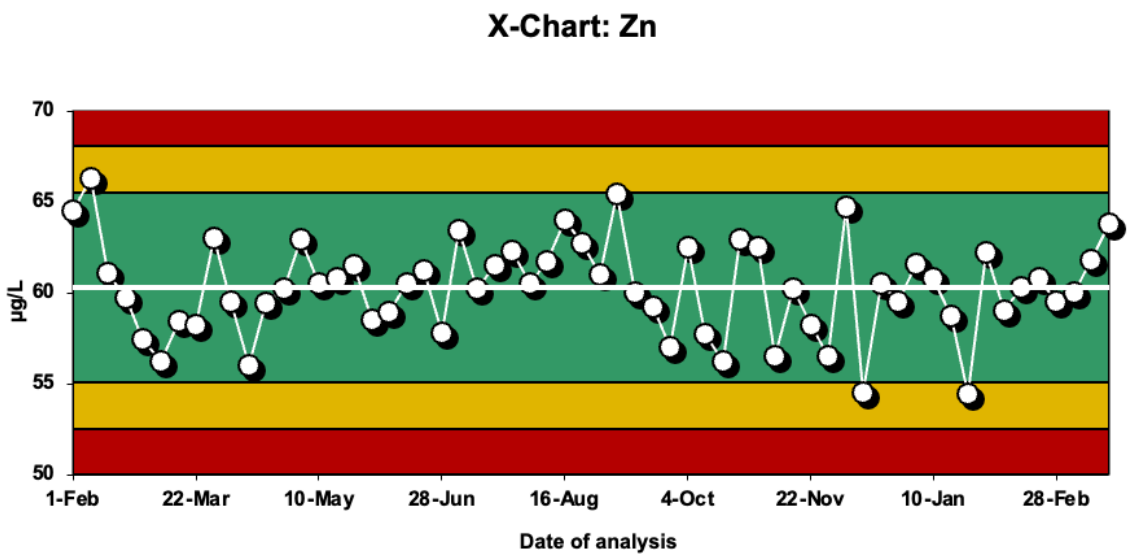


Figure 1. Example of an X control chart for the direct determination of zinc in water. All control values in the green area (within the warning limits) show that the determination of zinc performs within given limits and the test sample results are reported. Control values in the red area (outside the action limits) show clearly that there is something wrong and no test sample results are reported. A control value in the yellow area is evaluated according to a specific rule – see Chapter 9.

When a quality control (QC) programme is established, it is essential to have in mind the **requirement** on the analytical results and for what purposes the results are produced – the concept of *fitness for purpose*. From the **requirement** on the analytical results the analyst sets up the control program:

- type and number of QC samples;
- type of QC charts;
- control limits – warning and action limits;
- control frequency.

When the control sample similar to test samples encompasses the whole analytical process from the sample entering the laboratory to the analytical report the control results will demonstrate the *within-laboratory reproducibility*. The *within-laboratory reproducibility* indicates the variation in the analytical results if the same sample is given to the laboratory at different times.

The results of the control programme may be used in several ways: the analyst will have an important quality tool in his/her daily work, the customer can get an impression of the laboratory's quality, and the laboratory can use the results in the estimation of the measurement uncertainty [4].

The QC should be part of a quality management system and should be formally reviewed on a regular basis.

In practical work it is necessary that the quality control is limited to fulfilling the requirements on the analytical results – a good balance between control work and analysis of samples is essential. Wider control limits, than the statistical limits, can be set – *target control limits* when the requirements allow a higher spread in the analytical results. The aim of this handbook is to describe a *fit for purpose* system for internal quality control at analytical laboratories. The approach is general, but the examples are mainly from environmental analyses.

2. Measurement uncertainty and within-laboratory reproducibility

This chapter introduces the terminology used in quality of analyses and the statistical background for quality control.

A laboratory needs to demonstrate the quality of the analytical results. Depending on the customer's requirements it is either the spread in the results or the *measurement uncertainty* that is the important quality parameter. The internal quality control will normally give an indication of the *within-laboratory reproducibility* quantified by the standard deviation, s_{Rw}. The *within-laboratory reproducibility (intermediate precision)* will tell the customer the possible variation in the analytical results if the same sample is given to the laboratory in January, July or December. Measurement uncertainty, on the other hand, indicates for the customer the possible maximum deviation for a single result[1] from either 1) a true value, 2) a reference value, or 3) from a robust mean value of other competent laboratories analysing the same sample.

From the laboratory's point of view the possible deviation from a reference value for an analytical result may be described by the laboratory ladder of errors [5], Figure 2.[2]

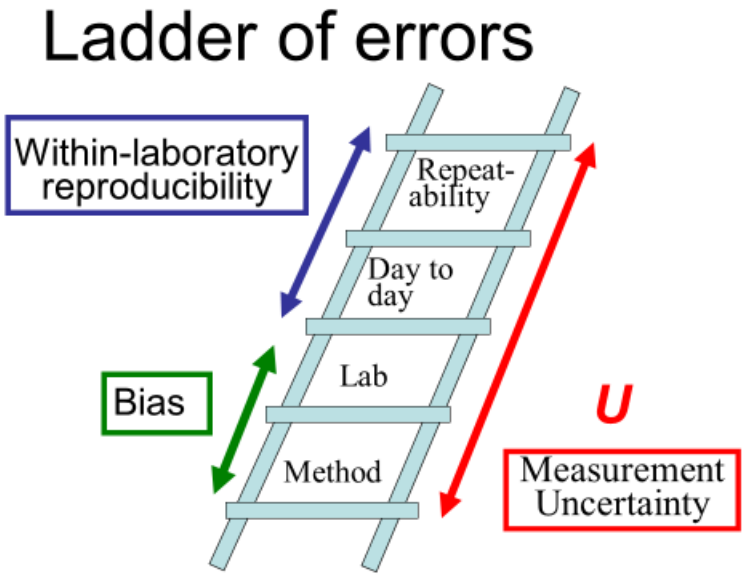


Figure 2. The ladder for a measurement procedure used in a laboratory.

Step 1 *The method bias – a systematic error owing to the method used*
Step 2 *The laboratory bias – a systematic error (for an individual laboratory)*
Step 3 *The day-to-day variation – a combination of random and systematic errors owing to, among other factors, time*
Step 4 *The repeatability – a random error occurring between replicate determinations performed within a short period of time; the sample inhomogeneity is part of the repeatability.*

[1] or more strictly *the range of possible values with a defined probability associated with a single result.*
[2] NOTE the bias here is the unknown bias that is included in the evaluation of the measurement uncertainty [4].

For an individual determination on a sample in a certain matrix the four different steps in the ladder are the following: 1) the method as such, 2) the method as it is used in the laboratory, 3) the day-to-day variation in the laboratory, and 4) the repeatability of the analysis of that sample. Each of these steps on the ladder adds its own uncertainty. The *within-laboratory reproducibility* consists of step 3 and 4 - day-to-day variation and the repeatability. Repeated inter-laboratory comparisons will show the laboratory bias, step 2, and if different methods are used also the variation in method bias, step 1. The *measurement uncertainty* reported to customer generally consists of all four steps.

Measurement uncertainty, as well as accuracy, is thus a combination of random and systematic effects. This is illustrated in Figure 3 where different requirements on measurement uncertainty are also illustrated with a small and a large green circle. For further reading about measurement uncertainty, we recommend the Nordtest TR 537 [4], ISO 11352 [6] and the Eurachem guide [7].

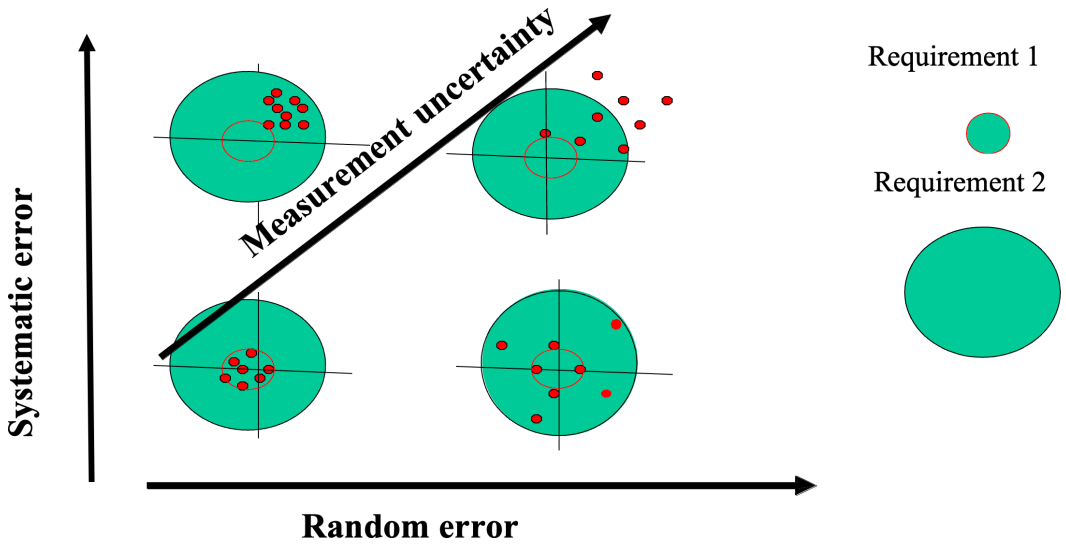


Figure 3. *Random and systematic errors of analytical results and measurement uncertainty may be illustrated by the performance of someone practising aiming at a target – the reference value or true value. Each red dot represents a reported analytical result. The two circles illustrate different requirements on analytical quality. In the lower left target requirement 1 is fulfilled and requirement 2 is fulfilled in all cases except the upper right. The upper left target represents a typical situation for most laboratories.*

Repeatability and reproducibility

We use the notion *repeatability conditions* when a sample (or identical samples) is analysed several times in a short time (e.g. the same day), by one person in one laboratory, and with the same instrument. The spread of the results under such conditions represents the smallest spread that an analyst will obtain.

We use the notion *reproducibility* conditions when a sample is analysed under varying conditions, for instance when the analyses are performed at different times, by several persons, with different instruments, different laboratories using the same analytical procedure.

The *within-laboratory reproducibility conditions* will be somewhere in between these two outermost cases.

Bias

Bias exists when an average is either higher or lower than a reference value. Variations on bias may occur over time because of changes in instrumental or laboratory conditions. For small differences it is often difficult to say if these errors are random or systematic.
Some typical sources of systematic errors are [8]:

- instability of samples between sample collection and analysis;
- loss of analyte;
- inability to determine all relevant forms of the analyte;
- interferences, e.g. a response for another substance in the matrix will cause an effect of this type;
- biased calibration;
  If samples and calibration standards are treated differently or if the matrix is different, this can represent a potentially serious source of error. Impurity of the material used to prepare calibration standards is, of course, another potential cause of systematic errors, as well as if the calibration curve is assumed to be linear in a concentration range when this is not true;
- blank correction too high or too low, if the blank and the sample are not treated in the same way.

Random variation and the normal distribution

Truly random variations from several sources *added* together can be described by a normal distribution. The unpredictable and uncontrollable variations in the many factors affecting the analytical result can for example be:

- small differences in the volume of reagents added;
- different reaction times;
- varying contamination from laboratory equipment and environment;
- instability in the instrument;
- uncertainty in the readings;
- temperature variations; and
- different calibration solutions used.

If we analyse a sample several times, we do not obtain a series of identical results. The results vary randomly, and we are not able to predict in which direction, and by how much. How may we describe the distribution of the results, and obtain a measure for the random variation? By visual evaluation of the control values in *Table 1*, we can hardly form a distinct picture of the analytical variation.

Table 1. Example of laboratory internal quality control values for a solution containing 60.0 µg/L of zinc. Figure 1 on page 1 shows these data in an X-chart.

64.5	66.3	61.1	59.7	57.4	56.2	58.4	58.2	63.0	59.5
56.0	59.4	60.2	62.9	60.5	60.8	61.5	58.5	58.9	60.5
61.2	57.8	63.4	60.2	61.5	62.3	60.5	61.7	64.0	62.7
61.0	65.4	60.0	59.2	57.0	62.5	57.7	56.2	62.9	62.5
56.5	60.2	58.2	56.5	64.7	54.5	60.5	59.5	61.6	60.8
58.7	54.4	62.2	59.0	60.3	60.8	59.5	60.0	61.8	63.8

A graphical representation of the results gives a much better picture of the spread. *Figure* 4 shows a histogram where the control values are collected into groups according to their concentration. Each group is represented by a column, the height of which is a measure of how many results this group consists of.

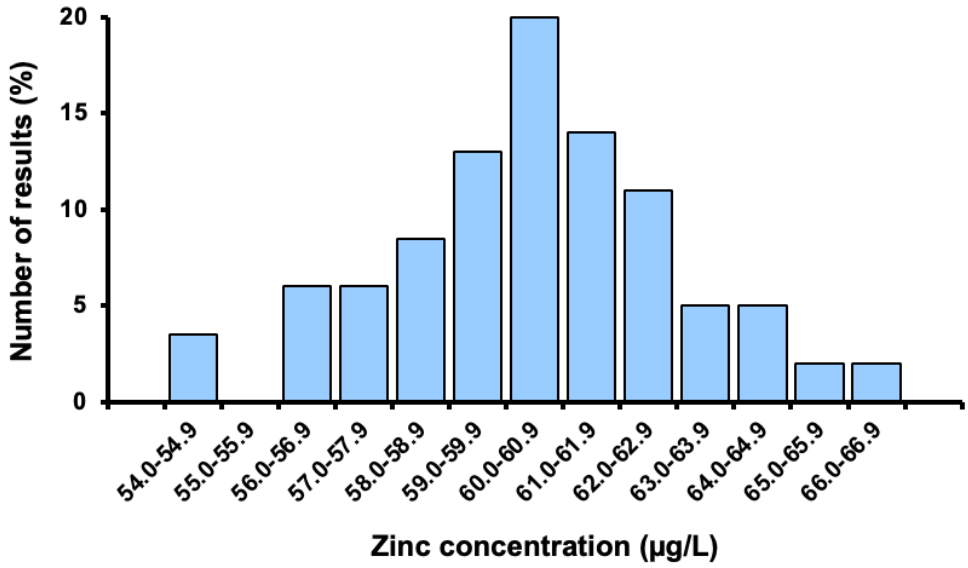


Figure 4. A histogram illustrating the distribution of the control values from Table 1. The results are sorted into groups defined by concentration ranges. Each group is represented by a column where the height represents the number of results in the group, calculated in percent of the total number of results.

If we increase the number of measurements and collect the values in groups with increasingly narrower columns, we will approach the smooth curve shown in *Figure 5*. This is an example of a frequency curve[3], the so-called normal distribution curve, constituting the basis of the control charts being used in the internal quality control.

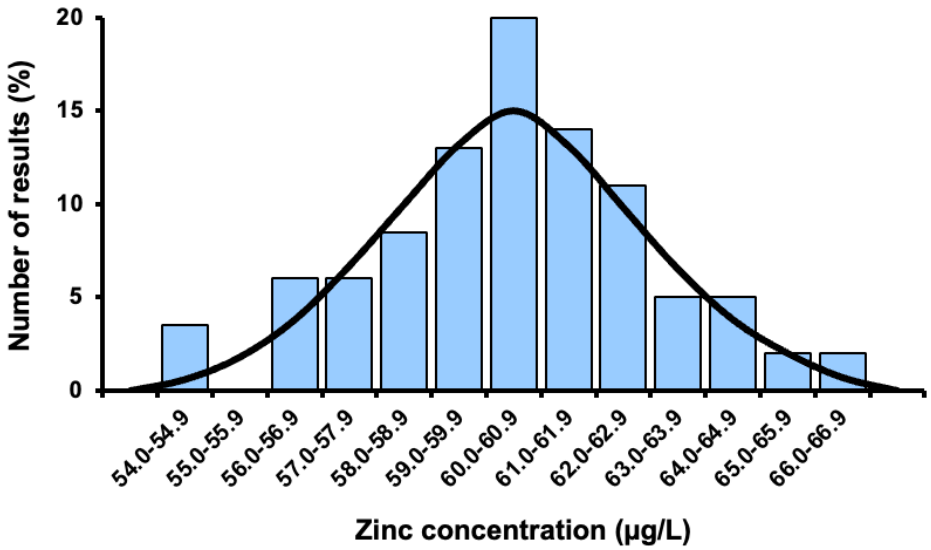


Figure 5. *The relation between the normal distribution curve and the histogram. The distribution curve is based on the same data as represented in the histogram (Figure 4).*

[3] Also called a *probability density function.*

It is a presupposition to apply the statistical methods, based on the normal distribution curve, for the treatment of the control data. However, over a longer period of time in a laboratory the bias may vary, resulting in all control values being over (or under) the mean value for a time. These control values are out of statistical control, but when they are within the warning limits test results can be reported.

When the results are normally distributed, the mean value $\bar{x}$ is defined by the position of the maximum of the curve. The shape of the curve is determined by the spread of the single results, expressed by the standard deviation, *s*. This is illustrated in *Figure 6*.

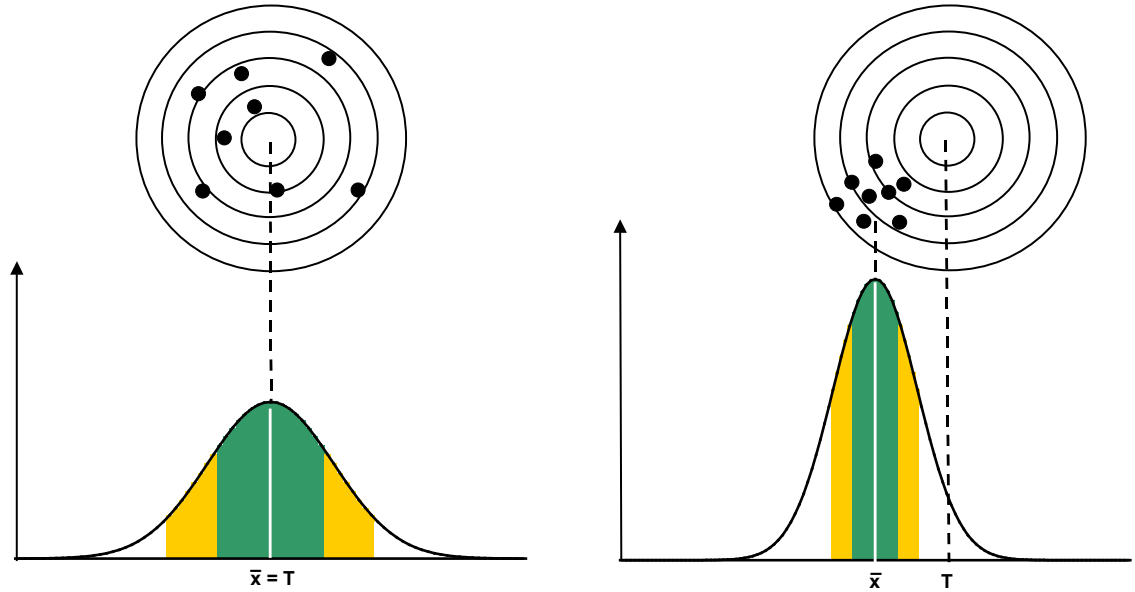


Figure 6. The shape of the normal distribution curve depends on the spread in the measurement results. A high within-laboratory reproducibility will give a large standard deviation, and the corresponding curve is broad (left). If the reproducibility is good, the standard deviation is small, and the normal distribution curve will be narrow (right). The position of the maximum demonstrates the trueness of the analysis. In the left example the mean value coincides with the true value or the reference value. In the example to the right the results are systematically too low ($\bar{x}$ is the mean value, and T is the true value or reference value). Bias is calculated as $\bar{x}$ - T or relative bias as ($\bar{x}$ - T)/T.

Based on the normal distribution we may calculate a theoretical spread of the results around the mean value, as shown in *Figure 7*. About 95 % of all results will be located within the mean value ± two times the standard deviation, and 99.7 % of the results are located within ± three times the standard deviation.[4] These properties are applied in the construction of the control charts.

When reporting within-laboratory reproducibility to a customer we will normally report it at the 95 % confidence level, that is ± two times the standard deviation. This means that on average, about 19 results out of 20 will be within this range. The 95 % confidence level is also often chosen when reporting an expanded measurement uncertainty to a customer and that will often be ± two times the combined standard uncertainty.

[4] For a normal distribution the proportion of control values outside ± 3 *s* is 0.27 % - first rule for out-of-control in Chapter 9. The proportion for the second rule, a control value between ± 2 *s* and ± 3 *s* and at least one of the two previous control values is also between ± 2 *s* and ± 3 *s* (on the same side) is estimated by probability calculus also to be 0.27 % and verified by simulation in program *R*.

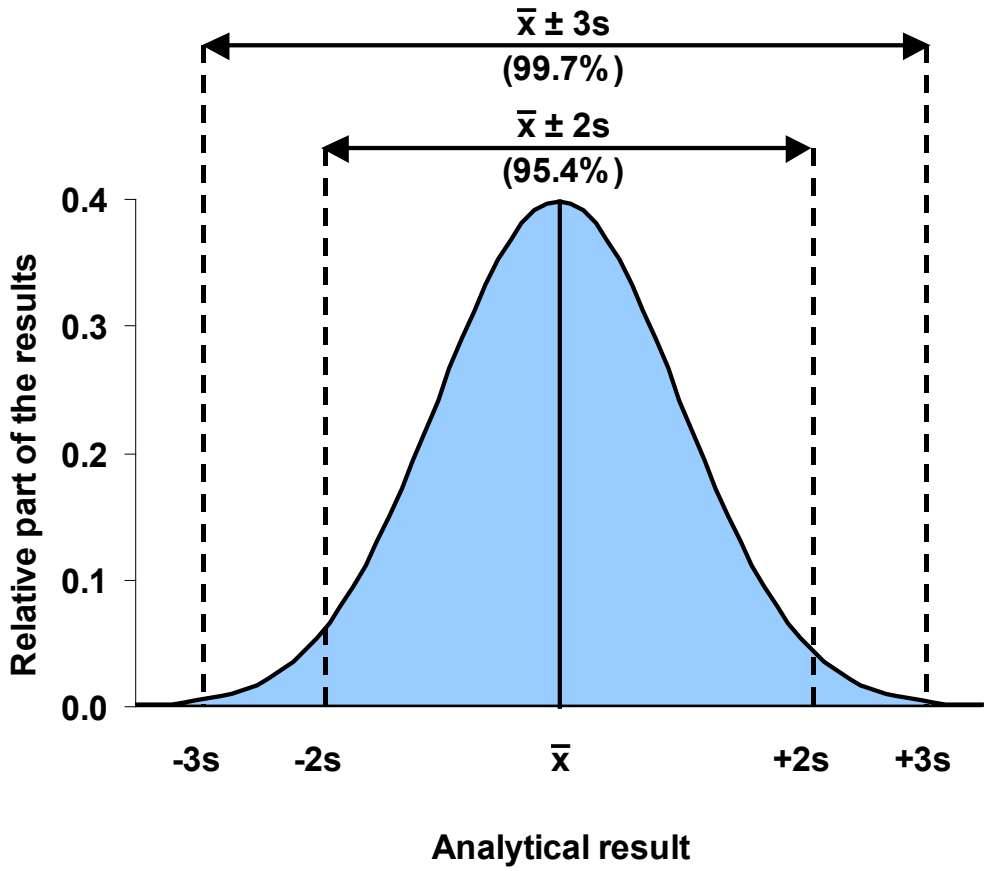


Figure 7. A normal distribution curve illustrating the probability for a result to be located within given limits ($\bar{x}$ is the mean value, s is the standard deviation).

3. Requirement for analytical quality

Here we describe how the laboratory can translate the customer's need for quality into terms applicable to internal quality control, i.e. within-laboratory reproducibility (s_{Rw}).

An analytical result can strictly speaking never be absolutely "correct". What ***is*** possible is to deliver a result with sufficiently small uncertainty for a given purpose, i.e. a result that is *fit for purpose*. Therefore, we need to know the intended use of the result before we can define the quality.

Figure 3 in Chapter 2 illustrates that the quality sufficient for one purpose is not necessarily sufficient for all other purposes. It is also extremely important to remember that it is always the intended use of the data, not the capability of the laboratory, that defines the necessary quality. Just as data can be too bad to be useful, it can also be too good, as too good often means too expensive or too slow to obtain!

An example: analysis of wastewater discharge is normally conducted to monitor discharges to check whether legally permissible quality limits are exceeded or not. These concentrations are relatively high compared to those in an unpolluted river or lake. Therefore, the required limit of detection can be relatively high. However, the measurement uncertainty must be adequate to ensure that the right decision is taken when comparing the result to the permissible limit.

Users of the results expect to be able to trust the data. In most cases, they do not have the expert knowledge necessary to explain exactly what they need. Hence the users rely on the laboratory to supply the right answer to the problem – that is to deliver a result that is fit for the purpose. Often, laboratories find it challenging to fully understand user needs. If the laboratory is accredited, the standard ISO/IEC 17025 requires the laboratory to evaluate the user's needs before any analyses are started.

Fortunately, most users for a specific parameter in a specific matrix, for example ammonium in drinking water, will need the analyses for the same purpose. Therefore, the requirements for quality will be the same. The laboratory does not need to think closely on the subject every day but can design its quality control programme such that the data delivered will have the right quality for the purpose.

However, the right quality requirements still need to be defined. In some cases, national or regional authorities have defined the required quality for regulatory analyses. For example, the European drinking water directive 2020/2184/EC [9] contains requirements for data quality. If no such national or regional requirements for quality exist, the laboratory must prepare its own requirements, preferably in cooperation with the end-users of the results.

Experience has shown that uncertainty in most analytical systems is proportional to concentration down to a limiting value at low concentration where the absolute uncertainty remains constant even though concentration in the sample decreases [4]. Requirements for quality will therefore often consist of two sets of values, one given in concentration units (describing the limiting minimum uncertainty at low concentration) and one in percent (describing the proportional component of uncertainty at higher concentrations).

Requirements for the target uncertainty are often described as a proportion (or percentage) of the concentration of primary interest. For example, the "concentration of primary interest" may, be a water quality limit or a similar maximum permissible concentration.

The requirement for quality may be given as a requirement for measurement uncertainty, but it is more common to give the requirements using quality characteristics that can be measured directly, for example by internal quality control. For internal quality control the quality characteristic needed is *within-laboratory reproducibility*, s_{Rw}. The example below shows how to start with quality requirements and from there estimate the demand for *within-laboratory reproducibility* to be used in internal quality control.

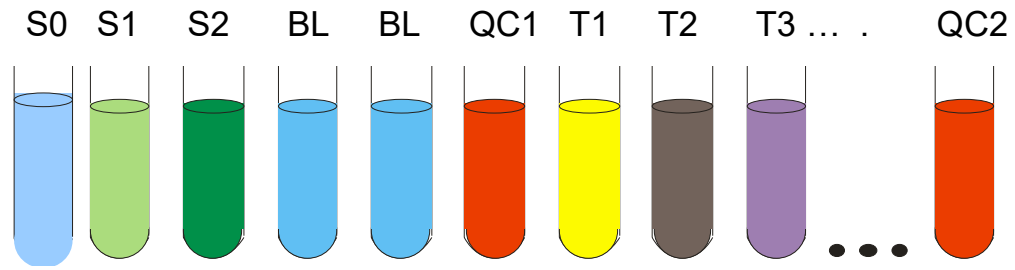
Example:
We are asked to determine ammonium in drinking water. The EU drinking water directive [9] states the required expanded measurement uncertainty at the level of 0.5 mg/L is 40 %. This guide proposes that a first estimate of s_{Rw} is the required U divided by a factor of 4 – see further Example 1 in Chapter 14.

Most laboratories will be able to determine ammonium with a relative s_{Rw} of 10 % at 0.5 mg/L. The result is the following requirements for s_{Rw}: 0.05 mg/L or 10 %, whichever is higher. In practice this means that for all concentrations below 0.5 mg/L the required s_{Rw} is 0.05 mg/L. From 0.5 mg/L and higher, the requirement is 10 % s_{Rw}.

4. Principles of quality control charting

This chapter describes the principles of quality control charts and what you do in the laboratory when running the control samples, plotting and evaluating the results.

Control charting is a powerful and a simple tool for the daily quality control of routine analytical work. The basis is that the laboratory runs control samples together with the test samples in an analytical run, as illustrated in *Figure 8*. The material used for control samples can be standard solutions, test samples, blank samples [18], in-house control materials and certified reference materials.



S0-S2 Standard solutions
BL Blank samples
QC Quality Control samples
T1… Test samples

Figure 8. Example of the analysis of two control samples in an analytical run.

In this case the first QC is QC1, a standard solution used to check the calibration - in this instance a chart with only action limits can be used; a decision is made as to whether the calibration is within a predefined requirement, and if so, the analyses can continue.
The second, QC2, is a control sample which is taken through the whole measurement procedure.

Immediately after the analytical run is completed the *control values* are plotted in a control chart. When reporting the control values, we recommend:

- giving one additional significant digit compared to test results;
- report **values** below reporting limit;
- report negative **values.**

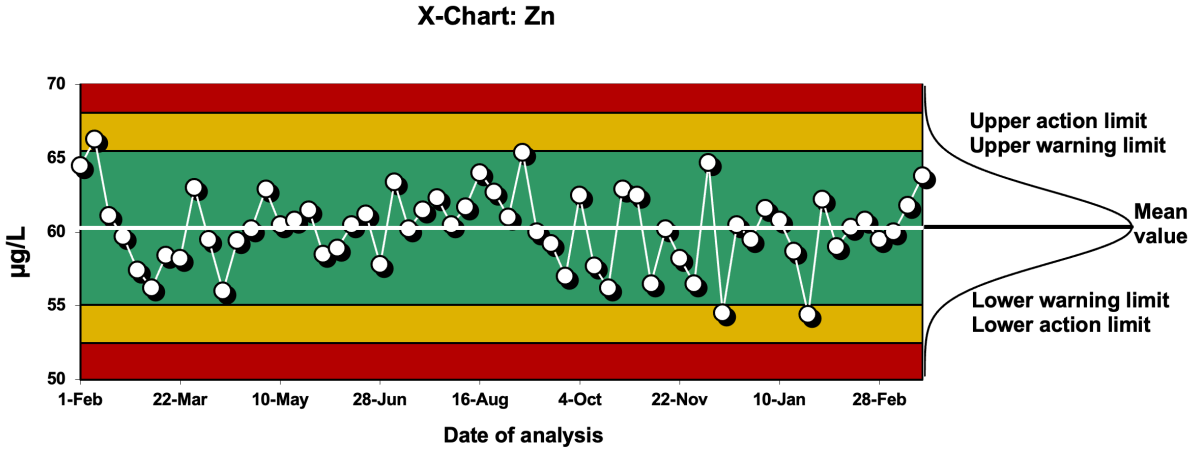


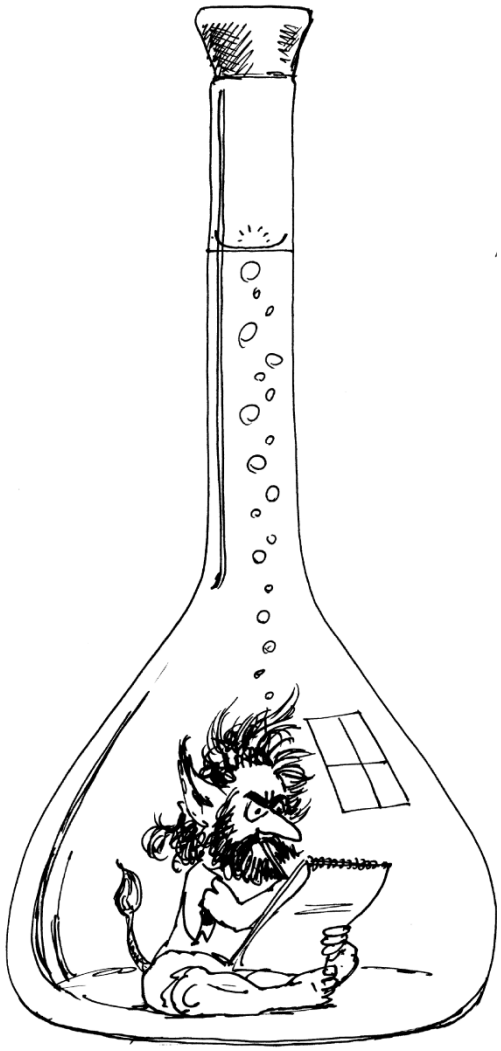
Figure 9. The relation between the normal distribution curve and the control chart. The central line is either a mean value or a reference value.

The chart is based on the statistical characteristics of random variations, defined by the normal distribution. The relation between the normal distribution curve and the control chart (X-chart) is illustrated in *Figure 9*.

The central line (CL) in the control chart represents the mean value of the control values or a reference value. In addition to the central line, the control chart normally has four lines. Two of these, the so-called warning limits, are located at a distance of ± two times the standard deviation from the central line (CL ± 2s). Provided that the results are normally distributed, about 95 % of the results should be within these limits. In the control chart two other lines are also drawn at a distance of ± three times the standard deviation from the central line (CL ± 3*s*). These lines are called the action limits and 99.7 % of the data normally distributed should be within these limits. Statistically only three out of 1000 measurement results are thus located outside the action limits for an X-chart.[5]

The warning and action limits can either be set as indicated above i.e. based on method performance data, *statistical control limits,* or based on independent quality criteria, i.e. *target control limits* – see Chapter 7.

Using the control charts, we should be alert if the control values are outside the warning limits or show trends. If values are outside the action limits no results should be reported – see Chapter 9.



[5] For a range chart nine out of 1000 measurement results from a normal distribution are located outside the action limit – false alarms of 0.93 %.

5. Different types of control charts

This chapter describes different types of control charts and when and what they can be used for.

The following types of control charts are the most important ones used for the internal quality control of chemical analyses [8, 10-15]:

- X-charts;
- R and r%-charts.

Charts normally have warning and action limits but charts with only action limits are also possible.

Charts with both warning and action limits

With the two limits, both changes in random error and systematic error are detected.

An X-chart with a central line, upper and lower warning limits and action limits.
One of the oldest and simplest types of control chart is the X-chart which is based on the distribution of the control values around a true or expected value. It can be used to monitor the combination of systematic and random effects for control values, based on single results or on a mean of multiple measurement results. Using a reference material similar to a test sample as the control sample, the bias may be monitored by comparing the mean control value over time with the reference value.

The *blank value chart* [8] is an X chart based on analysing a sample that can be assumed to contain the analyte at a very low level.[6] It offers valuable insights into contamination issues, reagent quality, and the measurement system's condition. A blank chart for Zn is presented in Example 9 in Chapter 14 showing the variation of a procedural blank. Remember that both positive and negative control values shall be plotted in the chart. Even though concentrations are normally entered into the blank value chart, it is also possible to use the measured signal.

Another special case is a *recovery chart* [8]. The analytical process may be tested for matrix influences by determining the recovery of additions of standards (spikes) to test samples. The % recovery can be plotted as the control value.

A range chart (R, r %) with a central line, an upper warning limit and an upper action limit.
The X-chart shows how well control values (mean values of multiple analyses or single values) are within control limits. By contrast, the range chart serves, above all, the purpose of demonstrating repeatability control. The range is defined as the difference between the largest and smallest result for two or more results. For practical applications in analytical laboratories the range chart mostly appears only in its simplest form, only duplicate determinations in each analysis series.

The most suitable samples for range charts are test samples chosen from those to be analysed in the analytical run. However, concentrations can vary because each analytical run involves different samples. As the range is normally proportional to sample concentration (at levels well above the reporting limit) it will be appropriate to use a control chart where the control value is the relative range, r%-chart (see Chapter 8). At levels close to the reporting limit, it is often appropriate to use the R-chart where the control values are the absolute range.

If, for test samples, single determinations are made, the control value for the range chart should be based on the difference between single determinations of two different sample aliquots. If on the other hand, test samples are run in duplicate for the X-chart, we recommend that the control value for a range chart be based on the absolute difference between mean value of duplicates – i.e. the same number of measurements for test samples as for control samples.

[6] In general, we have several different blanks, e.g. reagent blank, procedural blank and sample blank (19). The blank for the control value shall be different from blanks used for calibration or blank subtraction.

Charts with only action limits

These control charts are similar to the X and range charts but without any warning limits – only a central line and action limits [8].[7] These charts are suitable for deciding on instrument qualification - for example, whether a calibration is approved before continuing with the analyses, or a balance or a pipette is within tolerance. For an analytical balance the action limits can normally be set at ± 5 digits of the last digit; e.g. ± 0.5 mg for a 4-digit analytical balance. See further *Environmental parameter and instrument qualification* in Chapter 11. In Example 11 in Chapter 14 is shown a chart with only action limits for monitoring performance of a micropipette. The charts can also be used to demonstrate that a method is within a given tolerance – see Example 5. *For normal QC of a method, we recommend using charts with both limits.*

It has to be emphasised that a chart with only action limits will be less sensitive to drift, i.e. changes in the mean value, since the rule 2 out of 3 between warning and action limits (see Chapter 9) is not present. With no warning limits the estimated measurement uncertainty from control charts will increase. According to ISO 11352 [6] the estimated u_{Rw} (standard uncertainty component for the within-laboratory reproducibility) will increase with about 70 %.[8] Therefore the action limits when based on measured s_{Rw} can be set based on a factor lower than three – see further Chapter 7.

[7] Section 6.8 in ISO/TS 13530 describes *target control charts* with only action limits. The charts are according to the standard appropriate in the following cases when there are: no normal distribution of the values from the control samples (i.e. blank values), no available data for the statistical evaluation or external or internal prescribed bounds.

[8] How to evaluate measurement uncertainty based on control charts with only action limits will be addressed in the next version of Nordtest 537, which is scheduled for release in 2026.

6. Different control samples

This chapter describes the most common types of samples that are used as control samples in quality control.

Ideally, the control samples should go through the entire measurement procedure. The physicochemical properties should also be very similar to test samples and stable over time. There should also be a sufficient amount for years and a suitable analyte concentration. However, this is seldom the case and therefore we use several types of control samples:

I certified reference material – matrix CRM;
II standard solution or in-house material;
III blank sample;
IV test sample.

Control sample type I – certified reference material – matrix CRM

The results from repeated determinations of **a matrix** CRM will give a good indication of the presence of any systematic effect (bias). Repeated determinations in each analytical run provide the possibility of using the standard deviation (or range) as an estimate of the repeatability of the measurement. However, when a matrix CRM is used, there is generally a better repeatability compared to results obtained with a test sample, due to better homogeneity of the CRM.

A CRM is not always available for the desired sample matrix or concentration range. However, they are simple to use, and the results give immediate information on both systematic and random effects. Furthermore, the results provide the laboratories with an opportunity to calculate the measurement uncertainty of their results. Therefore, a CRM is recommended for use as often as practically and economically possible.

CRMs are purchased ready for use or with a procedure for preparation.

This control sample type is suitable for X-charts.

Control sample type II – standard solutions or in-house materials

Control sample type II, similarly, to type I, may give an indication of systematic effects as well as random effects.

A control sample type II is usually prepared by the laboratory. It can be either stable, homogeneous test samples or synthetic samples. For in-house matrix materials the laboratory collects the stable natural sample itself (or selects from samples received for analysis), making sure that the amount collected is sufficient for several years. Standard solutions (CRM, RM) or pure chemicals can be bought from external suppliers but are often prepared in-house. Synthetic control samples can be prepared from pure chemicals simulating the matrix of test samples. Due care should be taken to prepare this type of control sample – we recommend that the expanded uncertainty of the nominal value for the synthetic control sample should be less than one fifth of the standard deviation used to set up the control chart.

It is extremely important that chemicals used for preparation of synthetic materials are different from those used for calibration of the method. The difference can be either that the chemicals are purchased from different suppliers or for anions and cations that a different salt is used; for example, for nitrate that a Na-salt is used for calibration and a K-salt for control. Most laboratories prepare stock control solutions that are diluted daily or at intervals, according to the laboratory's experience for stability of the diluted solution. If the same chemical, or worse, the same stock solution, is used for calibration and control, any error in calibration, preparation or purity of the chemical will not be detected.

This control sample type is suitable for X-charts, and for R-charts if the control sample is a test sample.

Control sample type III – blank sample

Control sample type III may be used for the surveillance of the limit of quantification (LOQ). The blank may be a reagent blank, a procedural blank or a sample blank [18]. Furthermore, this type of control sample serves to reveal contamination.

The blank for the chart should be different from any blank used for calibration or for blank subtraction.

X-charts should be used, and R-charts can be used for this control sample type.

Control sample type IV – test sample

Control sample type IV is used when the spread for control sample Type I or II is less than for test samples, for example if only synthetic materials or extremely homogenized CRMs are available. It is also valuable when it is not possible to have a stable control sample – typical examples from environmental labs include the determination of dissolved oxygen, and chlorophyll *a*. Duplicate measurements give a realistic picture of the within-run random variations, the repeatability, for test samples.

The control sample will generally be randomly selected from the test samples submitted for measurement in the laboratory.

For this control sample type r%-charts are used in the higher concentration range and R-charts in lower concentration range.

NOTE The best control sample for X-charts (as well as range charts) is a stable in-house material similar to test samples. If a synthetic sample is used for X-charts and the repeatability for test samples is higher, it could be a good idea to include a control sample type IV.

7. Setting the control limits

Here we present how to set the central line and the control limits for X-charts and R-charts.

Control limits may be set according to the performance of the analytical method used irrespectively of the requirement on analytical quality – *statistical control limits*. This is the traditional method for setting the limits. An alternative is to start with the analytical requirements or intended use of the results. From the requirements a *target within-laboratory reproducibility* is estimated and if higher than the actual s_{Rw} for the method in routine use, target *control limits* can be set. Guidance on setting target s_{Rw} is given in Example 1 and 2 and in [15]. In the drinking water directive 2020/2184/EC [9] requirements on maximum measurement uncertainty are given. Preliminary target s_{Rw} can be based on an expanded uncertainty divided by four – see further Example 1.

Setting the control limits and the central line in X-charts

The control limits can be set based on method performance – **statistical control limits** or according to the requirement on *within-laboratory reproducibility* – **target control limits.**

Statistical control limits	**Target control limits**
The control limits are set based on the analytical performance of the control sample. From a long time period, e.g. a year, the standard deviation *s* is calculated from the control values. Warning limits will be ± 2 *s*. Action limits will be ± 3 *s*.	The control limits are set based on the requirement on the analytical quality. The standard deviation for the control chart, s_{target}, is estimated from the requirement on $s_{Rw.}$ The s_{target} is used to set the limits. Warning limits will be ± 2 s_{target}. Action limits will be ± 3 s_{target}.
Charts with only action limits	
Action limits **can** be ± 3 *s* or lower.[9]	Action limits will be ± tolerance.[10]

The central line in the control chart can be a determined mean value or a reference value.

Mean central line	**Reference central line**
The mean value is estimated from control values obtained over a long time period, e.g. a year. The central line is set to this mean value.	The control sample is a reference material with a nominal value. The central line is set to the nominal value.

If there is no reference value for the control sample, the central line is set to the mean value. When there is a reference value, you can choose how to set the central line. When the bias is small, it is recommended to set the central line to the reference or nominal value; wider control limits, target control limits, are then needed in order not to increase the number of out-of-control situations when the method is in control.

In the cases below, the control sample is a control sample similar to test samples and subjected to all steps of the analytical procedure. The number of samples analysed for control values is the same as used for routine measurements. The examples referred to below can be found in Chapter 14.

[9] The action limits, in a control chart without warning limits, can be set lower in order to get about the same frequency of out-of-control situations as for a chart with both limits. Simulations, with a shift of the mean value one standard deviation, *s,* show about 7 % out-of-control situations for a chart with both limits. For a chart with only an action limit the limit can be set to 2.5*s* instead of 3s in order to get about 7 % out-of-control situations for a chart with a shift of one *s* for the mean.

[10] The action limits (AL) are set based on the requirement.

Case 1. **Statistical control limits** and a **mean central line -** see Example 3 and Example 4.
The requirement on *within-laboratory reproducibility* is not set, and the method is performing with a s_{Rw} = 4 %. The warning limits are set to two times the standard deviation, ± 8 % and action limits to three times the standard deviation, ± 12 %. The mean value for the control sample is 59.2 µg/L so ± 8 % is equal to ± 4.7 µg/L and ± 12 % is equal to ± 7.1 µg/L. The warning limits will be at 59.2 ± 4.7 µg/L (54.5 and 63.9 µg/L)) and the action limits will be at 59.2 ± 7.1 µg/L (52.1 and 66.3 µg/L).

Case 2. **Target control limits** and a **mean central line –** see Example 1 and Example 2.
The requirement on *within-laboratory reproducibility* is, e.g. s_{Rw} = 5 % and the method is performing with a lower s_{Rw}. The warning limits are set to two times the standard deviation of the requirement, ± 10 % and action limits to three times the standard deviation, ± 15 %. The mean value for the control sample is 59.2 µg/L so ± 10 % is equal to ± 5.9 µg/L and ± 15 % is equal to ± 8.9 µg/L. The warning limits will be at 59.2 ± 5.9 µg/L (53.3 and 65.1 µg/L) and the action limits will be at 59.2 ± 8.9 µg/L (50.3 and 68.1 µg/L).

Case 3. **Target control limits and a reference central line** – see Example 5 and Example 7.
The requirement on *within-laboratory reproducibility* is, e.g. s_{Rw} = 5 % and the method is performing with a lower s_{Rw}. The warning limits are set to two times the standard deviation of the requirement, ± 10 % and action limits to three times the standard deviation, ± 15 %. The mean value for the control sample is 59.2 µg/L, but the reference value is 60.0 µg/L, so the warning limits will be at 60.0 ± 6.0 µg/L (54.0 and 66.0 µg/L) and the action limits will be at 60.0 ± 9 µg/L (51.0 and 69.0 µg/L).

Setting the control limit in R or r%-charts

For the range chart we only have upper limits – it is always positive. The control limits can be based on method performance – **statistical control limits** or according to the analytical requirement – **target control limits.** The control limits are calculated from a standard deviation. The factor used (2.83 & 3.69) for calculating the control limits for duplicates can be found in Table 4 in Chapter 13 and the background to these factors is also explained there.

Statistical control limits	**Target control limits**
The control limits are set based on the analytical performance of the method. From a long time period, e.g. a year, a pooled *s* is calculated. Central line is the mean range. For duplicates $n = 2$. Upper warning limit will be + 2.83 *s*. Upper action limit will be + 3.69 *s*.	The control limits are set based on the requirement on repeatability. From the requirement a standard deviation s_{target} is estimated for this control chart. Central line is the mean range. For duplicates $n = 2$. Upper warning limit will be + 2.83 s_{target}. Upper action limit will be + 3.69 s_{target}.
Charts with only action limits	
Action limit can be + 3.69 *s* or *lower*.[11]	Action limit will be + tolerance.

Case 1. Statistical control limits – see also Example 3 (R-chart) and Example 6 (r%-chart) in Chapter 14. The pooled standard deviation is = 0.356 %. The warning limit for the range chart will then be set at + 3·0.356 = 1.0 % and action limit 3.69·0.356 = 1.3 %.

Case 2. **Target control limits.**
The *repeatability limit*, *r* is often given in standard methods and in this case as 1 %; in 19 times out of 20 the difference between two results should be less than 1 %. From this limit the required repeatability standard deviation is calculated as s_r = r/2.8= 0.357 %.[12] The warning limit for the r%-chart will then be set at + 2.83·0.357 = 1.0 % and the action limit at 3.69·0.357 = 1.3 %.

Target control limits – estimating the *s* for the control sample

When the control sample encompasses the whole analytical process from the sample entering the laboratory to the analytical report the control values will demonstrate the *within-laboratory reproducibility,* s_{Rw}, and one can compare the obtained s_{Rw} with the requirement. With most other control samples, e.g. standard solutions, blank samples, the obtained standard deviation is only part of the s_{Rw}. Here the analyst should determine whether the *s* obtained for the control sample is sufficiently low to fulfil the analytical requirement.

[11] The action limits, in a control chart without warning limit, can be set to 3.5 *s* in order to get about the same frequency false out-of-control situations as a normal range chart with both limits. For a normal distribution the range chart, with action and warning limits, will have a frequency of false out-of-control situations of 1.31 %.

[12] The factor 2.8 comes from error propagation of a difference where the repeatability limit is equal to $2 \cdot \sqrt{2} \cdot s$.

Recommendations

Start of QC - To start the quality control of a new method preliminary control limits (set slightly wider) and central line can be estimated based on about 25 control values. Only after a longer time period, e.g. one year, can the control limits and the position of the central line be fixed. These first ***preliminary*** warning and action limits can also be based on results from method validation.

Fixed control limits – We do recommend fixed limits and not limits that are constantly changing for stable control samples. To obtain reliable control limits the calculated standard deviation should be based on control values over a one-year period and at least 60 control values. If the time period is shorter, the estimate of the standard deviation obtained is 1) unreliable (too few results) and 2) usually too low because not all variation over a longer time is taken into account.

Fixed central line – We recommend a fixed central line. To obtain a reliable mean value one-year period may be sufficient. If the time period is shorter, an unreliable estimate is often obtained.

Replicate analyses/samples - We also recommend the same number of sub-samples being used both for test samples and control samples – if we report the mean value of duplicates for test samples, we should plot the mean value of duplicate analyses for the control sample in the X-chart.

Multielement analyses – When many analytes are measured in the same analytical run in QC, e.g. ICP, XRF, GC, we strongly recommend using target control limits or wider statistical limits for those analytes that are less important. If for example using X-charts 20 analytes are determined[13] and statistical control limits are used for all analytes, the frequency of false out-of-control increases from 0.54 % to 10.3 % for a method in control. The same problem occurs if you have a long run with many control samples. By using target control limits based on the requirement the false out-of-controls can be drastically decreased while the analytical results are still fit for purpose.

[13] This applies to independent measurements and, to a lesser extent, also to measurements which are partially correlated such as ICP, XRF etc.

8. Setting up a quality control programme

This chapter describes how to start setting up QC for a measurement procedure: selection of number of control samples, type of charts and frequency of control analyses.

An example of setting up the QC - Cd determination in fresh waters

Cadmium concentration can normally vary between 0.01 µg/L and 100 µg/L in different types of waters. For quality control of Cd in fresh water using ICP-MS (LOD 0.01 µg/L) we have chosen the control samples as follows:

Control samples	Control chart	Control limits	Central line
A CRM, Cd: 2.28 µg/L (Type I)	X-chart	Statistical	Reference value
A standard solution, Cd: 20 µg/L (Type II)	X-chart	Statistical	Nominal value
An in-house material, Cd: 0.10 µg/L (Type II)	X-chart	Target	Mean value
Duplicate determinations of test samples (Type IV)	R-chart/r%-chart	Target	Mean value

Because of the rather wide concentration range in test samples, we have chosen 3 QC samples for X-charts. The standard solution of 20 µg/L is prepared from a stock solution, which is not the same stock solution as used for the preparation of the calibration solutions. The in-house material, acidified lake water was prepared for quality control of low Cd content in fresh water.

For a check of systematic effects in our measurement procedure, we use a CRM with a certified Cd content of 2.279 ± 0.096 µg/L.

To get a realistic picture of the repeatability for test samples we select at random one sample in each analytical run and this sample is analysed as duplicates (two different test tubes in the autosampler).

For measurement of Cd using ICP-MS we may carry out as many as 200 determinations in each analytical run. At the beginning and at the end of each run we analyse the CRM, the standard solution, the in-house material and the calibration standards. To check calibration drift during a run, we normally analyse one of our standard solutions about every 20 analyses.

All the results obtained for the control samples are plotted in X-charts using our LIMS.[14] The results of duplicates, the range, obtained in analysis of test samples are plotted in an R-chart at lower concentrations and in an r%-chart at higher concentrations.

NOTE If a control sample is analysed several times in the same run, either one or all control values can be plotted in the X-chart – this should be stated in the method description.

Practical points in setting up the QC

A method validation is normally performed before a measurement procedure is adopted. When setting up a programme for control charting, (such as selection of control samples, type of control charts and control frequency) the results of the initial tests for establishing performance of an analytical method may give valuable background information about, e.g. the concentration range, the stability and systematic effects. In particular, a within-laboratory reproducibility of measurements of different concentrations obtained during method validation forms the first basis for routine quality control.

Concentration range - In analysis of environmental samples concentrations of an analyte may vary considerably. In such cases it may be necessary to utilise separate X-charts and range charts for different concentration levels.

Range chart with test samples – To monitor repeatability using range charts (R-chart or r%-chart) we recommend analysing a test sample in duplicate in each analytical run. A test sample is selected at random and representative of the concentration range and matrix variations of the analytes being studied.

Frequency of control analyses - Generally, as a minimum, one control sample in each analytical run must be analysed for detecting possible systematic effects within the analytical run, for example from calibration. Stability of the measurement system can have an influence on the frequency of control analyses. If there are errors caused by calibration drift, the number of control samples to be analysed in each analytical run may need to be higher than under very stable measurement conditions. The principle guiding the decision on the number of times a control sample must be analysed in each analytical run is that all measurements performed after the last approved sample in the quality control may have to be reanalysed. The frequency of control is therefore a balance between the cost of the control and the cost of repeating analyses. When using automatic analysers, e.g. overnight, several control samples may be analysed in each analytical run.

Position of control samples in an analytical run - The analyses of control samples should in principle be carried out in random order to eliminate any systematic effects. However, we recommend that for an analytical run control samples are analysed at least at the beginning of each run and before finishing the run, in case a drift in the analytical process can cause errors.

A good balance between QC and test samples – QC fit for purpose. In this example, Cd in fresh water, we use several QC samples but, in most cases, fewer control samples will be sufficient.

QC programme in a method description and in a quality manual

The principles of the quality control programme covering the practical points mentioned above should be documented, e.g. described in the quality manual of the laboratory. Quality procedures should also be presented in detail in the procedure for each analytical method.

[14] Laboratory Information Management System

9. Daily interpretation of quality control

In this chapter we describe the interpretation after each analytical run. Can we report the results or not?

A practical procedure for the registration of the control data is to write down all information that may be significant for the interpretation of the control data. Typical examples are when new stock or control solutions have been prepared, e.g. a change of reagents, a change of measurement cell, and instrumental problems. If all information is properly documented it is, at a later time, possible to check the conditions for this measurement, e.g. in out-of-control situations.

For each batch of analysis there is normally at least one control value for each chart. In daily work it is essential to be alert if a control value is falling outside the control limits or if a certain systematic pattern is observed in the control values over a period.

Daily interpretation

There are two possible cases:
1. the method is in control;
2. the method is out-of-control.

For a chart with warning and action limits

1. The method is **in control** if:

- the control value is within the warning limits; or
- the control value is between the warning and the action limits[15] and the two previous control values were within the warning limits.

In this case the analytical result can be reported.

2. The method is **out-of-control** if:

- the control value is outside the action limits; or
- the control value is between the warning and the action limit and at least one of the two previous control values on the same side is also between warning and action limit on the same side of the central line – the rule two out of three – see for example March 22 in Figure 10.

In this case normally no analytical results can be reported. All samples analysed since the last in control value was obtained must, if possible, be reanalysed. Note that the false out-of-control frequency with these two rules, is about 0.54 % for an X-chart.[16]

For a chart with only action limits

1. The method is **in control** if:

- the control value is within the action limits

2. The method is **out-of-control** if:

- the control value is outside the action limits;

[15] The method is in control if one of the previous values was between warning and action limits on the other side.
[16] The frequency for false out-of-control for the 3*s* rule is about 0.27 % and for the rule *two out of three **on the same side*** the frequency is also 0.27 % for a method in control. In total you will have 0.27 % + 0.27 % = 0.54 % false out-of-control situations with statistical control limits and a central line set to the mean value.

Out-of-control situations

It is difficult to give general guidelines for how the laboratory should act when the analysis is out-of-control. The experience and common sense of the analyst is of vital importance when choosing remedial actions. However, if an out-of-control situation occurs, it is most likely that there is also an error in the analyses of test samples.

The normal action for an out-of-control situation is to repeat the control analyses twice. If the new control values are located *within the warning limits* the test samples can be reanalysed. If the control values are still outside the warning limits, the analyses shall be stopped, and remedial actions have to be taken to find and eliminate the cause(s) of error. If the test samples cannot be re-analysed, for example due to instability, and the customer still urgently needs a result the laboratory can decide (after careful consideration) to report the value, provided that a clear note on the increased uncertainty is given.

Usual remedial actions in out-of-control situations include 1) checking the reagents, 2) calibration of the equipment or 3) exchange of vessels and equipment. The problem, and the solution to this, should be documented.

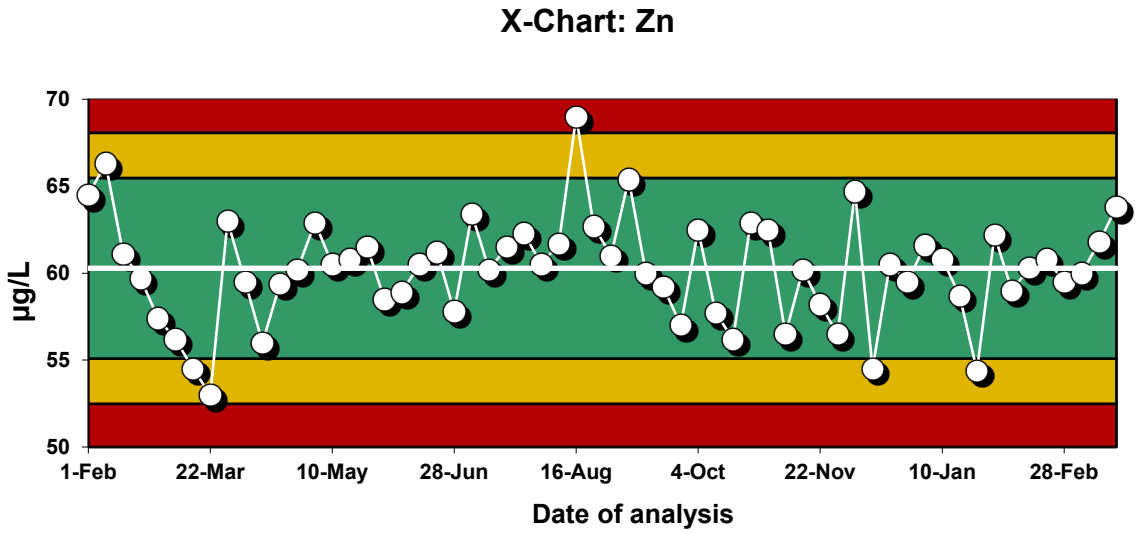


Figure 10. X control chart with two out of control situations.

Method out-of-statistical control

The method is **in control** but can be regarded as **out-of-statistical control** if:

- seven control values in consecutive order gradually increase or decrease;
- ten out of eleven consecutive control values are lying on the same side of the central line.

These rules originate from process control in production *to avoid waste by not producing unusable output* as stated in ISO 7870-2 [11]. However, these rules are normally not practical to the **daily** control of an analytical method, but important trends should be discovered as early as possible to avoid problems in the future – see the long-term evaluation of the results in Chapter 10 and also Example 8 in Chapter 14.

NOTE: When the central line is set at a reference value several values can be on one side due to a small bias. The laboratory must decide in the long-term evaluation if this bias is acceptable.

10. Long-term evaluation of quality control data

This chapter is about using the quality control data from a long period of time to answer two questions:

1. *what is the quality (random and systematic effects) currently in the laboratory? Has the quality significantly changed?*
2. *are control limits and central line in the control chart still optimal for detecting out-of-control situations?*

Note: This is one of the most difficult tasks in QC and we can only give general guidance. The frequency may vary but a yearly evaluation is suitable in most cases.
We will look at these two questions below.

Review of the current quality

The evaluation consists of a review of the quality control data (X, R and r% charts). Use at least all data from the last 12 months. If possible, there should be at least 60 control values. If there are fewer, use data from previous years as well; at least 20 should be from the last year. Check if there is any significant change:

1. in the standard deviation using an F-test;
2. in the mean using a *t*-test.

The equations for *F*-test and *t*-test are given in Chapter 12 and explained in detail in Example 8.

If the number of data points are about 60 for an X-chart with warning and action limits, the following simpler check applies [8]:

1. **standard deviation** – If you are using statistical control limits, count the number of cases outside the warning limits. If this number is 6 or lower there is no evidence (with 60 data points) that the standard deviation has increased. If the number is greater than 6 there is probably an increase in standard deviation (with 60 control values). In this case perform an *F*-test to verify the increase;
2. **mean value -** Calculate the mean and compare with the previous mean value. If the difference is less than 0.37 *s* no action is needed. If the difference is more than 0.37 *s* there is evidence that the mean value may have changed.[17] Then perform a *t*-test to verify this change.

How often should control limits be evaluated?

For successful use of control charts, it is important that the control limits and the central line remain stable over a long period of time – several years. The central line and control limits should not be changed frequently since this will make it difficult to detect gradual changes in analytical quality. The laboratory should have a policy for how often control limits are evaluated and how it is decided if a change is needed. We recommend that control limits and central line should be evaluated every year. For less frequent analyses, for example those performed once per month, we recommend evaluation after data from 20 control samples have been collected.

[17] Using Equation 7 a *t*-value can be calculated and compared with the critical *t*-value. If the difference is less than 0.37 *s* there is no significant difference (the calculated *t*-value us less than the critical *t*-value) between the means, with 60 control values for both the old and the new mean value.

When a change in control limits may be necessary

Target control limits are only changed if customers' requirements change. The recommendations below are therefore only relevant for statistical control limits.

A change of control limits should only be considered if a significant change in standard deviation has taken place. If an increase in spread is significant and if the change is acceptable compared to customers' requirements, calculate new warning and action limits as described in Chapter 7.

Special care must be taken when a control chart includes out-of-control situations (see Chapter 9). If an assignable cause for the out-of-control situation was identified at the time of the analysis, the control value should be excluded from the calculation of new control limits. However, there will inevitably be cases where out-of-control situations have existed, but no assignable cause identified. These data could probably be the result of an undetected mistake for that particular batch of analyses and including them in calculations may lead to a falsely large standard deviation. On the other hand, excluding such data, especially if there is more than one in the data set, may lead to a too optimistic standard deviation and falsely contract the control limits, leading to even more apparent out-of-control situations. A pragmatic approach is to exclude data that are more than 4 standard deviations away from the central line and retain the rest.

If there are two or more out-of-control situations in 60 points under consideration, this is more than could be expected and there is good reason to scrutinise the whole analytical procedure to search for the cause of the repeated out-of-control situations.

When a change in central line may be necessary

A reference central line is fixed. The recommendation below is therefore only relevant for a mean central line.

A change of the central line should only be considered if there is a significant change in the mean value. However, even if the change is significant, we do not recommend changing the central line unless there is a good explanation for the shift in data e.g. a new control sample.

11. Other uses of quality control data and control charts

The information obtained from the regular use of control charts can be used for purposes other than pure internal quality control. We here propose a few suggestions.

Measurement uncertainty

Results from the control charts can, together with other data be used for evaluating the measurement uncertainty. In most cases, the systematic effect and the random effect (the standard deviation) can be combined to evaluate the measurement uncertainty. How this can be done is described in detail in the Nordtest *Handbook for calculation of measurement uncertainty in environmental laboratorie*s (4) and partly in the Eurachem/CITAC guide (7).

Measurement uncertainty is estimated from control chart results combined with those from proficiency tests, method validations or information given in standard methods. This approach provides a practical and general way of utilising already existing information. Provided the whole analytical chain is included in the measurement of control samples for charting (including sample work-up such as filtration, concentration steps etc.), the estimate of the measurement uncertainty may be realistic.

Method validation

Normally, a method validation should be performed **before** a method is adopted in the laboratory. There might be situations, though, where a method is used after only partial validation or verification, and where information from the control charts can be used to complement the available data. Such situations could occur if a method has been changed only slightly, or if a standard method is adopted in the laboratory.

- If a matrix CRM similar to test samples is used as the control sample, the results will give direct information on the bias of the method, by comparing the resulting average result with the expected (certified) value.
- All types of control charts will provide information on the spread (random variation) from calculations of standard deviation.

Method comparison

Control charts can be used to compare different analytical methods using separate control charts for each method. This may for example give valuable method comparison information if the laboratory is in the process of changing from a manual to an automated method, or from a standard method to a non-standard method (e.g. a test-kit method). By running the two methods in parallel for some time, it is possible to compare important information such as:

- spread (from the standard deviation or from the range);
- bias (if a CRM is used);
- matrix effects (interferences), if spiking or a matrix CRM is used;
- robustness, i.e. if one method is more sensitive to small changes such as temperature shifts, handling etc.

Estimation of limit of quantification (LOQ)

The estimate of limit of quantification is usually a standard deviation multiplied by a factor. The factor is normally six to ten. For further guidance see reference [16].

Data from an X-chart with a test sample at low concentration will be useful for the estimation of the standard deviation for the method in routine use. Data from control sample type III (procedural blank sample) may in some cases be used for the estimation, provided that the laboratory has evidence that the standard deviation for the blank is representative for the standard deviation for test samples with low concentration – see Example 9.

Data from an R-chart will give the repeatability standard deviation, and if the concentration is low, this standard deviation, after correction for the number of measurements and the number of blank measurements, is useful for estimation of the LOQ. For the detailed calculation of LOQ see the Eurachem Guide [16].

Person comparison or assessing competence

In the same way as for methods, it is possible to compare the performance of different persons in the laboratory. Whereas this might be viewed as undesired policing, there is no doubt that control charts can be very useful tools when training and demonstrating competence of new staff in the laboratory. Part of the training activity will be to plot results from control samples analysed by the person under training in control charts and to set target values for permissible systematic errors and spread, then comparing this to what is reached by the experienced trained staff. In this way, the laboratory manager as well as the trainee will have a very objective tool for judging when the performance in the analytical work is sufficient to fulfil the requirements.

Evaluation of proficiency testing - PT

When the laboratory regularly participates in proficiency tests, plotting the PT results (*z*-scores or zeta-scores) in control charts (similar to an X-chart) provides a good overview of performance, including possible systematic effects or trends.

Here the *z*-score and *zeta*-score can be plotted in an X-chart. Normally CL = 0, WL = ± 2 and AL = ± 3.

$$z = \frac{(x_{lab} - x_{ref})}{s} \text{ and } zeta = \frac{(x_{lab} - x_{ref})}{\sqrt{u_{lab}^2 + u_{ref}^2}}$$

Example, the standard deviation in a proficiency test (all laboratories) was 0.08 mg/kg. Your result x_{lab} was 0.12 mg/kg lower than the reference value x_{ref} given by the PT provider. Your *z*-score becomes -1.5. Here we recommend that all values outside warning limits should be investigated. The maximum permissible error from authorities can also be used to calculate the *z*-score.

The standard deviation, *s,* for calculating the *z*-score is determined by the proficiency testing provider and can be different for different providers e.g. the *s* in the current round, based on previous experience or a fixed value.

Another possibility is to calculate the *zeta*-score using your own claimed standard uncertainty, u_{lab} and the standard uncertainty of the reference value, u_{ref}. A *zeta* value within ± 2 shows that the laboratory performs according to the stated measurement uncertainty.

Environmental parameters and instrument qualification

When monitoring environmental parameters in the laboratory, such as the temperature in the laboratory or in the refrigerators, it is very useful and easy to use a simple type of control chart with only action limits for plotting the observed control values. In such cases the target value will be used as the central line, and the permissible limits used as action limits. In Example 11 in *Chapter* 14 is an X-chart for controlling the bias for a 1 mL pipette. The control charts give a graphical presentation of any trends or unexpected variation that might influence the analyses and therefore might be worth investigating.

Similarly charts with only action limits are useful to plot the results of the verification of an analytical balance or other regular checks, partly to detect any trends as well as to see if the results are outside or inside the permissible limits.

12. Terminology and equations

*Here we try to describe terminology and the statistical equations we use in this handbook. Some of the definitions here are simplified. Exact definitions for terms used are found in VIM [17] and further explained in the Eurachem Guide [18]. Direct quotes from VIM are given below in italics. All terms defined here are given in **bold** text.*

Terminology

Accuracy of measurement
Closeness of the agreement between the result of a measurement and a true value of the **measurand** [17]. The accuracy is affected by both systematic and random errors.

Analyte
The substance subject to measurement.

Analytical run - batch of analyses
Analyses of a number of test samples and **control samples**. Normally one **control value** for each batch is entered into each **control chart**.

Bias – systematic error
Estimate of a systematic measurement error [17]. The bias is estimated as the difference between the **mean value** of a large number of test results and the accepted reference value. (*Figure 6*).

Confidence level[18]
The probability that test results will fall within a specified range.

Control chart
The principal tool in internal quality control is a chart where the x-axis is time, and the **control values** are entered and compared with **control limits**.

Control limits
Limits in a **control chart**. There are two control limits: action limits (AL) and warning limits (WL).

Control sample
Sample material whose test results are used to construct **control charts**, e.g., reference materials, standard solutions, test samples or blank samples.

Control value [8]
Result from the internal quality control. It can, e.g. be a single value, a mean value or a range. These values are reported differently from test results - values from analyses of test sample: **Control values** are reported with one extra significant figure, and also negative values are reported, e.g. a control value – 0.03 mg/L in an X-chart could for a test sample be reported as < 0.1 mg/L.

Degrees of freedom, *df*
The number of independent comparisons that may be made between individual results in a data set. In general terms the number of degrees of freedom, e.g. for an estimated standard deviation, provides an indication of the reliability of the estimate. As the number of degrees of freedom increases, the random error of the estimate itself, *s*, decreases. The degrees of freedom are used when comparing statistical quantities, see *F*- and *t*-test below.

Detection limit (LOD)
The lowest concentration of an **analyte** that can, with a given probability, be detected using a specified method.

[18] VIM uses **coverage probability** and defines the term as *probability that the set of true quantity values of a measurand is contained within a specified coverage interval.*

Limit of Quantification (LOQ)
LOQ is the lowest level of analyte that can be determined with acceptable performance [16]. The reporting limit (results reported <) is usually the LOQ or higher.

Measurand
Quantity intended to be measured [17]. e.g. the concentration of acid-soluble cadmium (the **analyte)** in a fresh-water sample.

Measurement procedure
The detailed description of an analytical method used in a laboratory. Also called method instruction or method documentation protocol [16].

Measurement uncertainty
Non-negative parameter characterizing the dispersion of values being attributed to a measurand, based on the information used [17]. Measurement uncertainty can be interpreted as a quantitative estimate of **accuracy** (**trueness** + **precision**) – see *Figure 3*.

Outlier rejection
In the statistical calculation we recommend rejecting outliers that are more than 4 *s* different from the mean. Another alternative is to use Grubbs test – see statistical textbooks [13].

Precision – see **spread**

Range
The difference between the highest and lowest result. The number of results for calculating the range can vary but are usually just two in QC ($n = 2$).

Repeatability
*Measurement **precision** under a set of repeatability conditions of measurement* [17].
Repeatability conditions of measurement refer to measurements being made on the same material by a single analyst, using the same procedure, under the same operating conditions over a short time period. The whole procedure should be repeated from taking a new test portion of a sample to the final reading or calculation of result.

Reproducibility
*Measurement **precision** under a set of reproducibility conditions of measurement* [17].
Reproducibility condition of measurement refers to measurements being made on the same material using the same procedure but by different analysts working in different locations.

Within-laboratory reproducibility (Intermediate precision)
The degree of agreement between individual results determined in a laboratory on a sample with the same **measurement procedure** over a long time period, i.e. at least a year. The time could be shorter if enough data are collected but in many cases a year is suitable to encompass all variations in reagents, personnel, instrument service, etc. Also called intermediate **precision** [17].

Spread/Precision
The variation between independent test results obtained under stipulated conditions.

Systematic error
Component of measurement error that in replicate measurement remains constant or varies in a predictable manner [17]. Systematic error is normally expressed in terms of **bias.**

Test result (response value)
The value obtained by applying the **measurement procedure**. The **control value** entered in the **control chart** is either the test result of a **control sample** or a value calculated from the test results, e.g. the range.

Trueness
Closeness of agreement between the average of an infinite number of replicate measured values and a reference value [17].

Equations

Mean value ($\bar{x}$)

The sum of all individual results (x_i), divided by the number (n) of results:

$$\bar{x} = \frac{\sum x_i}{n} \qquad \text{Eq. (1)}$$

Standard deviation (s)

A measure of the **spread** (precision) of individual results (x_i) around the **mean value** ($\bar{x}$):

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n-1)}} \qquad \text{Eq. (2)}$$

Degrees of freedom, $df = n - 1$

Coefficient of variation (CV) or **relative standard deviation** in % (RSD %). The standard deviation expressed in percent of the **mean value**:

$$CV\ (\%) = \frac{100 \cdot s}{\bar{x}} \qquad \text{Eq. (3)}$$

Standard deviation from mean range (n=2) Calculated for the application of range charts. For factors for n equal to 3 to 5 see Chapter 13, *Table 4*.

$$s_r \approx \frac{Range}{1.128}\ (n = 2) \qquad \text{Eq. (4)}$$

Note - A pooled standard deviation is more correct to use – see Equation 9 and 10. Using Equation 9, series of analyses with different numbers of results can be used to estimate the standard deviation.

***F*-test**

Used to evaluate whether the **standard deviations** (s_1 and s_2) from two series of determinations are significantly different:

$$F = {s_1}^2 \Big/ {s_2}^2, \text{ where } s_1 > s_2 \qquad \text{Eq. (5)}$$

When the calculated *F*-value is greater than the critical *F*-value found in *Table 3* the two standard deviations are significantly different.

***t*-test**

Used to evaluate whether there is a significant difference between the **mean value** ($\bar{x}$) for a series of determinations and the accepted reference value (T):

$$t = \frac{|\bar{x} - T|}{s} \cdot \sqrt{n} \qquad \text{Eq. (6)}$$

alternatively, between the mean values ($\bar{x}_1$ and $\bar{x}_2$) of two different series of analyses:[19]

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p} \cdot \sqrt{\frac{n_1 \cdot n_2}{(n_1 + n_2)}} \qquad \text{Eq. (7)}$$

where s_p is the pooled **standard deviation**, see Equation 9.

When the calculated *t*-value is greater than the critical *t*-value found in *Table 2*, the difference is statistically significant. **Degrees of freedom**, ***df*** are ***n-1*** for Equation 6 and $\boldsymbol{n_1+n_2-2}$ for Equation 7.

[19] This is the *t*-test, which assumes that the two measurements series have approximately the same standard deviation.

Weighted mean ($\bar{\bar{x}}$) for several series of analyses

Calculated from the mean values for k series of analyses with total of $n_1+n_2+\ldots= n_{tot}$ observations:

$$\bar{\bar{x}} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2 \ldots n_k \cdot \bar{x}_k}{n_{tot}} \qquad \text{Eq. (8)}$$

Pooled standard deviation (s_p) for several series of analyses

Calculated from the standard deviations for k series of analyses with total of $n_1+n_2+\ldots= n_{tot}$ observations:

$$s_p = \sqrt{\frac{(n_1-1)\cdot {s_1}^2 + (n_2-1)\cdot {s_2}^2 + \cdots + (n_k-1)\cdot {s_k}^2}{n_{tot}-k}} \qquad \text{Eq. (9)}$$

Degrees of freedom, $df = n_{tot} - k$.

If n is about the same for the different k series

$$s_p = \sqrt{\frac{{s_1}^2 + {s_2}^2 + \cdots + {s_k}^2}{k}} \qquad \text{Eq. (10)}$$

13. Tables

First table in this chapter is Table 2. Table 1 can be found on page 5.

*Table 2. Critical t-values (**2-sided test**). Usually, the 95 % confidence level is used.*

Degrees of freedom	Confidence level (%)				Degrees of freedom	Confidence level (%)			
	90	95	99	99.9		90	95	99	99.9
1	6.31	**12.7**	63.7	637	21	1.72	**2.08**	2.83	3.82
2	2.92	**4.30**	9.92	31.6	22	1.72	**2.07**	2.82	3.79
3	2.35	**3.18**	5.84	12.9	23	1.71	**2.07**	2.81	3.77
4	2.13	**2.78**	4.60	8.61	24	1.71	**2.06**	2.80	3.75
5	2.01	**2.57**	4.03	6.86	25	1.71	**2.06**	2.79	3.73
6	1.94	**2.45**	3.71	5.96	26	1.71	**2.06**	2.78	3.71
7	1.89	**2.36**	3.50	5.41	27	1.70	**2.05**	2.77	3.69
8	1.86	**2.31**	3.36	5.04	28	1.70	**2.05**	2.76	3.67
9	1.83	**2.26**	3.25	4.78	29	1.70	**2.05**	2.76	3.66
10	1.81	**2.23**	3.17	4.59	30	1.70	**2.04**	2.75	3.65
11	1.80	**2.20**	3.11	4.44	35	1.69	**2.03**	2.72	3.59
12	1.78	**2.18**	3.05	4.32	40	1.68	**2.02**	2.70	3.55
13	1.77	**2.16**	3.01	4.22	45	1.68	**2.01**	2.69	3.52
14	1.76	**2.14**	2.98	4.14	50	1.68	**2.01**	2.68	3.50
15	1.75	**2.13**	2.95	4.07	55	1.67	**2.00**	2.67	3.48
16	1.75	**2.12**	2.92	4.02	60	1.67	**2.00**	2.66	3.46
17	1.74	**2.11**	2.90	3.97	80	1.67	**1.99**	2.64	3.42
18	1.73	**2.10**	2.88	3.92	100	1.66	**1.98**	2.63	3.39
19	1.73	**2.09**	2.86	3.88	120	1.66	**1.98**	2.62	3.37
20	1.72	**2.09**	2.85	3.85	∞	1.64	**1.96**	2.58	3.29

Table 3. Critical F-values at the 95 % confidence level (2-sided test) for number of degrees of freedom, df, from 4 to 120 where s1 > s2.

Values of $F_{1-\alpha}$ (df_1, df_2), α = 0.025														
df_1	**4**	**5**	**6**	**7**	**8**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**
df_2														
4	9.60	9.36	9.20	9.07	8.98	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31
5	7.39	7.15	6.98	6.85	6.76	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07
6	6.23	5.99	5.82	5.70	5.60	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90
7	5.52	5.29	5.12	4.99	4.90	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20
8	5.05	4.82	4.65	4.53	4.43	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73
10	4.47	4.24	4.07	3.95	3.85	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14
12	4.12	3.89	3.73	3.61	3.51	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79
15	3.80	3.58	3.41	3.29	3.20	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.45
20	3.51	3.29	3.13	3.01	2.91	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.14
24	3.38	3.15	2.99	2.87	2.78	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01
30	3.25	3.03	2.87	2.75	2.65	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87
40	3.13	2.90	2.74	2.62	2.53	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72
60	3.01	2.79	2.63	2.51	2.41	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58
120	2.89	2.67	2.52	2.39	2.30	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43

df_1 = degrees of freedom in the numerator (s_1^2) and df_2 in the denominator (s_2^2) NOTE $s_1 > s_2$

Table 4. Factors for calculation of standard deviation from mean range (max-min), warning and action limits for construction of R and r%-charts. Factors obtained from ISO 7870-2 [11].

Number of replicates	**Standard deviation, *s***	**Warning limit WL***	**Action limit AL**
	Mean range/d_2	**$D_{WL} \cdot s$**	**$D_{AL} \cdot s$**
2	Mean range/1.128	$2.833 \cdot s$	$3.686 \cdot s$
3	Mean range/1.693	$3.470 \cdot s$	$4.358 \cdot s$
4	Mean range/2.059	$3.818 \cdot s$	$4.698 \cdot s$

*Calculated from

$$D_{WL} = d_2 + \frac{2}{3}(D_{AL} - d_2)$$

Formula originally developed for this handbook

Comments

False out-of-control situations for R and r%-charts.

The action limit (± 3 *s)* for X-chart based on a normal distribution corresponds to a confidence level of 99.73 %; frequency of false out-of-control situations is 0.27 %. Using uncertainty propagation, the factor for calculating the action limit for an R-chart based on duplicates at the same confidence level would be 4.243 ($3 \cdot \sqrt{2}$). However, in the ISO 7870-2 standard for control charts [11], the factor given is 3.686, which, for a normal distribution, corresponds to a confidence level of 99.07 %; frequency of false out-of-control situations is 0.93 %. This is what is normally used and works well.

The warning limits for R and r%-charts calculated with our proposed equation here is with the same confidence level (about 95.5 %) as for X-charts. Based on simulations, using the rule *2 out of 3* between warning and action limits, the percentage of false out-of-control situations becomes 0.38 %. For normally distributed data (method in. control) the total frequency of false out-of-controls for an R-chart is thus 0.93 % + 0.38 % ≈ 1.3 %.

14. Examples

In this chapter we present examples of different control charts from different sectors. All examples are data collected in the authors' laboratories.

In Examples 1 and Example 2 the control limits are based on requirements on measurement uncertainty and limit of quantification (LOQ).

How to proceed when no stable test samples are available is presented in Example 3.

In Example 4 the standard deviation of the control sample is used for setting the statistical control limits, while in Example 5 a requirement for maximum deviation from the reference value is used to set the action limit in a chart with only action limits - target control limits.

How to monitor the repeatability at higher concentrations is shown in Example 6. The r%-chart is chosen (and not the R-chart) since at concentrations well above the LOQ the relative standard deviation is approximately constant.

An out-of-control situation, two out of three control values between warning and action limits, on the same side of the central line, is shown in Example 7.

The long-term reviewing of the control limits is described in detail in Example 8.

Monitoring of blank values is shown in Example 9 – note that also negative values are plotted in the X-chart.

Example 10 shows a situation where nearly all values are above the central line. All values are in the green area and no out-of-control situations are observed.

Example 11 shows a chart with only action limits for monitoring bias of a micropipette.

In Example 12 pooling of standard deviation in order to obtain s_r and s_{Rw} from internal control data is described.

For setting up new QC charts there is a free Excel program, thanks to Michael Koch, available from https://qca-koch.de/en/excel-arbeitshilfen/.

Example 1

Determination of Ni in low-alloy steel with X-Ray Fluorescence (XRF)

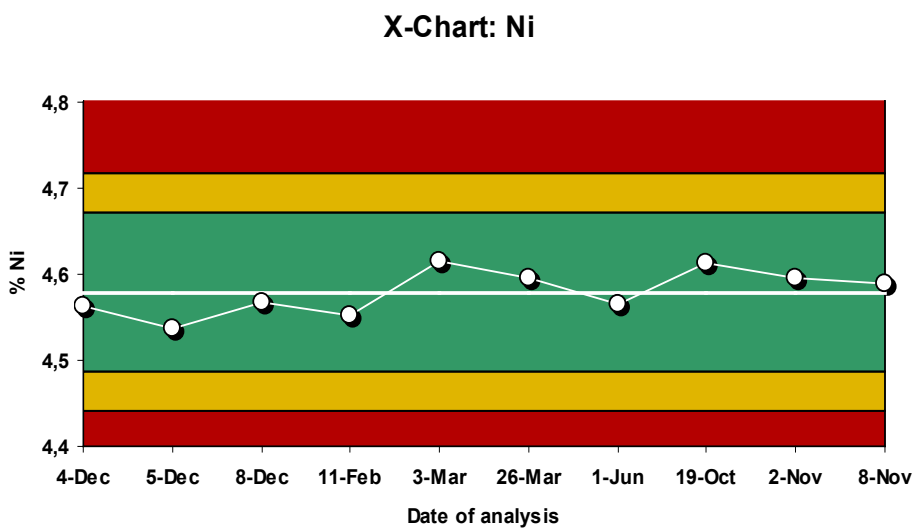
Sample type	Control chart	Control limits	Central line
Steel sample – test sample	X-chart	Target	Mean value

High concentration of nickel. The mean value for our control values over one year is 4.58 weight-%[20] with a standard deviation of 0.026 weight-%. The control sample is taken through the whole measurement procedure (polishing and measurement).

The requirement regarding expanded measurement uncertainty[21] (U) is 4 % (rel). This will be 2 % (rel) as combined standard uncertainty u_c. The requirement of s_{Rw} can normally be set to half or 50 % of the standard uncertainty[22] so we obtain an estimate of the requirement from:

$$s_{Rw} = \frac{u_c}{2} = \frac{U}{4} = \frac{4\%(rel)}{4} = 1\%(rel) \text{ or } 0.0458 \text{ weight-\%}$$

From the requirement on s_{Rw} we calculate the target control limits.



$\bar{x}$ = 4.58 weight-%
s_{target} = 0.0458 weight-%

CL: 4.58 weight-%
WL: 4.58 ± 2·0.0458 = 4.49 and 4.67 weight-%

AL: 4.58 ± 3·0.0458 = 4.44 and 4.72 weight-%

.

[20] The X-chart concentration unit is weight % of nickel, and the demand is given in relative percent of the nickel value (% rel).
[21] Further information on expanded and standard uncertainty is available in the Eurachem/CITAC guide [7].
[22] Due to the way standard deviations are combined this will, with an s_{Rw} of 1 % (rel), allow 1.7 % (rel) for u(bias) for a requirement of expanded uncertainty of 4 % (rel).

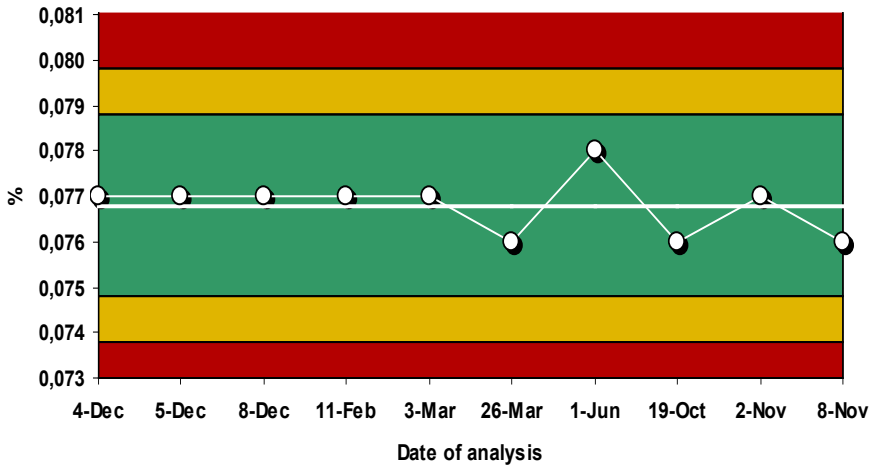
Example 2

Determination of Co in low-alloy steel with XRF

Sample type	Control chart	Control limits	Central line
Steel sample	X-chart	Target	Mean value

Low concentration of cobalt. The mean value for our control values over one year is 0.0768 weight-% with a standard deviation of 0.00063 weight-%. The control sample covers the whole measurement procedure (polishing and measurement).

The requirement for limit of quantification, LOQ, is 0.01 weight-% and this is normally set to 6 to 10 times the standard deviation of a blank or a sample at low concentration. This will require 0.001 weight-% as a standard deviation and this value can be used to set the control limits:



$\bar{x}$ = 0.0768 weight-%
s_{target} = 0.001 weight-%

CL: 0.0768 weight-%
WL: 0.0768 ± 2·0.001 = 0.0748 and 0.0788 weight-%

AL: 0.0768 ± 3·0.001 = 0.0738 and 0.0798 weight-%

Comment
The concentration of the control sample is 8 times the LOQ. In this case this reflects the concentration of interest and is therefore suitable.

Example 3

Determination of N-NH_4 in water with indophenol blue method

Sample type	Control chart	Control limits	Central line
Standard solution	**X-chart**	**Statistical**	**Mean value**
Low test samples	**R –chart**	**Statistical**	**Mean range value**

Low concentration (20 µg/L) in a synthetic solution. $(NH_4)_2SO_4$ was used for preparation of the stock solution of 100 mg/L, and from this the control sample for the X-chart was prepared. The stock solution was different from the solution used for preparation of the calibration standards (which is prepared from NH_4Cl). For the R-chart the sample is one test sample with a concentration < 100 *µg/L* selected among the samples to be analysed in that analytical run.

The X-chart and R-chart were established as follows:

- The mean value of the duplicates was used for plotting the X-chart and the mean value of all results was used as the central line (CL). The standard deviation of the mean values was used for calculating the control limits.
- The range value of the duplicates was used for plotting the R-chart. The mean range was used as the central line (CL). The pooled repeatability standard deviation for the duplicates was used for calculating the control limits.

$\bar{x}$ = 19.99 µg/L s = 0.521 µg/L CL: 19.99 µg/L WL: 19.99 ± 2·0.521 = 19.99 ± 1.04 µg/L (18.95 and 21.03 µg/L) AL: 19.99 ± 3·0.521 = 19.99 ± 1.56 µg/L (18.43 and 21.55 µg/L)	$\bar{x}$ = 0.559 µg/L (mean range) s = 0.496 µg/L CL: 0.559 µg/L WL: 2.83·0.496 = 1.40 µg/L AL: 3.69·0.496 = 1.83 µg/L

Comment
On the X-chart the mean value was not statistically different from the calculated concentration 20 µg/L – no systematic effects were observed for the analyses. There were no results that exceeded the control limits (Chapter 9). On the R-chart there was one control value that exceeded the action limit. The control sample as well as the test samples were reanalysed on 10 Dec with positive outcome.

Example 4

Determination of Pb in water with ICP-MS

Sample type	Control chart	Control limits	Central line
In-house lake water	**X-chart**	**Statistical**	**Mean value**

Low concentration of Pb (0.29 µg/L) in an in-house material. The control sample was prepared from lake water for analysis of low concentrations of Pb (< 1 µg/L). The sample was preserved with HNO_3. The control was performed once in each analytical run.

The X-chart was established as follows:

- the individual results were used for plotting the X-chart;
- the mean value of all results was used as the central line (CL);
- the standard deviation of the control values was used for calculating the control limits.

$\bar{x}$ = 0.294 µg/L
s = 0.008 µg/L

CL: 0.294 µg/L
WL: 0.294 ± 2·0.008 =
0.294 ± 0.016 µg/L
(0.278 µg/L and 0.310 µg/L)

AL: 0.294 ± 3·0.008 =
0.294 ± 0.024 µg/L
(0.270 µg/L and 0.318 µg/L)

Comment
The control values were within the limits. There are 12 consecutive results above the central line. This is out of statistical control but as described in Chapter 9 not out-of-control and can be regarded as acceptable.

Example 5

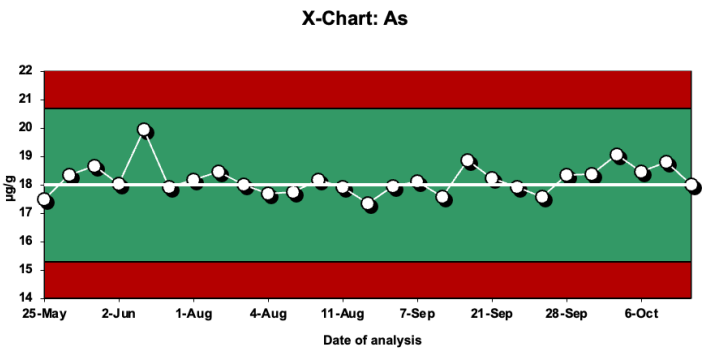
Determination of As in biological material with ICP-MS

Sample type	Control chart	Control limits	Central line
CRM	X-chart with only action limits	Target	Reference value

High concentration of As (18 µg/g) in the CRM (Dogfish muscle NRC/DORM-2). The control sample was used for the determination of As in biological material. The control sample was analysed once in each run. Maximum deviation from the reference value was set to 15 % - the tolerance. Here the laboratory chose to use a chart with only action limits for monitoring the deviation from the reference value.

The X-chart with only action limits was established as follows:

- the individual results were used for plotting in the X-chart
- the reference value was used as the central line (CL);
- the tolerance of ± 15 % was used for setting the action limits.



Reference value = 18.0 µg/g
Tolerance ± 15 % or 2.7 µg/g

CL: 18.0 µg/g

AL: 18.0 ± 2.7 µg/g
(15.3 µg/g and 20.7 µg/g)

Comment
All control values are within the tolerance of ± 15 %.

Example 6

Determination of total P in water using spectrophotometric method

Sample type	Control chart	Control limits	Central line
Test samples	**r%-chart**	**Statistical**	**Mean relative range**

Test samples (>10 µg/L). According to method validation the detection limit (3 *s*) was 2 µg/L. In each run one test sample was analysed in duplicates. The results were applied for r%-charting.

The r%-chart was established as follows:

- the absolute difference of duplicates as percent of the mean value was used for plotting;
- the mean of the r%-values was used as the central line (CL);
- the pooled relative standard deviation, CV, of the r%-values was used for calculating the control limits.

$\bar{x}$ = 1.88 %
CV = 1.67 %

CL = 1.88 %
WL = 2.83 · 1.67 %
= 4.73 %
AL = 3.69 · 1.67 %
= 6.16 %

Comment
In the r%-chart two control values exceeded the control limit. In one instance the action limit was exceeded, and the repeatability was out-of-control. New measurement of duplicates showed repeatability within control. The test samples were reanalysed.

Note – for lower concentrations close to the reporting limit the R-chart is recommended.

Example 7

Determination of β-HCH in biological material with Gas Chromatography with EC (Electron Capture) Detector

Sample type	Control chart	Control limits	Central line
CRM	**X-chart**	**Target**	**Reference value**

The CRM used was cod liver oil BCR/598 with β -HCH (16 µg/kg). The CRM was used for analysis of β-hexachlorocyclohexane in biological material. This control sample was analysed once in each run.

The X-chart was established as follows:

- the individual results were used for plotting the X-chart;
- the reference value was used as the central line (CL);
- the target standard deviation of 15 % was used to calculate the control limits.

Reference value = 16.0 µg/kg
$U = 3$ µg/kg

$s_{target} = 0.15 \cdot 16.0 = 2.4$ µg/kg

CL: 16.0 µg/kg

WL: $16.0 \pm 2 \cdot 2.4$
$= 16.0 \pm 4.8$ µg/kg
(11.2 µg/kg and 20.8 µg/kg)

AL: $16.0 \pm 3 \cdot 2.4$
$= 16.0 \pm 7.2$ µg/kg
(8.8 µg/kg and 23.2 µg/kg)

Comment
A trend was detectable in the results: From September 18 (point number 16 in the graph) results were above the CL and once two control values out of three were above the warning limit. This time (about 1st of January) the analyses were out-of-control and an action was taken.

Example 8

Determination of Cu in water with ICP-OES

Sample type	Control chart	Control limits	Central line
Synthetic standard	**X-chart**	**Statistical**	**Mean value**
Test samples	**R-charts**	**Statistical**	**Mean range**

Synthetic standard (1.00 ± 0.02 mg/L). The control sample was prepared from a commercial standard. The sample was preserved with HNO_3. Test samples and the control sample were measured in duplicates.

X- and R-charts were established and **preliminary** control limits and central line were estimated from the first 60 analytical runs. Now more results are available three months back.

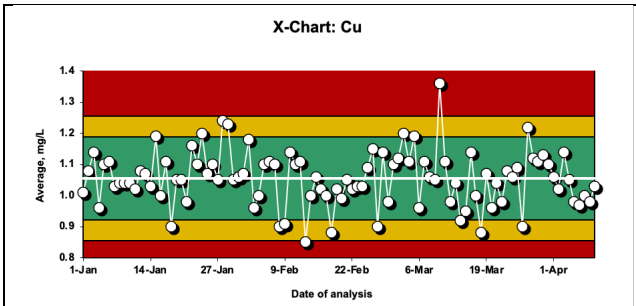
X-chart:

- the average of the two results for the control sample in each run was plotted;
- the mean value was used as the central line (CL);
- the standard deviation was used for calculating the control limits.

R-chart:

- the range for duplicates (highest value minus lowest value) was used for plotting;
- the mean range was used as the central line;
- the pooled repeatability standard deviation (s_r) from duplicates was used to establish control limits by multiplication with factors D_{WL} and D_{AL} (Chapter 13, Table 4).

The control charts were established, and analyses were continued.



$\bar{x}$ = 1.055 mg/L and s = 0.0667 mg/L Preliminary values CL: 1.055 mg/L WL: 1.055 ± 2·0.0667 mg/L (0.92 and 1.19 mg/L) AL: 1.055 ± 3·0.0667 mg/L (0.85 and 1.26 mg/L)	Mean Range = 0.11 mg/L Preliminary values CL: 0.11 mg/L and s_r = 0.0975 WL: 2.833·0.0975 = 0.28 mg/L AL: 3.686·0.0975 = 0.36 mg/L

Review of the data

It is now time for the review of the control charts. As described in Chapter 10 we look at the last 60 data. These are the data plotted since 9 February.

We count the number of times that the control values were outside the warning limits **since** 9 February. On the X-chart we find four cases where the upper warning limit is exceeded, one of these even outside the action limit, and seven cases below the lower warning limit. This makes a total of 11 incidents when the warning limits were exceeded since 9 February. There is thus reason to change the preliminary control limits.

On the R-chart we find five control values outside the warning limit where two of them also are outside the action limit. This indicates that the repeatability may have changed.

One control value on the X-chart on 11 March was outside the upper action limit. On this date the results of test analyses were rejected, and the test samples were afterwards re-analysed. This control value is regarded as an outlier in the review since it differs from the central line by more than 4 standard deviations; see discussion on outliers in Chapter 10. We have therefore excluded this point from the data review.

We calculate a new average and standard deviation, s_{Rw}, from the last 59 points on the X-chart (only 59 since an outlier has been excluded). From the R-chart we calculate a new standard deviation, s_r, for the last 60 points.

New $\bar{x}$ = 1.041 mg/L and new s_{Rw} = 0.0834 mg/L	New repeatability standard deviation s_r = 0.0957 mg/L

X-chart

We compare the new standard deviation to the original standard deviation using an *F*-test:

$$F = 0.0834^2 / 0.0667^2 = 1{,}563$$

The *s* values have 59 and 58 degrees of freedom since they are based on 60 and 59 data points.

In Chapter 13, *Table 3* we cannot find 58 or 59 degrees of freedom, but we can find 60. Since the difference between the values in the table for 40 and 60 degrees of freedom is small, we do not bother to interpolate. Using 60 degrees of freedom for df_1 (new *s*) and df_2 (original *s*) we find that the critical value for *F* is 1.67. This is larger than our calculated value for F (1.563) and therefore the new *s* is not significantly higher than the original value for *s*. However, this *F* value is close to the critical value as would be expected from the number of times that the warning limits are exceeded (11 times with 60 data points) and we recommend recalculating the control limits based on all the data. This will result in slightly higher *s* to be used for calculating the control limits. It is always good to have well determined control limits based on as long a period as possible, preferably over a year.

We will now investigate if the central line has changed significantly. This we do using a *t*-test. The equation in Chapter 12 is:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p} \cdot \sqrt{\frac{n_1 \cdot n_2}{(n_1 + n_2)}}$$

This equation uses s_p, which is the pooled standard deviation for the two sets of data giving the original and the new mean value. The equation for calculation of s_p is given in Chapter 12:

$$s_p = \sqrt{\frac{(n_1-1)\cdot s_1{}^2 + (n_2-1)\cdot s_2{}^2}{(60+59-2)}} = \sqrt{\frac{(60-1)\cdot 0.0667^2 + (59-1)\cdot 0.083^2}{(60+59-2)}} = 0.07545 \text{ mg/L}$$

Since s_p is now based on both sets of data it has 59 + 58 = 117 degrees of freedom.

$$t = \frac{|1.055 - 1.041|}{0.07545} \cdot \sqrt{\frac{60 \cdot 59}{(60 + 59)}} = 1.012$$

In Chapter 13, *Table 2* we find the critical value for the *t*-test at 95 % confidence level. The critical value is the same for 100 and 120 degrees of freedom and therefore also for 117 degrees of freedom: 1.98. The calculated t-value in our test is small compared to the critical value and therefore we see no significant difference between the central line (original mean value) and the mean for the last 60 data points. However, even if there is no significant difference we can calculate more reliable control limits and central line for the X-chart based on all available data.

Previous preliminary X-chart	New X-chart based on longer time period
$\bar{x}$ = 1.055 mg/L and s = 0.0667 mg/L	$\bar{x}$ = 1.048 mg/L and s = 0.0822 mg/L
CL: 1.055 mg/L	CL: 1.048 mg/L
WL: 1.055 ± 2·0.0667 mg/L (0.92 and 1.19 mg/L)	WL: 1.048 ± 2·0.0822 mg/L (0.88 and 1.21 mg/L)
AL: 1.055 ± 3·0.0667 mg/L (0.85 and 1.26 mg/L)	AL: 1.048 ± 3·0.0822 mg/L (0.80 and 1.29 mg/L)

R-chart

We compare repeatability standard deviations by using the *F*-test:

$$F = {s_1}^2 / {s_2}^2 = {0.0975^2} / {0.0957^2} = 1.037$$

The critical value for *F* from *Table 3* in Chapter 13 is 1.67. This is larger than our calculated value for *F* and therefore the repeatability standard deviation has not changed significantly, but we still recommend recalculating the control limits based on all the data.

NOTE The repeatability standard deviation, s_r, is here higher than the within-lab reproducibility s_{Rw}. This can be explained by that the control values for the X-chart are based on a mean of two results.

Conclusion

These results show that the spread and bias of the analyses have not changed *significantly.* We have taken advantage of the larger data set to calculate new and more reliable control limits.

Here is a +5 % bias for a standard solution (1.00 ± 0.02 mg/L) at relatively high concentrations. If this bias is regarded as important, we recommend investigating this and to change the method to minimise the bias.

Example 9

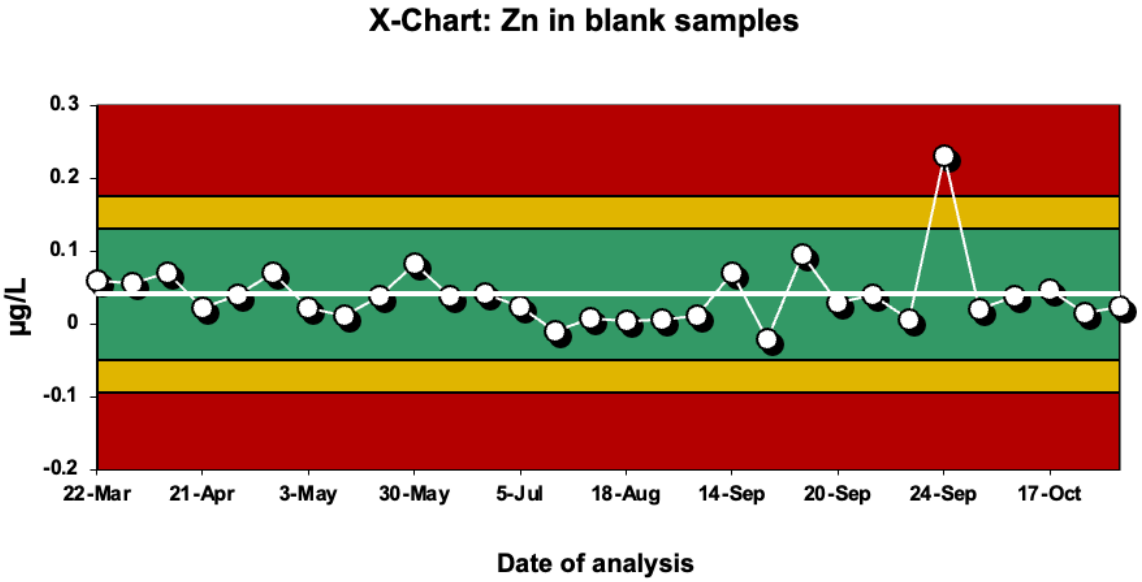
Determination of Zn in hydrogen peroxide with ICP-OES – procedural blank

Sample type	Control chart	Control limits	Central line
Procedural blank	**X- chart**	**Statistical**	**Mean value**

Procedural blank of ultrapure water. The procedural blank [19] determinations were carried out to check for contamination; following the whole procedure using ultrapure water as a sample. In the procedure 50 ml H_2O_2 is evaporated to near dryness, 0.5 ml conc. HCl is added and diluted with pure water to 5 ml and analysed with ICP-OES.

The X-chart was established as follows:

- the single value of the control sample in each run was plotted;
- the mean value of the results was used as the central line (CL);
- the standard deviation was used for calculating the control limits.



$\bar{x}$ = 0.039 mg/L and s = 0.045 mg/L
CL: 0.039 mg/L
WL: 0.039 ± 2·0.045: –0.051 mg/L and 0.129 mg/L
AL: 0.039 ± 3·0.045: –0.096 mg/L and 0.174 mg/L

Comment

There was one result (24-Sep) that exceeded the action limit. Test samples and control samples were reanalysed the following day. Note also that all control values, even the negative ones, are plotted.

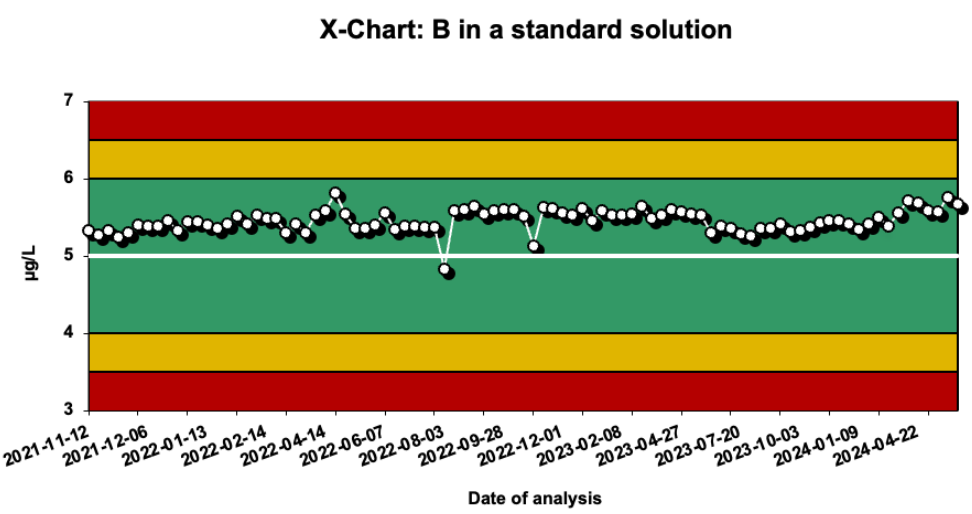
Example 10

Safety of toy material - determination of B with ICP-OES

Sample type	Control chart	Control limits	Central line
In-house synthetic standard	X-chart	Target	Reference value

Migration of elements from toys are determined according to EN 71-3 *Safety of toys - Part 3: Migration of certain elements* [20]. No control sample or reference material to take through the whole procedure is available. A control sample was prepared from a commercial standard. Duplicate determinations of test samples monitor the repeatability.

For monitoring drift an X-chart is used with target control limits. The reproducibility given in the standard for boron is 15 % (Table C1 in EN 71-3). The s_{target} for the chart is set arbitrary to 10 %, i.e. 2/3 of the reproducibility.[23] Here the X-chart for boron is presented.



Reference value = 5.0 µg/L

s_{target} = 10 % or 0.5 µg/L

CL: 5.0 µg/L

WL: 5.0 ± 1 µg/L (4 and 6 µg/L)

AL: 5.0 ± 1.5 µg/L (3.5 and 6.5 µg/L)

Comment

The target limits are here based on the reproducibility given in the EN 71-3 standard. The bias of about +6 % relative (mean value 5.3 µg/L) is regarded as acceptable when the required measurement uncertainty is about 30 % (the double of the reproducibility s_R)

[23] The s_{Rw} is somewhere between the repeatability and the reproducibility standard deviation as stated in Chapter 2.

Example 11

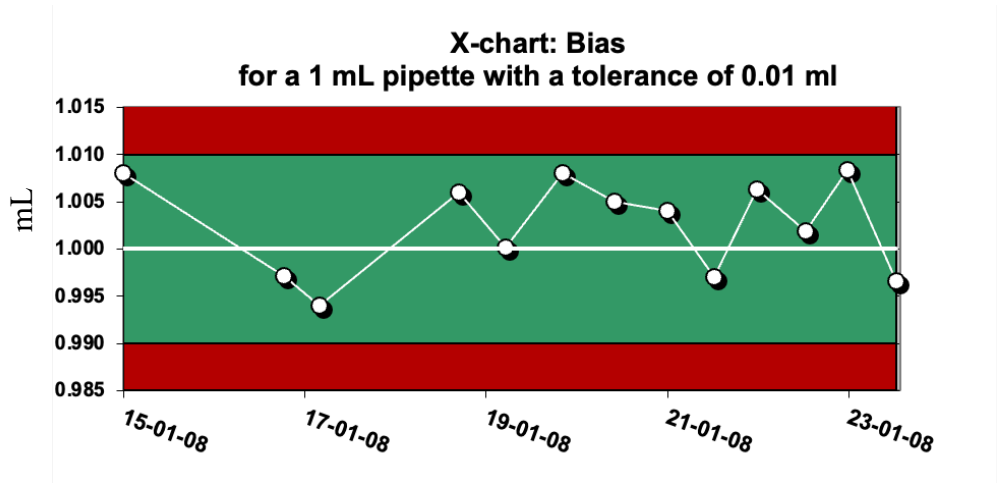
Controlling bias for a 1 mL micropipette

Object	Control chart	Control limits	Central line
Micropipette – 1 mL	**X-chart with only action limits**	**Target**	**Nominal value**

A fixed 1 mL micropipette with a tolerance of ± 1 %. The tolerance is set by the laboratory. The testing of the pipette is performed 10 times according to Nordtest 626 [21]. The mean volume, *V* is calculated and the bias, *b* is calculated as the difference from the nominal volume of 1 mL. This control value, the bias, is plotted in the X-chart with only action limits.

The X-chart was established as follows:

- the nominal value of 1 mL was used as the central line (CL);
- the tolerance of 0.01 mL (1 %) was used for setting the action limits (AL)



Nominal value = 1.00 mL
Tolerance ± 0.01 mL

CL: 1.00 mL

AL: 1.00 ± 0.01 mL

Comment
The micropipette has performed within the tolerance for over an eight-year period from 2015 to 2023.

Example 12

Pooling of standard deviation for s_r and s_{Rw} from internal control

In this example three replicates of a QC sample are measured every day on eight different days. If all results are used to calculate s_{Rw} the estimate obtained will be too low resulting in control limits that are too narrow. Below is shown how to pool the standard deviations within the same day, repeatability s_r, and a simplified way (see note below) to pool between days, within-lab reproducibility s_{Rw}.

Measurement				Day #				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
First	7.1	6.9	6.6	6.7	7	7.3	7.1	7
Second	7.1	6.7	6.5	6.5	6.9	7.4	7.1	6.5
Third	7	6.8	6.9	6.6	6.6	7.3	6.9	6.8
Repeatability								
s	0.058	0.100	0.208	0.100	0.208	0.058	0.115	0.252
s^2	0.003	0.010	0.043	0.010	0.043	0.003	0.013	0.063
s_r	**0.15**							

Within-lab reproducibility		
	s	s^2
	0.226	0.051
	0.342	0.117
	0.226	0.051
s_{Rw}		**0.27**

Repeatability
Calculate s and s^2 for each day for the three measurements. The first day gives $s = 0.058$ and $s^2 = 0.003$. Pool the eight standard deviations using equation 10. Since every standard deviation is obtained from measurements *the same day and analytical run* the pooled estimate, 0.15 is the s_r.

Within-lab reproducibility
Calculate s and s^2 for each measurement day 1-8. The first measurement gives $s = 0.226$ and $s^2 = 0.051$. Pool the three standard deviations using equation 10. Since every standard deviation is obtained from measurements on *different* days the pooled estimate, 0.27 is the s_{Rw}.

NOTE – A simplified way to estimate s_{Rw} is shown here giving $s_{Rw} = 0.270$. The correct estimate using ANOVA (Analysis of Variance) is 0.272.[24] More info on ANOVA can be found in the Eurachem Guide *Fitness for Purpose of Analytical methods* [16].

[24] From ANOVA the repeatability s_r is combined with the between group variation s_i to obtain the within laboratory standard deviation; $s_{Rw} = \sqrt{(s_r^2 + s_I^2)} = \sqrt{0.01541^2 + 0.02246^2} = 0.0272$

15. References

1. V. B. Jensen, K. Haapala, H. Hovind and Ö. Lindgren, *Handbook of internal quality control in water laboratories*. Copenhagen: Water Quality Institute, 1984.
2. B. Magnusson, H. Hovind, M. Krysell, U Lund and I Mäkinen, *Handbook of Internal Quality Control,* Nordtest 2005.
3. ISO/IEC 17025:2017 – *General requirements for the competence of testing and calibration laboratories*, Geneva: International Organization for Standardization.
4. B. Magnusson, T. Näykki, H. Hovind, M. Krysell and E. Sahlin, *Handbook for calculation of measurement uncertainty in environmental laboratories*, Nordtest Report, TR 537 (ed. 4) 2017. Available from www.nordtest.info.
5. E. Hund, D. Luc Massart, J. Smeyers-Verbeke. *Operational definition of uncertainty*, TRAC 2001; 20(8):394-406.
6. ISO 11352:2025 – *Water quality — Estimation of measurement uncertainty based on validation and quality control data.* Geneva: International Organization for Standardization, 2025.
7. S L R Ellison and A Williams Eurachem/CITAC Working Group. *Quantifying Uncertainty in Analytical Measurement*. Eurachem/CITAC Guide, 3rd edition 2012. Available from www.eurachem.org.
8. ISO/TS 13530:2009. *Water quality — Guidance on analytical quality control for chemical and physicochemical water analysis. Guide to analytical quality control for water analysis*. Geneva: International Organization for Standardization, 2009.
9. European Union Council Directive 2020/2184/EC of 16 December 2020 on *the quality of water intended for human consumption.*
10. ISO 7870-1:2007. *Control Charts – Part 1: General Guidelines*. Geneva: International Organization for Standardization, 2007.
11. ISO 7870-2:2023. Part 2: Shewhart control charts. Geneva: International Organization for Standardization, 2023.
12. ISO 7873. *Control Charts for arithmetic average with warning limits*. Geneva: International Organization for Standardization, 1993.
13. J. N. Miller, J. C Miller and R.D Miller, *Statistics and chemometrics for analytical chemistry*. 7 ed. Pearson Education Limited, 2018.
14. W. Funk, V. Dammann, G and Donnevert, *Quality assurance in analytical chemistry. Applications in Environmental, Food and Materials Analysis, Biotechnology and Medical Engineering*. 2nd revised Edition. Weinheim: Wiley-VCH, 2006.
15. *Requirements on analytical quality*, under *Resources* at www.trollboken.se.
16. H. Cantwell (ed.) Eurachem Guide: *The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics*, (3rd ed. 2025). Available from www.eurachem.org.
17. ISO/IEC Guide 99, *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms*, International Organization for Standardization (ISO)/ International Electrotechnical Commission (IEC), Geneva, 2007. Available as JCGM 200:2008 from www.bipm.org.
18. V. J. Barwick (ed.), *Eurachem Guide: Terminology in Analytical Measurement – Introduction to VIM 3* (2nd ed. 2023). ISBN 978-0-948926-40-2. Available from www.eurachem.org.
19. H. Cantwell (ed.) *Blanks in Method Valid*ation - Supplement to Eurachem Guide the Fitness for Purpose of Analytical Methods, (1st ed. 2019). Available from www.eurachem.org.
20. *EN 71-3 Safety of toys - Part 3: Migration of certain elements.* Geneva: International Organization for Standardization, 2013.
21. M Jönsson and B Magnusson, *In-house calibration and control of piston operated pipettes*, Nordtest Report, 626 (1st ed. 2025). Available from www.nordtest.info.

Blank page

NT TECHN REPORT 569
Edition 6.1 English
Approved 2026-08

Return address:
Nordtest,
Slettetoften 9
DK-2630 Taastrup, DK

NORDTEST

NORDTEST, founded in 1973, is a Nordic institution acting as a joint body in the field competence and expertise in the field of harmonizing of norms and methods, a large Nordic net-work of experts, more than 650 recommended Nordic testing methods and 550 published technical reports.

www.nordtest.info