NORDTEST NT TR 569 edition 5.1 2018:09

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 5.1 September 2018

Authors
Bertil Magnusson, Trollboken AB and RISE Research Institutes of Sweden, Sweden
Håvard Hovind, Norway
Mikael Krysell, Swedish Agency for Marine and Water Management
Ulla Lund, Denmark
Irma Mäkinen, Finland

Drawings by Petter Wang, Norway and Timo Vänni, Finland

Valuable comments on the content have been provided by:
Eskil Sahlin, RISE, Sweden
Håkan Marklund, SNV, Sweden,
Annika Norling, SWEDAC, Sweden
Roger Wellum, IRMM, Belgium
Elizabeth Prichard, United Kingdom
Marina Patriarca, Antonio Menditto and Valeria Patriarca, ISS, Italy
and many, many devoted analytical chemists.

Downloading address:
www.nordtest.info

Recommended citation
This document should be cited* as "B. Magnusson, H. Hovind, M. Krysell, U. Lund and I. Mäkinen, Handbook - Inter Quality control, Nordtest Report TR 569 (ed. 5) 2018. Available from www.nordtest.info.
*Subject to journal requirements

Preface

The aim of the Troll book is to give practical guidance on internal quality control for the analysts in their daily work with the analytical methods. This 5th version of the Handbook is a minor revision. The main updates are:

- more focus on *target control limits.* In cases where the client's demand is lower than the performance of the method, wider control limits can be set. Since this will results in fewer out of control values we recommend laboratories consider this option – see example 1,2,5 and 7;
- long-term evaluation (chapter 10) is revised and now discusses changing of control limits and central line in separate paragraphs;
- pooled standard deviation
  - combined standard deviation is now more correctly called pooled standard deviation;
  - pooled standard deviation is recommended to obtain the standard deviation for range charts;
  - an example of pooling standard deviation for s_r and s_{Rw} from an internal control measuring three replicates in every analytical run, example 10, is added. If all results were used to calculate s_{Rw} too low estimate is obtained resulting in too narrow control limits;
- for range chart the best samples to be used are test samples selected among the samples to be analysed in that analytical run – examples have been updated.

The first version of *Internal Quality Control (1) – Handbook of Internal Quality Control in Water Laboratories* was prepared in Nordic cooperation in the1980s, best known under the name *Trollboken* (2). Later it has been translated to several other languages and has been widely used as a tool in chemical routine laboratories – not only in environmental laboratories.

During the years since the first version was prepared, there have been many developments in the field of analytical quality. First of all, the requirement for accreditation of analytical laboratories has put ~~a~~ 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 both in chemical and microbiological methods. When a laboratory estimates measurement uncertainty for a test method knowledge of the within-laboratory reproducibility (intermediate precision) is essential.

The task of compiling this book has been made possible by the financial support from Nordic Innovation Centre/Nordtest through the project 04038, the Swedish Environmental Protection Agency and Trollboken AB.

This version 5 of the handbook TR569 can be downloaded from www.nordtest.info. Current Frequently Asked Questions (FAQ) about Nordtest guides can be found at www.trollboken.se under menu item *Resources*.

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 the data of internal quality control are used is described in the following two chapters. *Chapter 9* explains the interpretation of quality control data to be performed after every analytical run, whereas *Chapter 10* explains how the quality control programme should be reviewed periodically to investigate if 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 10* explains how information on the within-laboratory reproducibility, bias and repeatability is derived from quality control data. *Chapter 11* gives examples of other uses of quality control data and the principles of control charting.

Chapters 12 and 13 give definitions and useful equations and statistical tables for internal quality control and use of data from control charts.

Chapter 14 contains nine examples illustrating how control charts can be started as well as practical application of the control rules and the yearly review described in Chapters 9 and 10. In example 8 we present a detailed review of preliminary control limits and setting new control limits based on more data. Example 10 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.
Full explanation is given in Chapter 12.

n	Number of measurement values
s	Standard deviation
$\bar{x}$	Mean value
R_w	Within-laboratory reproducibility (Intermediate precision)
CRM	Certified Reference Material
AL	Action Limit
WL	Warning Limit
CL	Central line
QC	Quality Control

CONTENTS

Blank page

1. Introduction

According to ISO/IEC 17025 7.7.1 *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 *7.7.3 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.*

Internal quality control at the chemical 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 outside the limits, no analytical results are reported, and remedial actions must be taken to identify the sources of error, and to remove such errors. Figure 1 illustrates the most common type of control chart, the X-chart.

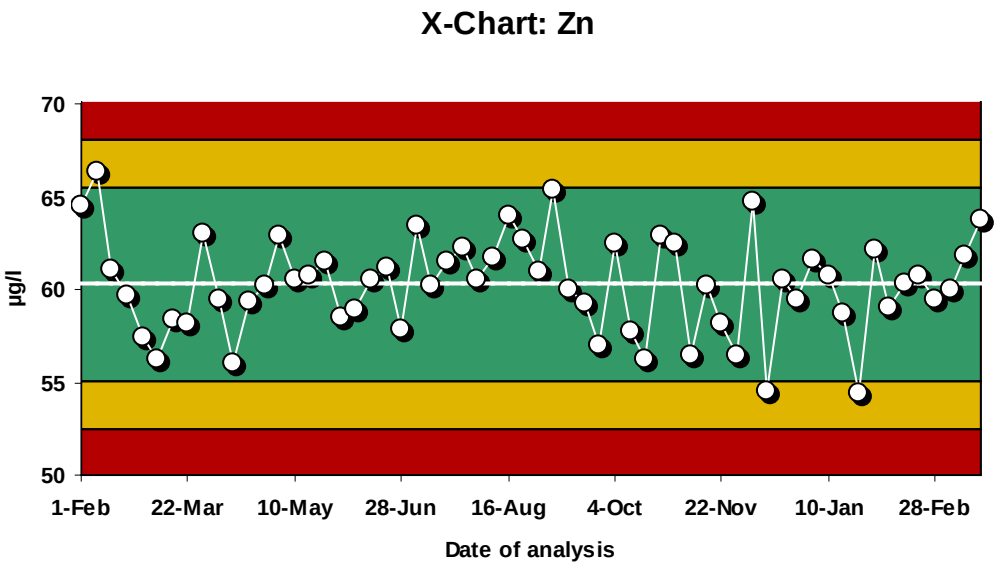


Figure 1. Example of an X control chart for the direct determination of zinc in water. All control values in the green area (within the warning limits) show that the determination of zinc performs within given limits and the test sample results are reported. Control values in the red area (outside the 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 specific rules.

When a quality control (QC) program is established, it is essential to have in mind the **requirement** on the analytical results and for what purposes the analytical results are produced – the concept of *fit 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 program 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 program 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 (5537).

The QC should be part of a quality system and should be formally reviewed on a regular basis. Other important elements of the quality system are the method validation and the participation in proficiency testing.

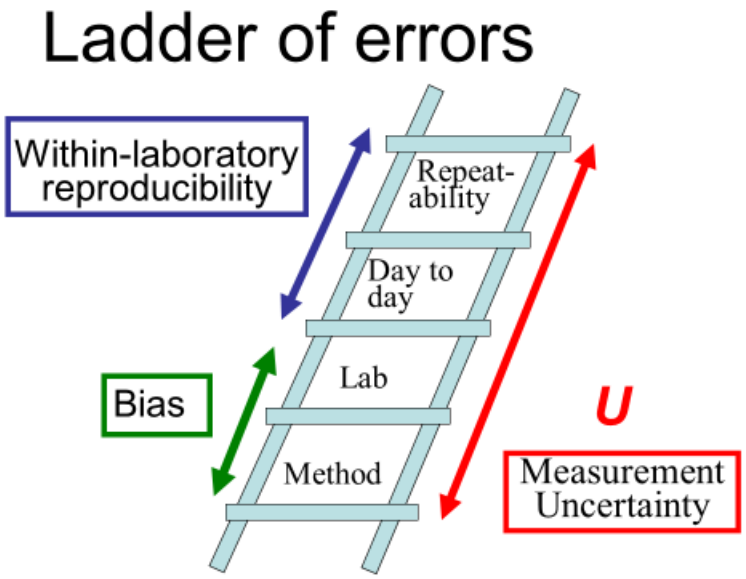
In practical work it is necessary that the quality control is limited to fulfilling the requirements on the analytical results – a good balance between control work and analysis of samples is essential. When the requirements are lower than the performance of the method wider control limits can be set – *target control limits*. The aim of this handbook is to describe a *fit for purpose* system for internal quality control at analytical laboratories that are performing chemical analysis. 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.

Analytical chemists know that 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 (repeatability or reproducibility) or the *measurement uncertainty* that is the important quality parameter. The internal quality control will normally give an indication of the *within-laboratory reproducibility*, 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. The *measurement uncertainty* will tell the customer the possible maximum deviation for a single result[1] from a reference value or from the 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 (4), Figure 2.



o

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

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

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,

[1] or more strictly *the range of possible values with a defined probability associated with a single result*

3) the day-to-day variation in the laboratory, 4) the repeatability of that sample. Each of these steps on the ladder adds its own uncertainty. The *within-laboratory reproducibility*, R_w, 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* normally 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 also different requirements on measurement uncertainty are illustrated with a small and a big green circle. For further reading about measurement uncertainty we recommend the Nordtest report (5) and the Eurachem guide (6).

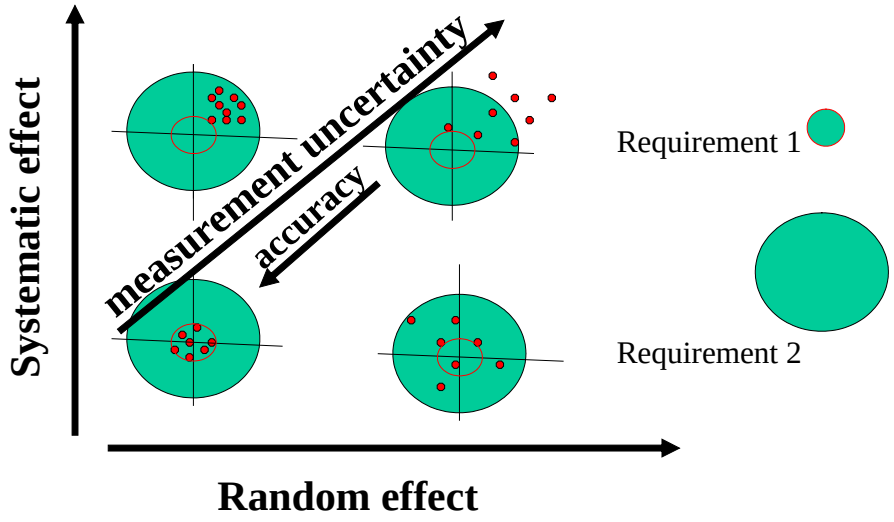


Figure 3. *Random and systematic effects on analytical results and measurement uncertainty may be illustrated by the performance of someone practising aiming at a target – the reference value or true value. Each point represents a reported analytical result. The two circles are illustrating different 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 is representing 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

There is a bias when an average is either greater than the reference value or lower. Variations on bias may occur over time because of changes in instrumental and laboratory conditions. For small changes it is often difficult to say if these effects are random or systematic.
Some typical sources of systematic effects (7):

- 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 effects, as well as if the calibration curve is supposed to be linear in a concentration range where this is not true;
- blank correction too high or too low, if the blank and the sample are different and 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 irregular and uncontrollable variations in the many factors affecting the analytical result can 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 etc.

Table 1. Example of laboratory internal quality control values for a solution containing 60.0 µg/l of zinc. Figure 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

If we analyse a sample several times, we do not obtain a series of identical results. The values are spread within certain limits. The results are varying 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 achieve 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.

A graphical presentation of the results gives a much better understanding of the spread. Figure 4 is 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.

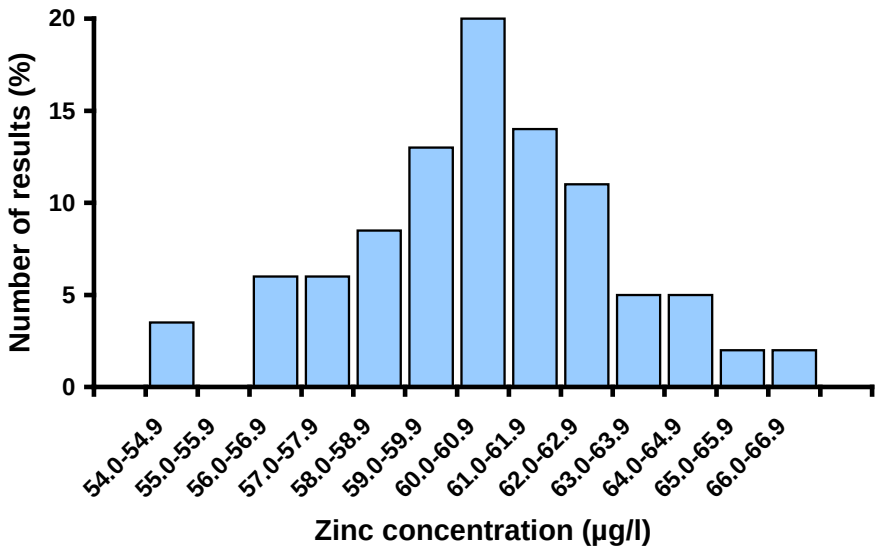


Figure 4. A histogram illustrating the distribution of the control values from Table1. The results are sorted in groups defined by the concentration range. Each group is represented by a column where the height represents the number of results in the group, calculated in percent of the total number of results.

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

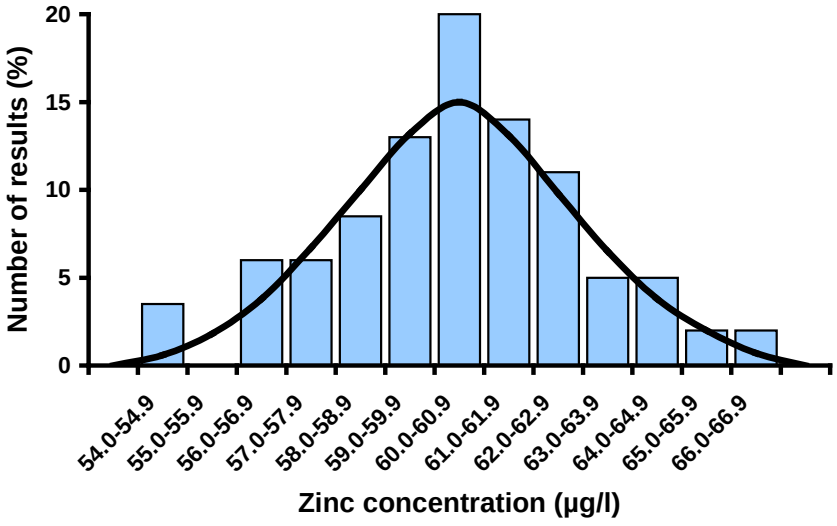


Figure 5. *The relation between the normal distribution curve and the histogram. The distribution curve is based on the same data as represented in the histogram (Figure 4).*

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*.

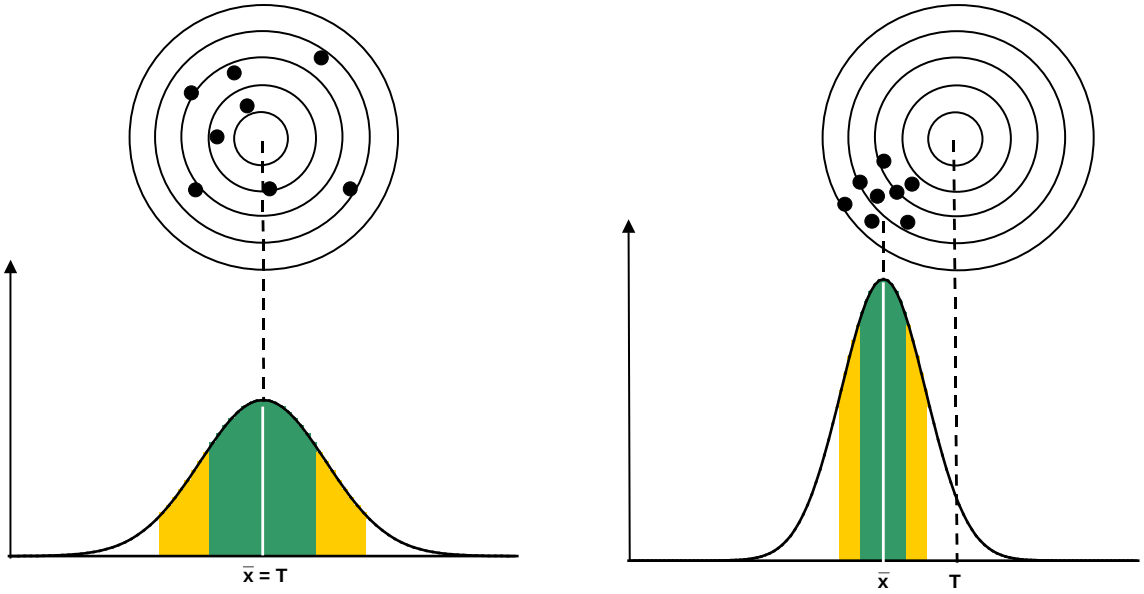


Figure 6. The shape of the normal distribution curve depends on the spread in the measurement results, i.e. within-laboratory reproducibility: A poor reproducibility will give a large standard deviation, and the corresponding curve is broad (left). If the reproducibility is good, the standard deviation 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. 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, see *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. 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, for chemical measurements, will often be ± two times the combined standard uncertainty.

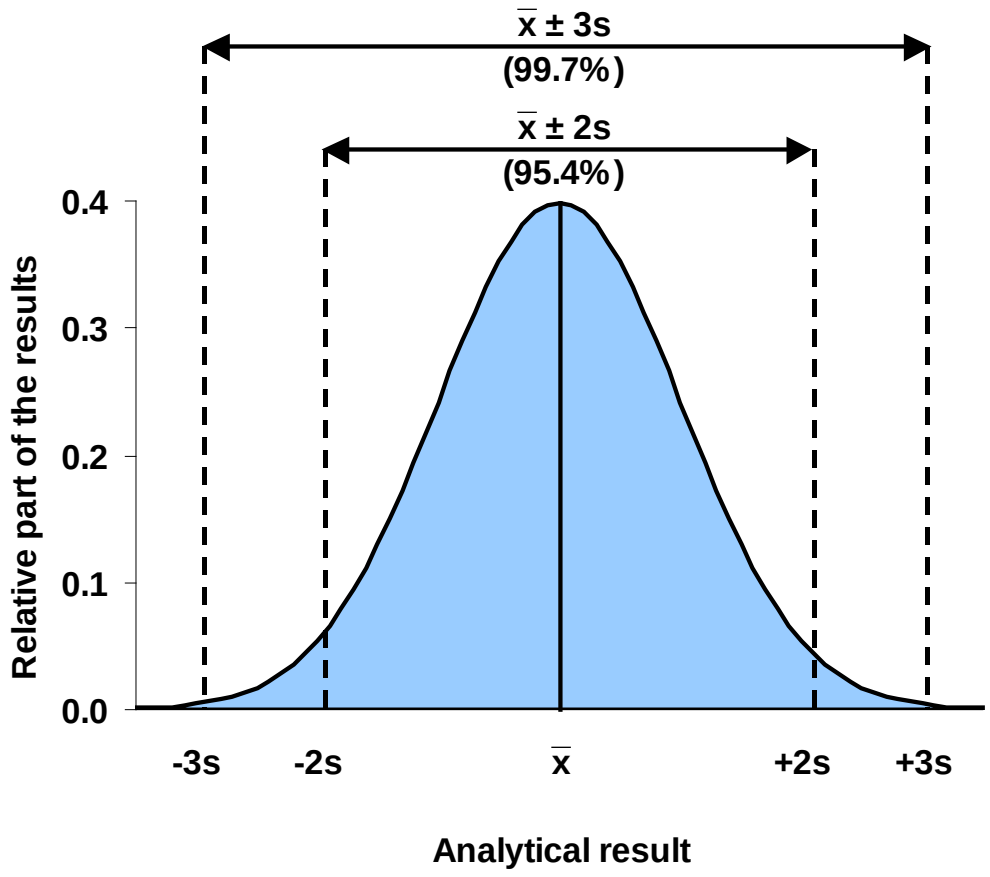


Figure 7. A normal distribution curve illustrating the probability for a result to be located within given limits ($\bar{x}$ is the mean value, s is the standard deviation).

3. Requirement for analytical quality

Here we describe how the analyst can translate the customer's requirement 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 requirements for 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 carried to monito discharges to check whether legally allowable 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, but the measurement uncertainty must be adequate to ensure that the right decision is taken when comparing the result to the allowable concentration limit.

The users of the results expect to be able to trust the data, but in most cases, they do not have the expert knowledge necessary to explain exactly what they need, and they rely on the laboratory to supply the right answer to the problem – that is to deliver a result that is fit for the purpose. It is a challenge for the laboratory to understand the needs of the user. If the laboratory is accredited, the standard ISO/IEC 17025 requires that the laboratory evaluates 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 and therefore have the same requirements for quality. The laboratory therefore does not need to think closely on the subject every day but can design its quality control programme so that the data delivered will have the correct quality for the purpose.

However, the correct quality still needs to be defined. In some cases, national or regional authorities have defined the required quality for regulatory analyses (19). For example, the European drinking water directive 98/83/EC [8] contains requirements for 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 uncertainty remains constant even though concentration in the sample decreases. 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 limiting minimum uncertainty are often described as a proportion (or percentage) of the concentration of primary interest. The "concentration of primary interest" may, e.g., be a water quality limit or a similar allowable 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 that estimate the demand for *within-laboratory reproducibility* to be used in internal quality control.

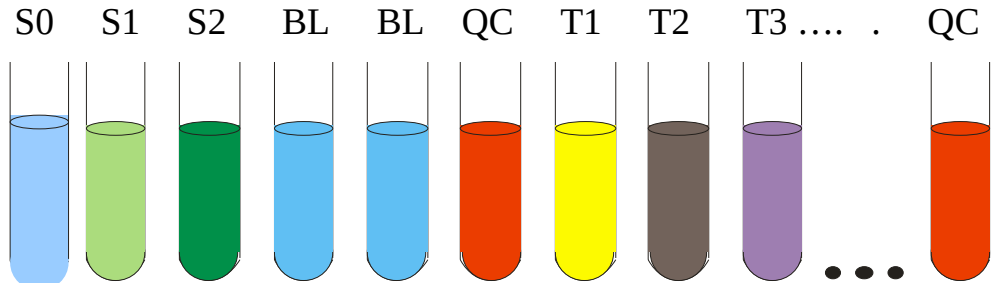
Example:
We are asked to determine ammonium in drinking water. The EU drinking water directive [8] states the required expanded measurement uncertainty at the level of 0.5 mg/l is 40 %. This guide proposes that a first estimate of s_{Rw} is the required U divided by a factor of 4 – see further example 1.

Most laboratories will be able to determine ammonia 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 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 (*Figure 8*). Material of control samples can be standard solutions, test samples, blank samples (20), 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.

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

- giving one more significant digit compared to test results;
- report **values** below reporting limit (LOQ);
- report negative **values.**

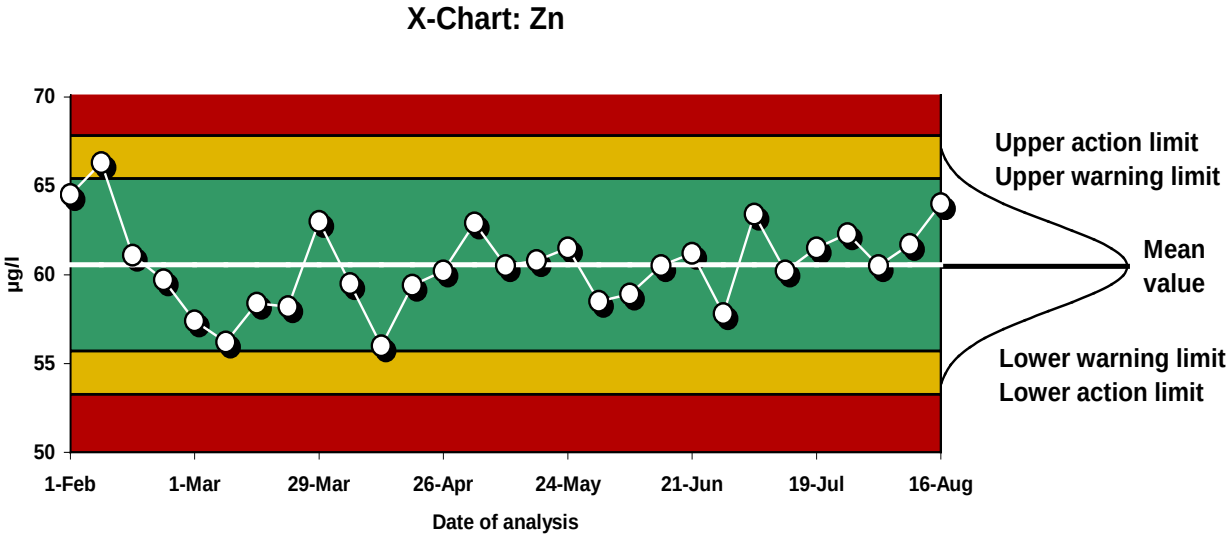


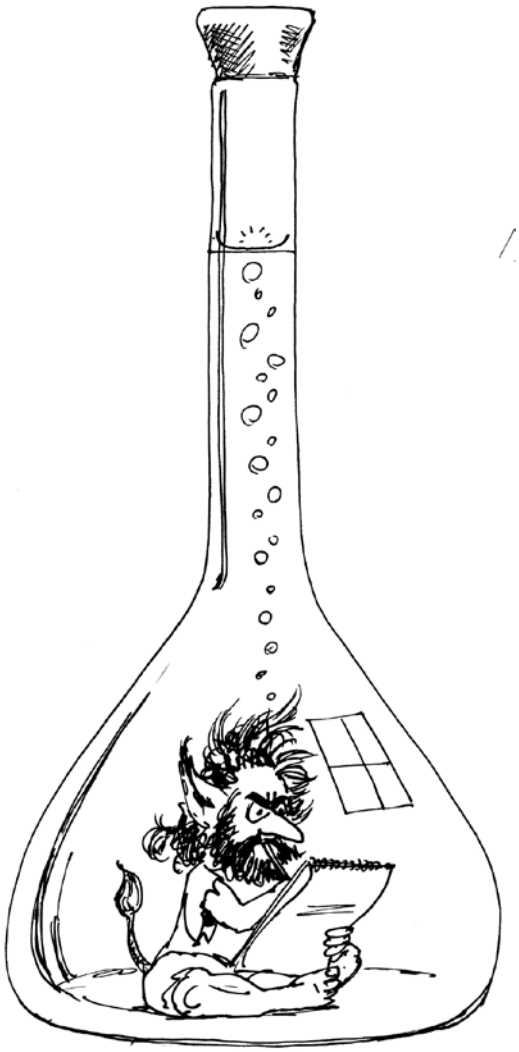
Figure 9. The relation between the normal distribution curve and the control chart. The central line is either a mean value or a reference value.

The chart is based on the statistical characteristics of random variations, defined by the normal distribution curve. The relation between the normal distribution curve and the equivalent 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. If the result for the control sample is outside the action limits, there is a high probability that the analysis is in error.

The warning and action limits can be set either as above on method performance, *statistical control limits* or using independent quality criteria based on fitness for purpose – *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 are reported – see Chapter 9.



5. Different types of control charts

This chapter describes the different types control charts, when they will be used, 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:

- X-charts;
- Range-charts, R or r %.

X-charts

An X-chart has 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 (9,10,11,12,13,14,15) 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* is a special application of the X chart based on analysing a sample that can be assumed to contain the analyte at a very low level[2]. It provides special information about contamination of the reagents used, and the state of the measurement system. Even though concentrations are normally entered into the blank value chart, it is also possible to use the value of the measured signal. Remember that both positive and negative control values shall be plotted in the chart.

Another special case is a *recovery chart*. The analytical process may be tested for matrix influences by determining the recovery of additions of standards (spikes) to test samples.

Calibration parameters such as slope and intercept, in so far as they are determined daily, can also be monitored by means of he X chart.

Range charts

A range chart (R, r %) has 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. In 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 single result for two or more separate analyses of the same sample. For practical applications in analytical laboratories the range chart mostly appears only in its simplest form, only duplicate determinations (of samples to be analysed) in each analysis series.

The best samples to be used are test samples selected among the samples to be analysed in that analytical run. However, the concentrations may vary, because the samples are different in every analytical run. The range is normally proportional to sample concentration (at levels well above the reporting limit) and then it will be more 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 R-chart where the control values are the absolute range.

[2] In general, we have several different blanks, e.g. reagent blank, procedural blank and sample blank (18).

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 we recommend that the control value is based on the mean value of duplicated determinations of two different sample aliquots – i.e. the same number of measurements for test samples as for control samples.

6. Different control samples

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

Ideally the control samples should go through the whole measurement procedure. They 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. This is however 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 any systematic effect (bias). Repeated determinations in each analytical run give a possibility of using the standard deviation (or range) as an estimate of the repeatability of the measurement. However, when a 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, and if multiple analyses are performed may be used for range charts. For range charts we generally recommend control sample type IV.

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

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

If the initial validation of the method has proved that the random effects, when analysing control samples, are approximately the same as for test samples, this type of control will provide a direct measure for the within-laboratory reproducibility. However, in most cases the spread of the analytical results of a synthetic and a test sample will not be the same; therefore, a stable real control sample should be chosen whenever possible.

A control sample type II is usually prepared by the laboratory. It can be either stable, homogeneous test samples or synthetic samples. Standard solutions can be bought from external suppliers but are often prepared in-house. 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 to last for several years. Synthetic in-house materials are prepared from pure chemicals and purified solvent (e.g. water) 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 preparation or purity of the chemical will not become evident ~~be seen~~.

This control sample type is suitable for X-charts, and if multiple analyses are performed, also for R-charts.

Control sample type III - blank sample

Control sample type III may be used for the surveillance of the limit of quantitation (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. Errors in the blank cause systematic effects at low concentrations, which can also be detected with control sample type III.

Control sample type III can be the blank sample used for blank correction according to the procedure for analysis.

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

Control sample type IV test (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 (type II) – typical examples are for the determination of dissolved oxygen and chlorophyll *a*. Duplicate measurements give a realistic picture of the within-run random variations for natural samples.

The control sample will generally be selected at random among the test samples submitted for measurement in the laboratory.

If a synthetic sample is used for the X-charts, it could be a good idea to include a control sample type IV, if the repeatability for synthetic and test samples is different.

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

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 most common method to set the limits. An alternative is to start with the analytical requirements or intended use of the results. From the requirement 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 reference [7, 19].

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

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**[3]
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* and – 2 *s*. Action limits will be +3 *s* and – 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}. Warning limits will be +2 *s* and – 2 *s*. Action limits will be +3 *s* and – 3 *s*.

The central line in the control chart can be the calculated mean value of the control values or a reference value for the control sample. In most cases a mean is used as the central line.

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 or a well-characterised material. The central line is set to the nominal value of the material.

In the cases below the control sample is an ideal control sample similar to test samples and subjected to all steps of the analytical procedure and therefore the target s_{Rw} may be used to set the target limits. The examples referred to below are presented in Chapter 14.

Case 1. **Statistical control limits** and a **mean central line -** see also Example 3 and Example 4.
The requirement on *within-laboratory reproducibility* is not set and the method is performing with a s_{Rw} = 6 %. The warning limits are set to two times the method standard deviation, ± 12 % and action limits to three times the standard deviation, ± 18 %. The mean value for the control sample is 59.2 µg/l so ± 12 % is equal to ± 7.1 µg/l and ± 18 % is equal to ± 10.7 µg/l. The warning limits will be at 59.2 ± 7.1 µg/l (52.1 and 66.3 µg/l) and the action limits will be at 59.2 ± 10.7 µg/l (48.5 and 69.9 µg/l).

[3] In the examples we always assume that the number of samples analysed for control values is the same as used for routine measurements. If, however, a control value is based on duplicates (the mean of two response values) and a routine result is based on a single sample, and the major part of the spread is repeatability, the *s* used for setting the limits may have to be reduced.

Case 2. **Statistical control limits** and a **reference central line.**
If the mean value is very close to the nominal or the reference value, statistical control limits can be used otherwise we recommend Case 4.

Case 3. **Target control limits** and a **mean central line –** see also 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 4. **Target control limits and a reference central line** – see also 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-chart or r%-chart

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 can be found in Table 4 in Chapter 13 and the background to these factors is explained in a comment to Table 4.

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, or the *s* is calculated from the mean range. For duplicates (n = 2) the *s* = mean range/1.128.	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.
For duplicates $n = 2$. Central line is the mean range. Upper warning limit will be + 2.83 *s*. Upper action limits will be + 3.69 *s*.	For duplicates $n = 2$. Central line is 1.128 *s*. Upper warning limit will be + 2.83 *s*. Upper action limits will be + 3.69 *s*.

Case 1. **Statistical control limits** – see also Example 3 (R) and Example 6 (r%) in Chapter 14.
The pooled standard deviation is = 0.356 %. The warning limit for the range chart will then be set at + 2.83 · 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 % abs). From this limit the repeatability standard deviation is calculated as s_r = r/2.8= 0.357 %.[4] The warning limit for the range chart will then be set at + 2.83 · 0.357 = 1.0 % and the action limit at 3.69 · 0.357 = 1.3 %.

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

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 estimate if the *s* obtained for the control sample is sufficiently low to fulfil the analytical requirement - see Chapter 3.

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 statistical 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 usually too low because not all variation is taken into account.

Fixed central line – We recommend a fixed central line. To obtain a reliable mean value one-year period may be a good time period. 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 (e.g. the whole process) for test samples we should also in the X-chart plot the mean value of duplicate analyses for the control sample. 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.

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 20 analytes are determined[5] and statistical control limits are used for all analytes, on average one control value (equal to 5 % of the control values) can be expected to be outside the warning limits in each analytical run. Also, in about 1 out of 17 analytical runs a control value for one of the analytes is expected to be outside the action limit, making ordinary interpretation very unpractical.

[5] 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 program

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

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

Setting up the QC can best be described by a practical example: 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 waters 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	Standard value
An in-house material, Cd: 0.10 µg/l (Type II)	X-chart	Target	Mean value
Replicate determinations of test samples in two concentration ranges, (Type IV)	R-chart r%-chart	Target	Target s_r · 1.128

Because of the rather wide concentration range in test samples we have chosen 3 QC samples Type I and II. 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 direct check of systematic effects in our measurement procedure we use the 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 two samples in each analytical run representing two concentration ranges and these samples are analysed as duplicates (two different test tubes in the autosampler).

In 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 system. The results of duplicates, the range, obtained in analysis of test samples are plotted in R-chart at lower concentrations and r%-chart at higher concentration.

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 a long period of time in 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 analyte 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 control samples are analysed at least at the beginning of each run and before finishing the analytical 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 program in a method description and in a quality manual

The principles of the quality control program 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 of each analytical method.

9. Daily interpretation of quality control

In this chapter we describe the interpretation after each analytical run. Can we report the results or not? Is the method out of statistical control?

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. the change of reagents, the 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 three possible cases:

1. the method is in control;
2. the method is in control, but the long-term evaluation shows that the method is *out of statistical control;*
3. the method is out of control.

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

1. the control value is within the warning limits;
2. the control value is between warning and action limit and the two previous control values were within warning limits.

In this case the analyst can report the analytical results.

2. The method is **in control** but can be regarded as **out of statistical control** if all the control values are within the warning limits (maximum one out of the last three between warning and action limit) and if:

3. seven control values in consecutive order gradually increase or decrease (7);

4. 10 out of 11 consecutive control values are lying on the same side of the central line (7).

In this case the analyst **can report the analytical results,** but a problem may be developing. Important trends should be discovered as early as possible to avoid serious problems in the future. Examples of important trends are when most of the control values lie far away from the central line though still within the warning limits. **In other words, each laboratory must document how to treat these trends.**

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 has to decide if this bias is acceptable.

3. The method is **out of control** if:

5. the control value is outside the action limits;
6. the control value is between the warning and the action limit and at least one of the two previous control values is also between warning and action limit – 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 for the control sample was obtained must be reanalysed.

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 different analytical variables cannot be treated exactly in the same manner. 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 an error also in the analyses of test samples.

If there is an out-of-control situation the normal action is to do some more (at least two) control analyses. 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 test analyses shall be stopped, and remedial actions have to be taken to find and eliminate the cause(s) of error.

Controlling the reagents and the calibration of the method or exchange of vessels and apparatus are usual remedial actions in out-of-control cases. The problem, and the solution of this, should be documented. Analyses which have been carried out since the last acceptable control value was obtained must, if possible, be repeated. If the repeated control values still are out-of-control, the results of test samples shall not be reported. 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 decreased reliability is given.

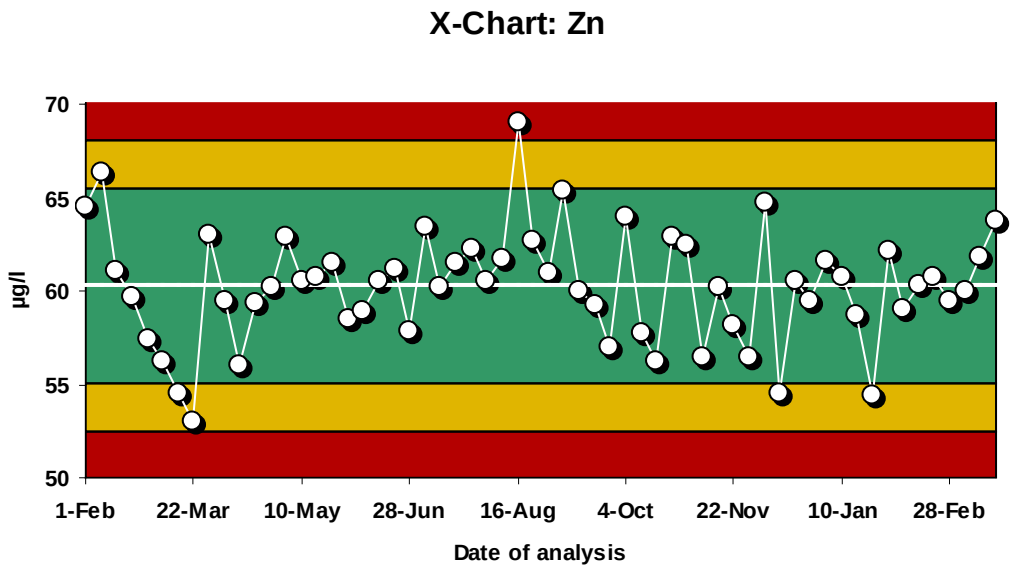


Figure 10. X control chart with two out of control situations.

10. Long-term evaluation of quality control data

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

- *What is the quality (random and systematic effects) currently in the laboratory? Has the quality significantly changed?*
- *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.

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) from last year (7) and compares the current quality with a) the year before and b) with the current control chart. Use all data from last year and check if there is any significant change. There should be at least 60 control values. If less use also data from previous years, but at least 20 control values should be from last year:

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

The F-test and t-test are explained in detail in Example 8.

If the number of data points are about 60 the following simpler check applies:

1. **standard deviation** - If you have a control chart using statistical control limits, then count the number of cases where the results are outside the warning limits. If this number is greater than 6 or less than 1 there is clear evidence (with 60 data points) that the standard deviation has changed (7). In this case perform an F-test;
2. **mean value in an X-chart** - Calculate the mean of the last results and compare with the previous mean value. If the difference is more than 0.35 *s* there is clear evidence (with 60 data points) that the mean value has changed. Then perform a t-test.

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 limit 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.

What makes a change in control limits necessary?

Target control limits are only changed if customers' requirements change. This section is 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 (7) is to exclude data that are more than 4 standard deviations away from the central line and retain the rest. If more than one out of control situation exists in 60 points under consideration, it is more than you would expect and there is good reason to scrutinise the whole analytical procedure to search for the cause of the repeated out of control situations.

What makes a change in central line necessary?

A reference central line is fixed. This section 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. Depending on which type of control chart that is used, a few suggested uses are listed in this chapter.

Measurement uncertainty

Results from the control charts can, together with other data be used for estimating the measurement uncertainty. In most cases, the systematic effect and the random effect (the standard deviation) can be combined to calculate 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 (5) and partly in the Eurachem/CITAC guide (6).

Measurement uncertainty is estimated from control charts results combined with results from proficiency tests, data from 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 (i.e. includes sample work-up such as filtration, concentration steps etc.) the estimate of the measurement uncertainty may be realistic.

Method validation

Normally, a full 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, 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 to the expected (certified) value. With an in-house or purchased RM, a rough estimate of the bias will be given, though with less certainty than when using a CRM.
- All types of control charts will provide information on the spread (random variation) from calculations of standard deviation or from estimates using the range.

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 detection (LOD)

The estimate of limit of detection used by many sectors is a standard deviation multiplied by a factor. The factor is normally 3. For further guidance see reference [18].

Data from an X-chart with a test sample at low concentration will be useful for the estimation of the detection limit 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.

Data from an R-chart will give the repeatability standard deviation, and if the concentration is low, this standard deviation, after correcting for number of measurement and blank determinations, is useful for estimation of the limit of detection [18].

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 allowable systematic effects and spread, then comparing this to what is reached by the experienced trained staff. 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 tests

If 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 over performance, including possible systematic effects or trends.

Here the z-score is plotted in an X-chart. CL = 0, WL = 2 and AL = 3.

$$z = \frac{(x_{lab\ value} - x_{assigned\ value})}{s} \text{ or } zeta = \frac{(x_{lab\ value} - x_{assigned\ value})}{\sqrt{u_{lab}{}^2 + u^2_{assigned\ value}}}$$

Example: The total standard deviation in a proficiency test (all laboratories) was 0.08 mg/kg; your result was 0.12 mg/kg lower than the assigned value. Your *z*-score becomes -1.5. Here we recommend that all values outside warning limits should be investigated. The maximum allowed error from authorities (see also Chapter 3) can also be used to calculate the *z*-score.

Another possibility is the zeta score using your own claimed measurement uncertainty (U_{lab}) where u_{lab} is the combined standard uncertainty.

Environmental parameters and similar checks

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 target control chart for plotting the observed control values. In such cases the ideal, expected, temperature will be used as the central line, and the allowable limits used as action limits. The control charts give a very simple graphical presentation of any trends or unexpected variation that might influence the analyses and therefore might be worth considering.

Similarly, it is useful to plot the results of the frequent 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. Exact definitions for terms used are found in VIM Ref (16) and further explained in the Eurachem Guide (17). Direct quotes from this reference 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** (16). The accuracy is affected by both **systematic** and random effects.

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 (16). The bias is estimated by the difference between the **mean value** of a large number of test results and the accepted reference (*Figure 6*).

Confidence interval
The range about the mea**n value** within which a stated percentage of values would be expected to lie. For example, for a normal distribution, approximately 95 % of values are between ± 2 *s* (*Figure 7*).

Control chart
The principal tool in internal quality control. A chart where 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, blank samples.

Control value
Test result from the internal quality control samples entered in the **control chart**. 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 – 0.07 mg/l in a X-chart could for a test sample be reported < 0.1 mg/l.

Degrees of freedom, df
The number of independent comparisons that may be made between individual results in a 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 with a specified method.

Limit of Quantification
When an analytical result is below this limit it is reported as less than (<). Another term used is Reporting limit.

Measurand
Quantity intended to be measured (16), e.g. the amount of acid-soluble cadmium (the **analyte)** in a fresh-water sample.

Measurement procedure
The detailed description of an analytical method used in a laboratory.

Measurement uncertainty
Non-negative parameter characterizing the dispersion of values being attributed to a measurand, based on the information used (16). 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 (7). This is a practical approach. Another alternative is to use Grubbs test – see statistical textbooks.

Repeatability
Measurement precision under a set of repeatability conditions of measurement (16).
Repeatability condition of measurement refers 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 (16)
Reproducibility conditions 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 period 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 (16).

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** (reported with one more significant figure and not less than) or a value calculated from the test results, e.g. the range. Dependent on the type of **control sample,** maybe only a part of the measurement procedure can be applied to the **control sample**.

Spread
The variation between independent test results obtained under stipulated conditions. The opposite is closeness of agreement between **test results** - also called **precision.**

Systematic error
Component of measurement error that in replicate measurement remains constant or varies in a predictable manner (16).

Trueness
Closeness of agreement between the average of an infinite number of replicates measured values and a reference value (16). Trueness is normally expressed in terms of **bias.**

Equations

Mean value ($\bar{x}$)

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

$$\bar{x} = \frac{\sum x_i}{n} \qquad 1)$$

Standard deviation (s).

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

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

Degrees of freedom, df = n – 1

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

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

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

$$s_r = \frac{Range}{1{,}128} \quad (n = 2) \qquad 4)$$

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

F-test

(see Chapter 13, *Table 3*). Used to evaluate whether the **standard deviations** (s_1 and s_2) from two series of determinations are significantly different:

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

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

(see Chapter 13, *Table 2*). 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):

$$t = \frac{\left|\bar{x} - T\right|}{s} \cdot \sqrt{n} \qquad 6)$$

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

$$t = \frac{\left|x_1 - x_2\right|}{s_p} \cdot \sqrt{\frac{n_1 \cdot n_2}{(n_1 + n_2)}} \qquad 7)$$

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

When the calculated t-value is greater than the critical t-value found in *Table 2*, the difference between the two values is statistically significant. **Degrees of freedom**, **df** = n_1+n_2-2.

Mean ($\bar{x}$) 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:

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

Pooled standard deviation (*s*) 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:

$$s_P = \sqrt{\frac{(n_1 - 1) \cdot s_1^{\,2} + (n_2 - 1) \cdot s_2^{\,2} + \ldots + (n_k - 1) \cdot s_k^{\,2}}{n_{tot} - k}} \qquad 9)$$

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

If *n* is about the same for the different *k* series

$$s_p = \sqrt{\frac{s_1^{\,2} + s_2^{\,2} + \ldots + s_k^{\,2}}{k}} \qquad 10)$$

13. Tables

First table in this section is Table 2. Table 1 you can find on page 5.

Table 2. Critical t-values (2-sided test). Normally 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 df from 4 to 120.

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 numerator (s_1^2), df_2 = degrees of freedom in denominator (s_2^2), $s_1 > s_2$

Table 4. Factors for calculation of central line, warning and action limits for construction of R-charts (11). Factors obtained from ISO 8258.

Number of replicates	**Standard deviation, *s***	**Warning limit WL***	**Action limit AL**
	Mean range/d_2	**D_{WL} • *s***	**D_{AL} • *s***
2	Mean range/1.128	2.83 • *s*	3.69 • *s*
3	Mean range/1.693	3.47 • *s*	4.36 • *s*
4	Mean range/2.059	3.82 • *s*	4.70 • *s*

*Calculated from

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

Formula originally developed for this handbook

Comments

Confidence levels for the control limits in X and R-charts

The action limit (± 3 *s)* for X-chart is for a normal distribution with a confidence level of 99.73 %. Using uncertainty propagation, the action limit for R-chart based on duplicates at the same confidence level would be 4.25 ($\pm\ 3 \cdot \sqrt{2} = 4.25$). However, in the ISO standard 8258 for control charts (11) the factor given is 3.686, which corresponds to a confidence level of 99.1 % for a normal distribution. This is what is normally used and works well.

The warning limits for R-charts calculated with our proposed equation here is with the same confidence level (about 95.5 %) as for X-charts.

14. Examples

In this Chapter gives examples of different control charts from different sectors. All examples are data collected in the authors' laboratories. The yearly reviewing of the control limits is described in detail in example 8. In example 10 pooling of standard deviation for s_r and s_{Rw} from internal control is described.

Example 1

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

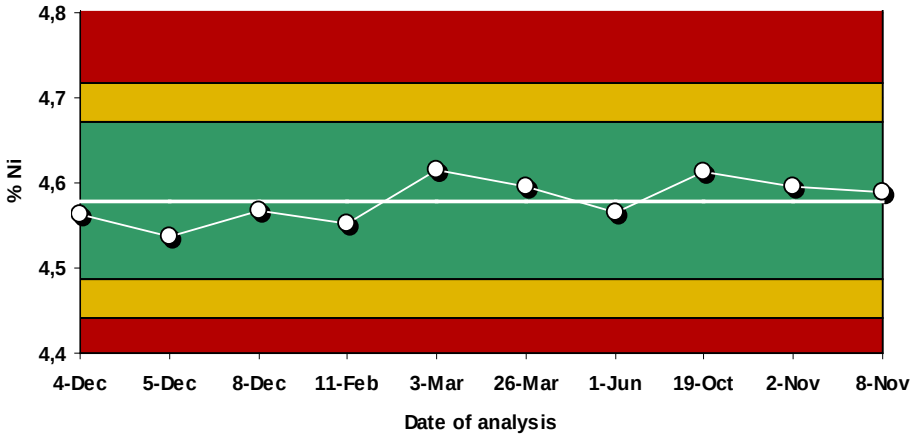
Sample type	Control chart	Control limits	Central line
Steel sample – test sample	X-chart	Target	Mean value

High concentration of nickel. The mean value for our control values over one year is 4.58 % (abs)[6] with a standard deviation of 0.026 % (abs). The control sample is taken through the whole measurement procedure (polishing and measurement).

The requirement regarding expanded measurement uncertainty[7] (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[8] so we obtain an estimate of the requirement from:

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

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



$\bar{x}$ = 4.58 % (abs)
s_{target} = 0.0458 % (abs)

CL: 4.58 % (abs)

WL: 4.58 ± 2 · 0.0458 = 4.67 and 4.49 % (abs)

AL: 4.58 ± 3 · 0.0458 = 4.72 and 4.44 % (abs)

.

[6] The X-chart concentration unit is in weight % of nickel (% abs) and the demand is given in relative percent of the nickel value (% rel).
[7] Further information on expanded and standard uncertainty is available in the Eurachem/CITAC guide (6).
[8] Due to the way standard deviations are combined this will result in a about max 36 % contribution from s_{Rw} allowing up to 64 % contribution from bias uncertainty to the standard uncertainty,

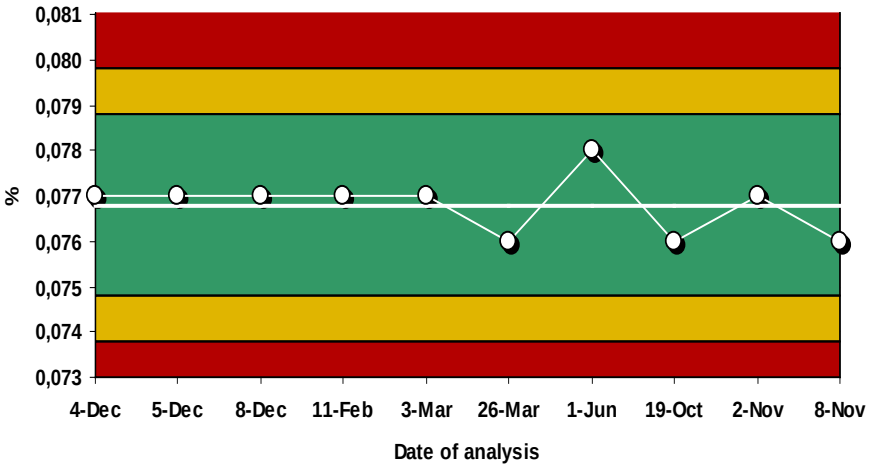
Example 2

Determination of Co in low-alloy steel with XRF

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

Low concentration of cobalt. The mean value for our control values over one year is 0.0768 % (abs)[9] with a standard deviation of 0.00063 % (abs). The control sample is covering the whole measurement procedure (polishing and measurement).

The requirement for limit of quantification LOQ is 0.01 % (abs) 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 % (abs) as a standard deviation and this value can be used to set the control limits. From the limit of quantification (LOQ) we therefore calculate the control limits to be:



$\bar{x}$ = 0.0768 % (abs)
s_{target} = 0.001 % (abs)

CL: 0.0768 % (abs)

WL: 0.0768 ± 2 · 0.001 = 0.0788 and 0.0748 % (abs)

AL: 0.0768 ± 3 · 0.001 = 0.0798 and 0.0738 % (abs)

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.

[9] See footnote 6 on page 35.

Example 3

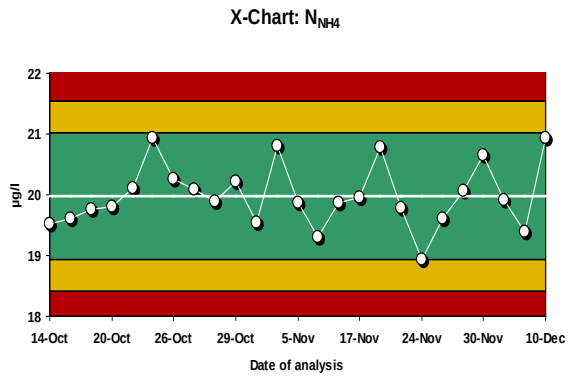
Determination of N-NH_4 in water with indophenol blue method

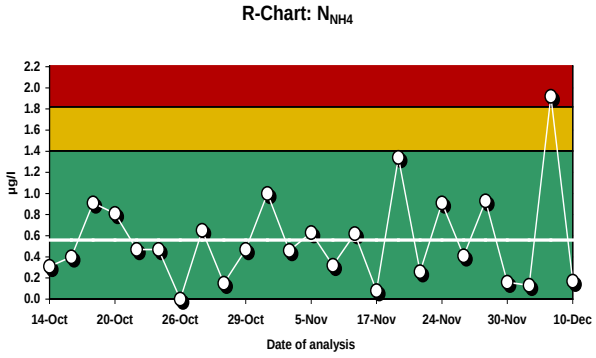
Sample type	Control chart	Control limits	Central line
Standard solution	**X-chart**	**Statistical**	**Mean value**
Low test samples	**R –chart**	**Statistical**	**Mean range value**

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 < *30 µg/l* selected among the samples to be analysed in that analytical run. The control was used for analyses of waters in the concentration range between 2 µg/l and 100 µg/l.

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

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





$\bar{x}$ = 19.99 µg/l and s = 0.521 µg/l CL: 19.99 µg/l WL: 19.99 ± 2·0.521 = 19.99 ± 1.04 µg/l (18.95 & 21.03 µg/l) AL: 19.99 ± 3·0.521 = 19.99 ± 1.56 µg/l (18.43 & 21.55 µg/l)	Mean range = 0.559µg/l and 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 in 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. This control value outside the action limit should therefore be rejected when reviewing the R-chart (Chapter 9 and 10).

Example 4

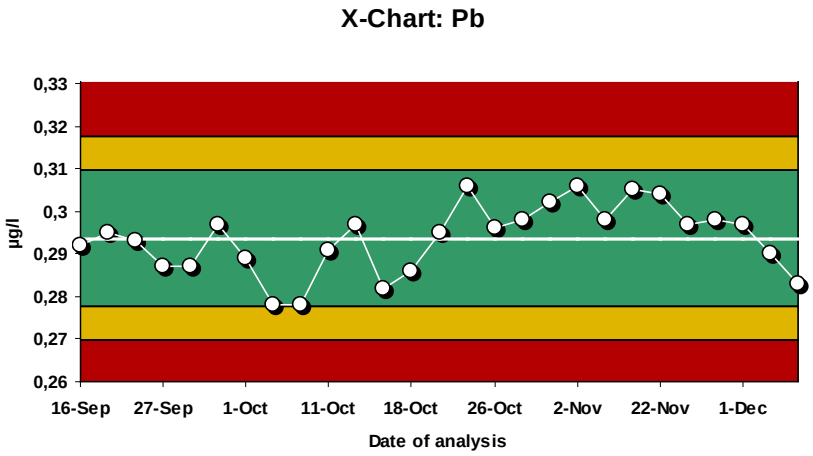
Determination of Pb in water with ICP-MS

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) in waters. 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 of 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

On the X-chart the control values were within the limits. No systematic effects were detected in the results.

There are 12 consecutive results above the central line. This is out of statistical control but as described in Chapter 9 regarded as acceptable.

Example 5

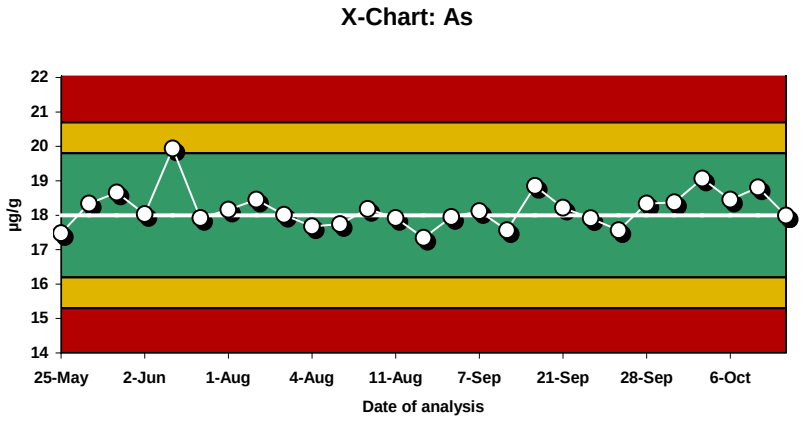
Determination of As in biological material with ICP-MS

Sample type	Control chart	Control limits	Central line
CRM	**X-chart**	**Target**	**Certified 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.

The X-chart was established as follows:

- the individual results were used for plotting of X-chart;
- the certified value was used as the central line (CL);
- the target standard deviation of 5 % was used to calculate the control limits



Certified value = 18.0 µg/g

s_{target} = 0.05·18.0 = 0.9 µg/g

CL: 18.0 µg/g

WL: 18.0 ± 2·0.9 =
= 18.0 ± 1.8 µg/g
(16.2 µg/g and 19.8 µg/g)

AL: 18.0 ± 3·0.9 =
= 18.0 ± 2.7 µg/g
(15.3 µg/g and 20.7 µg/g)

Comment

On the X-chart one control value exceeded the warning limit. However, the previous value and the next one were both within the warning limits – the method was in control (Chapter 9).

Example 6

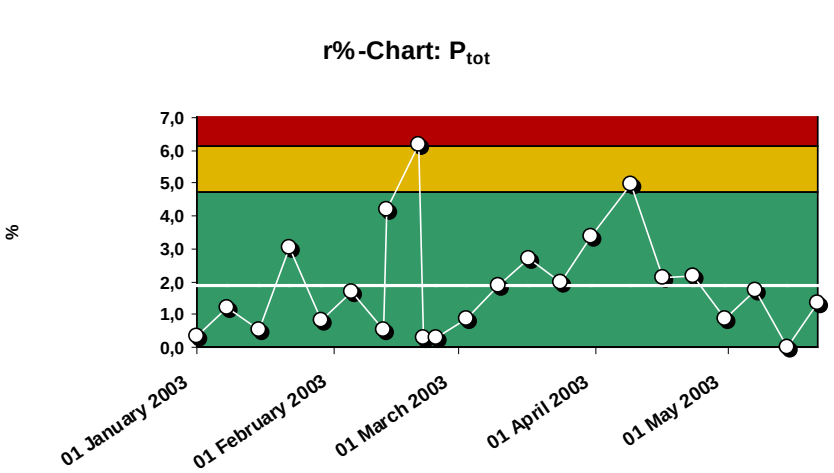
Determination of total P in water using spectrophotometric method

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

Test samples (10 - 50 µg/l). According to method validation the detection limit (3 *s*) was 2 µg/l. In each run one test sample was analysed as duplicates. The results were applied for r%-charting.

The r%-chart was established as follows:

- the 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 standard deviation of the r%-values was used for calculating the control limits.



$\bar{x}$ % = 1.88 %
s = 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 the first instance also the action limit was exceeded. The repeatability was out of control (Chapter 9) and after taking care of the problem the QC sample and the test samples were reanalysed.

Example 7

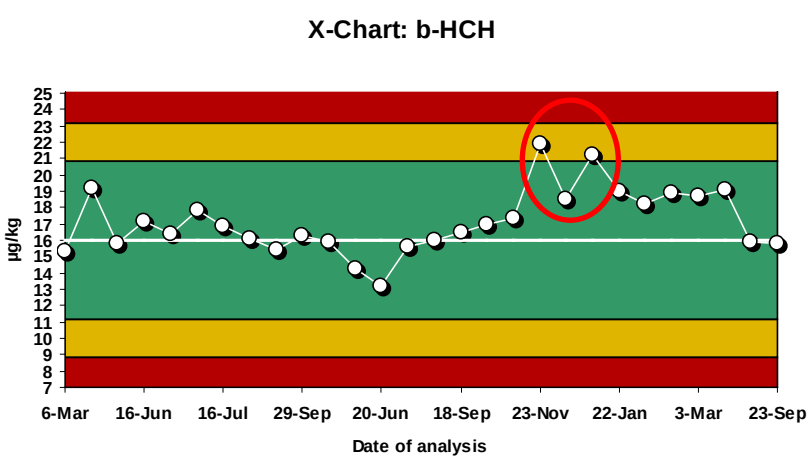
Determination of b-HCH (b-hexachlorocyclohexane) in biological material with Gas Chromatography

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

The CRM Cod liver oil BCR/598 with b-HCH (16 µg/kg). The control sample was used for analysis of b-HCH in biological material. The 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 certified value was used as the central line (CL);
- the target standard deviation of 15 % was used to calculate the control limits.



Certified value = 16.0 µg/kg

s_{target} = 0.15·16.0 = 2.4 µg/kg

CL: 16.0 µg/kg

WL: 16.0 ± 2·2.4 = 16.0 ± 4.8 µg/kg (11.2 µg/kg and 20.8 µg/kg)

AL: 16.0 ± 3·2.4 = 16.0 ± 7.2 µg/kg (8.8 µg/kg and 23.2 µg/kg)

Comment

A trend was detectable in the results: From September 11 (point number 15 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.

Example 8

Determination of Cu in water with ICP-OES

Sample type	Control chart	Control limits	Central line
In-house synthetic standard	**X-chart**	**Statistical**	**Mean value[10]**
	R-charts	**Statistical**	**Mean range**

In-house synthetic standard (1.00 ± 0.02 mg/l). The control sample was prepared from a commercial standard. The sample was preserved with HNO_3. Control was performed twice in each analytical run.

X- and R-charts were established in 2003. Preliminary control limits and central line were estimated from the first 60 analytical runs.

X-chart:

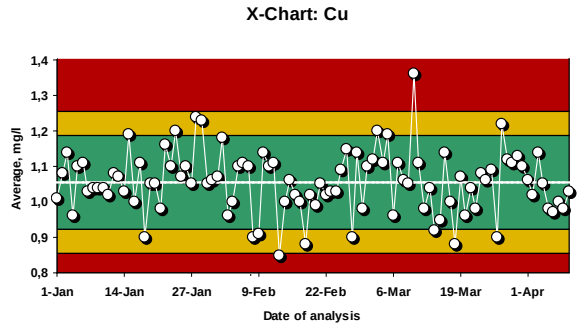
- the average of the results for the control sample in each run was plotted;
- the mean value was used as the central line (CL);
- the standard deviation was used for calculating the control limits.

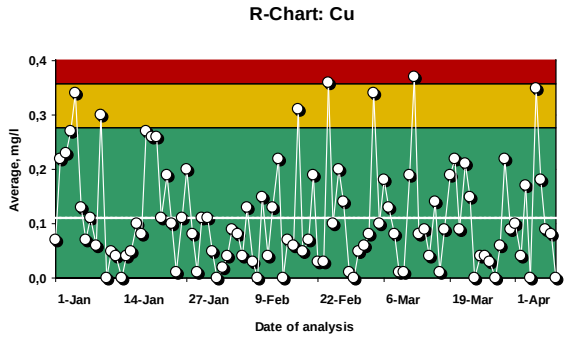
R-chart:

- the range for duplicates (highest value minus lowest value) was used for plotting;
- the mean range for the same 60 analytical runs that were used to establish the X-chart was used as the central line;
- the repeatability standard deviation (s_r) 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 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 1.255 mg/l)	Mean Range, R = 0.11 mg/l 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 9 we look at the last 60 data. These are the data plotted since 9 February 2004.

We count the number of times that the control values were outside the warning limits **since** 9 February. On the X-chart we find three cases where the upper warning limit is **clearly**

[10] In this old example the mean value is used for the central line in the X-chart as there is a 5 % bias. Today we would normally recommend the standard value for the central line and then wider control limits e.g. target control limits

exceeded, one of these even outside the action limit, and seven cases clearly below the lower warning limit. This makes a total of 10 times where the warning limits have been exceeded. There is thus reason to change the preliminary control limits. On the R-chart we find five cases outside the warning limit. This is less than the required number of more than six times, but we will review the limits in both control charts anyway.

One control value on the X-chart on 11 March was clearly 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 because 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 all statistical analysis of the data.

We calculate a new average and standard deviation from the last 59 points on the X-chart (only 59 since the outlier has been excluded) and a new average range for the last 60 points on the R-chart.

New $\bar{\bar{x}}$ = 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:

$s^2_{new}/s^2_{original} = 0.0834^2 / 0.0667^2 = 1.563$

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 (10 times with 60 data points). Since there was not a significant change we recommend recalculating the control limits based on all the data. 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:

$$s = \sqrt{\frac{(n_1 - 1) \cdot s_1^{\,2} + (n_2 - 1) \cdot s_2^{\,2} + \ldots + (n_k - 1) \cdot s_k^{\,2}}{n_{tot} - k}} =$$

$$\frac{(60 - 1) \cdot 0{,}0667^2 + (59 - 1) * 0{,}0834^2}{(60 + 59 - 2)} = 0{,}07545 \text{ mg/l}$$

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.

Previous preliminary X-chart	New X-Chart based on longer time period
$\bar{x}$ = 1.055 mg/l and s = 0.0667 mg/l 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 1.255 mg/l)	$\bar{x}$ = 1.048 mg/l and s = 0.0822 mg/l CL: 1.048 mg/l WL: 1.048 ± 2*0.0822 mg/l (0.884 and 1.212 mg/l) AL: 1.048 ± 3*0.0822 mg/l (0.801 and 1.295 mg/l)

R-chart

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

$F = s^2_{original}/s^2_{new} = 0.09752 / 0.0957^2 = 1.037$

The critical value for F from *Table 3* in Chapter 13 is 1.67 (see further under X-chart). This is larger than our calculated value for F and therefore the repeatability standard deviation has not changed significantly, and we recommend recalculating the control limits based on all the data. The new calculation gave the same mean range so no changes to the R-chart.

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 based on all available data.

However, there is a 5 % bias in comparison with the expected value of the control sample, a standard solution at a high level (1.00 ± 0.02 mg/l) and we would recommend investigating this and changing the procedure to reduce this bias. When the bias is reduced we recommend setting the central line to the value of the standard solution, 100 mg/l.

Example 9

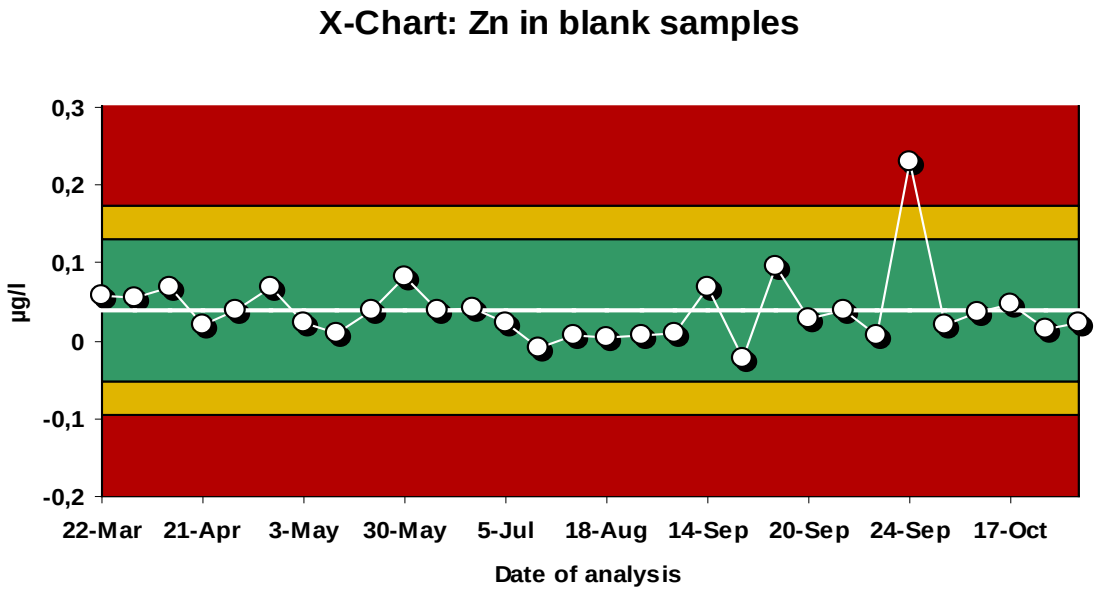
Determination of Zn in hydrogen peroxide with ICP-OES – procedural blank

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

Procedural blank of ultrapure water. The procedural blank 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 added and diluted with pure water to 5 ml and analysed with ICP-OES.

The X-chart was established as follows:

- 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 $s = 0.045$ mg/l

CL: 0.039 mg/l

WL: 0.039 + 2·0.045: 0.129 mg/l and –0.051 mg/l

AL: 0.039 + 3·0.045 = 0.174 mg/l and –0.096 mg/l

Comment

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

Example 10

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

Sample type	Control chart	Control samples each day	Different days
Any	X- chart	3	8

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 days 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 her giving s_{Rw} = 0.270. The correct estimate using ANOVA (Analysis of Variance) is 0.272. More info on ANOVA can be found in the Eurachem Guide *Fitness for Purpose of Analytical methods* (18).

15. References

1. V. B. Jensen, K. Haapala, H. Hovind & Ö. Lindgren. *Handbook of internal quality control in water laboratories*. Copenhagen: Water Quality Institute, 1984 (in Danish).
2. H. Hovind: *Internal Quality Control. Handbook for Water Analysis Laboratories*. Oslo: Norwegian Institute for Water Research, 1986 (in Norwegian).
3. ISO/IEC 17025:2017 – *General requirements for the competence of testing and calibration laboratories*, Geneva: International Organization for Standardization.
4. E. Hund, D. Luc Massart, J. Smeyers-Verbeke. *Operational definition of uncertainty*, TRAC 2001; 20(8):394-406.
5. B. Magnusson, T. Näykki, H. Hovind, M. Krysell, E. Sahlin, *Handbook for calculation of measurement uncertainty in environmental laboratories*, Nordtest Report, TR 537 (ed. 4) 2017. Available from www.nordtest.info, accessed 2018-03-12.
6. Eurachem/CITAC Working Group. *Quantifying Uncertainty in Analytical Measurement*. EURACHEM/CITAC Guide, 3rd Internet edition 2012. www.eurachem.org, accessed 2018-03-09.
7. ISO 13530:2010. *Water quality Guide to analytical quality control for water analysis*. Geneva: International Organization for Standardization; 2010.
8. European Union Council Directive 98/83/EC of 3 November 1998 with amendment 2015/1787 (2015).
9. A. J. Duncan. *Quality control and industrial statistics*. 5th Edition. Homewood. Illinois: Richard D. Irwin, 1986.
10. R. V. Cheeseman & A. L. Wilson (revised by M. J. Gardner). *A manual on analytical quality control for the Water Industry*. Swindon (UK): Water Research Centre; 1989, Publication NS30.
11. ISO 8258. *Shewhart Control Charts*. Geneva: International Organization for Standardization; 1991 0 Corr 1:1993.
12. ISO 7870-1. *Control Charts – Part 1: General Guidelines*. Geneva: International Organization for Standardization; 2007.
13. ISO 7873. *Control Charts for arithmetic average with warning limits*. Geneva: International Organization for Standardization; 1993.
14. J. N. Miller & J. C Miller. *Statistics and chemometrics for analytical chemistry*. 6t ed. Harlow: Prentice Hall, 2010.
15. W. Funk, V. Dammann, G. Donnevert. Funk W. *Quality assurance in analytical chemistry. Applications in Environmental, Food and Materials Analysis, Biotechnology and Medical Engineering*. 2nd revised Edition. Weinheim: Wiley-VCH, 2006.
16. International Vocabulary of Metrology – Basic and General Concepts and Associated Terms, ISO/IEC Guide 99, International Organization for Standardization (ISO)/ International Electrotechnical Commission (IEC), Geneva, 2007. Available as JCGM 200:2008 from www.bipm.org.
17. V. J. Barwick and E. Prichard (Eds) Eurachem Guide: Terminology in Analytical Measurement – introduction to VIM 3 (2011). Available from www.eurachem.org.
18. B. Magnusson and U. Örnemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014), ISBN 978-91-87461-59-0. Available from www.eurachem.org.
19. Requirements on analytical quality, under *Resources* at www.trollboken.se.

NT TECHN REPORT
569 Edition 5.1
Approved 2018-09

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