# LABORATORY DATA MANAGEMENT GUIDANCE

## Out of Expectation (OOE)

## and Out of Trend (OOT) Results

| Name and Role | Date |
|---|---|
| **Author:**<br>Dr Christopher Burgess<br>on behalf of the Expert Drafting Group | |
| **Technical Review:**<br>Dr Phil Nethercote<br>On behalf of the ECA Analytical Quality Control Working Group | |
| **Approved by:**<br>Dr Günter Brendelberger<br>On behalf of the ECA Analytical Quality Control Working Group | |

# Table of Contents

## Document Revision History

| Version | Date | Reason for Change | Status |
|---|---|---|---|
| V 0.1 | April 2014 | First structural draft | Draft |
| V 0.2 | 14 July 2015 | First full draft for Core Team Review | Draft |
| V 0.3 | 15-Aug-2015 | First full draft for Peer Review | Draft |
| V 0.4 | 02-Nov-2015 | Final draft for Core Team Review | Draft |
| V 1.0 | 16-Nov-2015 | Version 1 for OOT/OOE Forum December 2015 | Released |
| V 1.1 | 03-Nov-2016 | Additional regulatory references, minor updates for clarification and typographical errors | Released |
| V 1.2 | 11-Jul-2023 | Correction to equation 1.26 | Released |

## Scope & Application

This guideline applies primarily to physicochemical -based laboratory testing resulting in continuous data (variables, for example assay, impurity values, hardness etc., which may be assumed to be normally distributed) or discrete data (attributes for example, particle counts, identity tests or cosmetic quality defects derived from AQLs which are not normally distributed). However, for discrete data, it may also be applicable to the microbiological laboratory.

Laboratory tests are performed on active pharmaceutical ingredients, excipients and other components, in-process materials, and finished drug products. It is applicable to PAT (Process Analytical Technology) or RTR (Real Time Release) approaches. This SOP is complementary to, and should be used in conjunction with, the ECA SOP on OOS Results[1]

If a number of measurements are made over a short period of time and an anomalous or unexpected value is observed within these measurements then it is designated OOE (Out of Expectation). An OOE is defined as a parameter value which lies outside the expected variation of the analytical procedure variation with respect to either location or dispersion.

A trend can occur in a sequence of time related events, measurements or outputs. Trend analysis refers to techniques for detecting an underlying pattern of behaviour in a time or batch sequence which would otherwise be partly or nearly completely hidden by noise. These techniques enable specific behaviours (OOT; Out of Trend) such as a shift, drift or excessive noise to be detected.

There are two distinct types of trend situations;

1. Where the expectation is that there will be no trend, for example for production or analytical process data which are known or assumed to be under statistical control.

or

2. Where the expectation is that there is will be trend; for example in stability testing.

There is a fundamental difference between these two situations in that the variance increases with time in the second situation.

Therefore in this guideline there are three distinct sections covering OOE and the two types of OOT. Each section is supported by examples given in the appendices. The methods used in examples are intended to be advisory as to represent recommended practice but should not be mandatory. Other statistically sound procedures may be used as alternatives.

[1] STANDARD OPERATING PROCEDURE Laboratory Data Management; Out of Specification (OOS) Results, Version 2, 14-Aug-2012

## Expert Drafting Group

This guideline is the result of a collaborative effort involving

- members of the ECA AQCWG core team in the first instance
- and review/critique by many ECA attendees at the OOT Forum held in Prague in October 2014
- colleagues on the USP Validation and Verification Panel and the USP Statistics Subcommittee.

Those involved in the core team were;

| Team Member | Affiliation | Primary area of activity/role |
|---|---|---|
| **Dr Christopher Burgess** | Burgess Analytical Consultancy Limited UK | Chairman of the AQCWG of ECA and coordinating author |
| **Dr Milan Crnogorac** | Roche, Switzerland | SPC, attributes |
| **Dr Lori A. McCaig** | Roche, USA | Stability Trending |
| **Dr Peter Rauenbuehler,** | Roche, USA | Stability Trending |
| **Dr Bernd Renger** | Bernd Renger Consulting, Germany | OOE results |
| **Lance Smallshaw** | UCB Biopharma sprl , Belgium | SPC variables |
| **Dr Bianca Teodorescu** | UCB Biopharma sprl , Belgium | SPC oversight & statistician |
| **Stephen Young** | MHRA, UK | Regulatory aspects |

## Regulatory References

1. Guidance for Industry; Investigating Out-of-Specification (OOS)
   Test Results for Pharmaceutical Production, US **F**ood and **D**rug **A**dministration, Center for Drug Evaluation and Research (CDER), October 2006
2. Guidance for Industry Process Validation: General Principles and Practices, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Veterinary Medicine (CVM) January 2011
   *'An ongoing program to collect and analyze product and process data that relate to product quality must be established (§ 211.180(e)). The data collected should include relevant process trends and quality of incoming materials or components, in-process material, and finished products. The data should be statistically trended and reviewed by trained personnel. The information collected should verify that the quality attributes are being appropriately controlled throughout the process.'*
3. Out Of Specification Investigations, **M**edicines and **H**ealthcare products **R**egulatory **A**gency, UK, (MHRA) November 2010 updated 2013
4. "The Rules Governing Medicinal Products in the European Union", Volume 4, Good Manufacturing Practice (GMP) Guidelines 2015
   **Part I - Basic Requirements for Medicinal Products**
   a. Chapter 1 Quality Management System 1; 10 Product Quality Review
   b. Chapter 6 Quality Control; Documentation 6.7 & 6.9
      Testing 6.16
      On-going stability programme 6.32, 6.32 & 6.36
   c. Chapter 8 Complaints, Quality Defects and Product Recalls
      Root Cause Analysis and Corrective and Preventative Actions 8.19

   **Part II - Basic Requirements for Active Substances used as Starting Materials**
   a. Chapter 15 Complaints and Recalls; 15.12

   **Annex 2 Manufacture of Biological active substances and Medicinal Products for Human Use**
   Seed lot and cell bank system 42, 49
   Quality Control 70

   **Annex 6 Manufacture of Medicinal Gases** Manufacture 2

   **Annex 15 Qualification and Validation**
   **Ongoing Process Verification during Lifecycle** 5.29, 5.30 & 5.31
   Manufacturers should monitor product quality to ensure that a state of control is maintained throughout the product lifecycle with the relevant process trends evaluated.
   Statistical tools should be used, where appropriate, to support any conclusions with regard to the variability and capability of a given process and ensure a state of control.

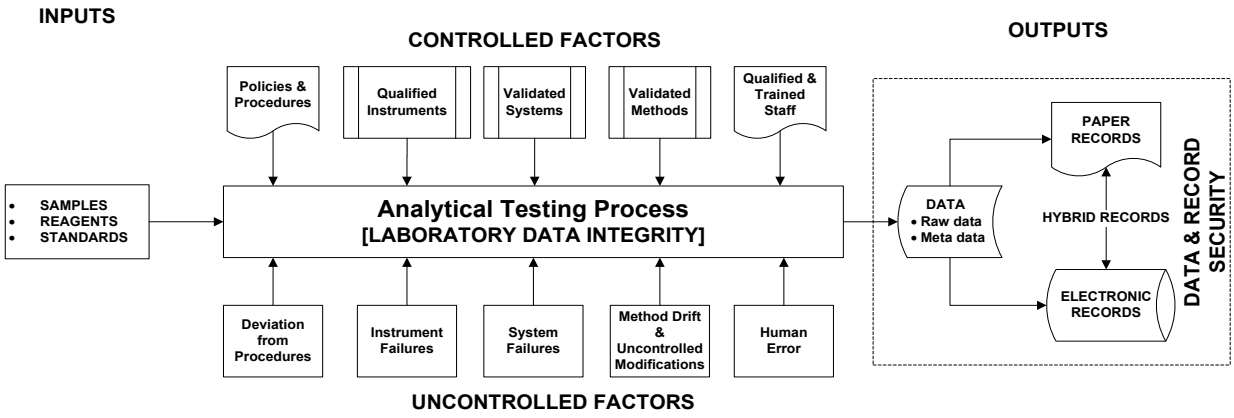
   **Annex 16 Certification by a Qualified Person and Batch Release** 1.7.16
5. USP 38 (2015) General Chapter <1010>, ANALYTICAL DATA; INTERPRETATION & TREATMENT
6. ISO/IEC 17025 2nd edition (2005) General requirements for the competence of testing and calibration laboratories Section 5.9 – assuring the quality of test and calibration results.
7. ICH Harmonised Tripartite Guideline, Q10, Pharmaceutical Quality System (2008); Control Strategy '*A planned set of controls, derived from current product and process understanding that assures process performance and product quality*'.

8. WHO Technical Report Series 996, Annex 5 Sections 6 and 11 (2016)
9. PIC/S Draft Guidance PI041-1 Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments; 10th August 2016

## Overview of Laboratory Data Management & the Analytical Process

Laboratory data quality management processes are a part of the overall Quality Management System as required by Chapter 1 of EU GMP and the FDA cGMPs as specified in 21 CFR §210 & §211.

Analytical processes and procedures are managed as part of a lifecycle concept. Laboratory data integrity and security are critical requirements under the GMPs. Such a process is illustrated below.




The purpose of this guidance document is to define the procedures for managing laboratory data which are Out-of-Expectation (OOE) or Out-of-Trend(OOT). Any confirmed OOE or OOT should trigger a deviation and appropriate investigation. The investigation should follow the principles laid down in the Out-of-Specification (OOS) SOP, ECA_AQCWG_SOP 01. This guidance document does not cover the evaluation of trend data with respect to specification. Process capability is mentioned briefly but the details are a topic beyond the scope of this document.

The pharmaceutical industry lags far behind many other manufacturing industries in the area of process evaluation and control. This guidance document is intended to assist in the simple implementation of trending techniques to meet regulatory requirements particularly in the areas of Product Quality Review (EU) and Annual Product Review (US).

In 1960, Dr Genichi Taguchi introduced a new definition of "World Class Quality" namely;

*On target with Minimum Variance*

This contrasts with the traditional Conformance with Specification, previously adopted by the FDA and other authorities.

Indeed it is not in technical accordance with the principles of Six Sigma which allowed the mean to vary ±1.5σ.

However, his revolutionary definition ensured that the application of statistical process control techniques was in the forefront of the tools required to achieve this life cycle objective.

## QU involvement/Responsibilities

Quality Control testing is considered an integral part of the Company's Quality Unit as explicitly required by EU GMP. Formal Quality involvement, e.g. by a separate QA function, should be kept to the minimum consistent with US & EU regulatory expectations and requirements based upon published legislation and guidelines.

The extent of Quality oversight is very dependent on individual company requirements. Organisation and nomenclature of Quality Control and Assurance functions and assignment of responsibilities are also highly company specific. This Guideline does not dictate or recommend specific steps that must be supervised by specific quality functions other than those required by regulation. Therefore the term Quality Unit (QU) as used in the revised chapter 1 of EU GMP Guide, is used here.

The initial OOE or OOT investigation, however, should be performed directly under the responsibility of the competent laboratory.

## Overview & purpose of trend analysis

The approaches set out in this guidance document are dependent on the applicable shape (mathematical distribution model) of the data. The data types under consideration here are variables and attributes. A continuous random variable is the one which can take any value over a range of values for example an assay value or an impurity level. An attribute is an integer where the set of possible values for a discrete random variable is at most countable for example a cosmetic defect on a tablet or the number of particles in a solution.

**Hence the selection of the appropriate mathematical distribution may be shown as a decision tree. For example**

Figure 1, which is an illustrative example only and not exhaustive.

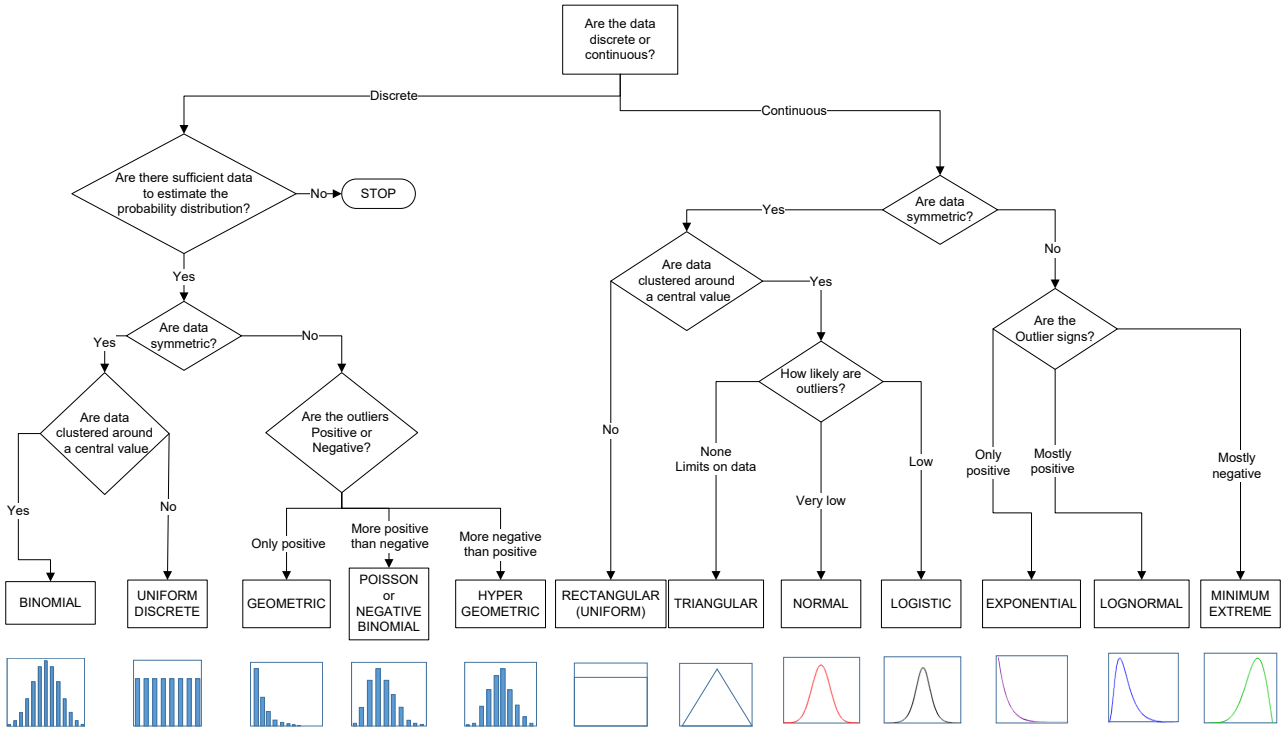



**Figure 1 Decision tree for the selection of an appropriate mathematical data model based on data shape[2]**

For our purposes, the most useful distribution for continuous variables is the Normal or Gaussian distribution for a population whose properties are well known.

For a true mean value ($\mu$) of zero and a standard deviation ($\sigma$) of 1 then the probability distribution is given by

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]} \qquad (1.1)$$

and shown graphically in Figure 2. The areas under the curve indicate the probability of values lying ±σ, ±2σ and ±3σ from the mean. This distribution is the basis for control charting of continuous random variables and stability trending as discussed later.

[2] Adapted and redrawn from a paper by a Prof Aswath Damodaran at the Stern School of Business at New York University http://people.stern.nyu.edu/adamodar/New_Home_Page/home.htm

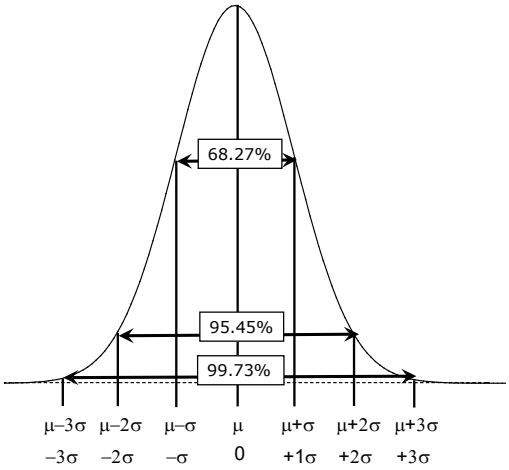



**Figure 2 Normal distribution for a mean value ($\mu$) of zero and a standard deviation ($\sigma$) of 1**

For attribute data, the Binomial or Poisson distributions are preferred. If counted defects are to be used, the Binomial distribution is used. If the data are defects expressed as a % for example then the use of the Poisson distribution is indicated.

## Control Charting Concept

Conceptually, a control chart is simply a plot of a response variable against time or batch whereby the degree of variation is predicted by the chosen distribution (mathematical model) around a mean or target value. Hence for a continuous variable which is assumed to be normally distributed the trend plot is shown in Figure 3 The decision rules regarding an out of trend result come from the likelihood of the pattern of responses or the distance from the target or mean value.

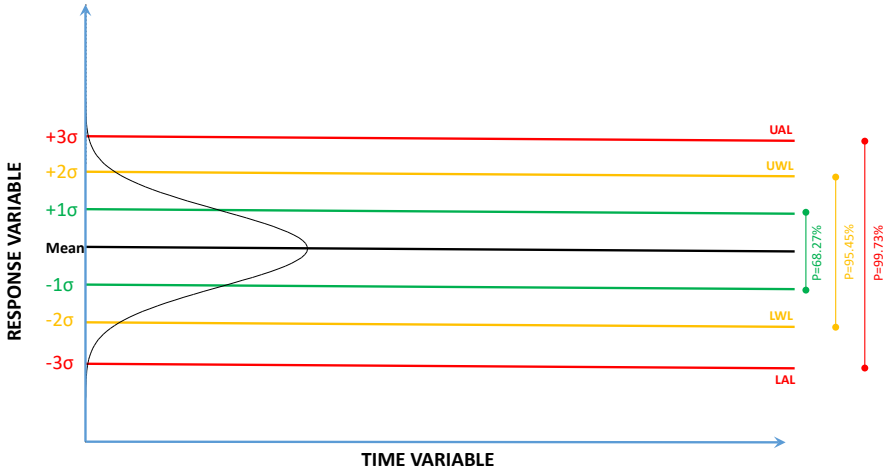
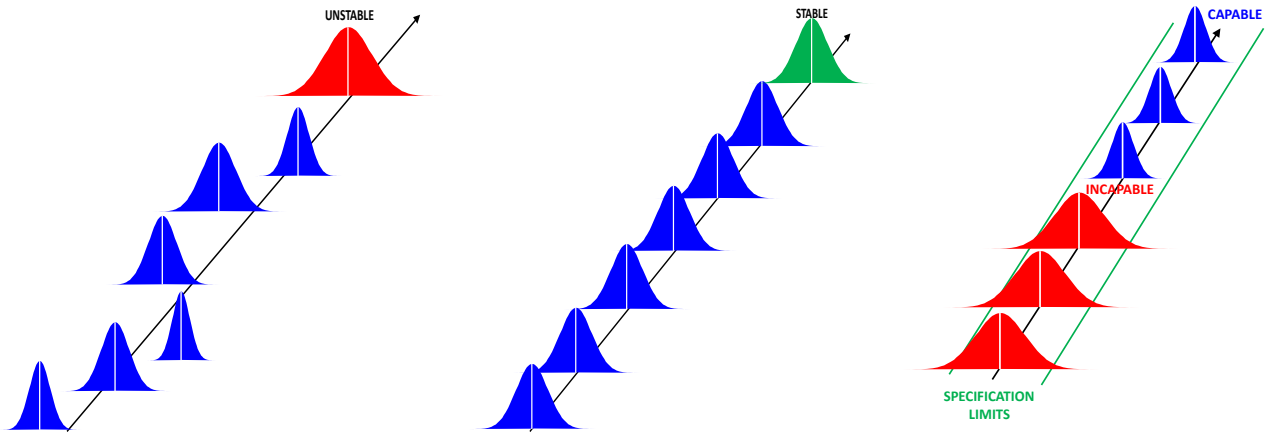



Figure 3 Idealised control chart for a continuous variable under the normal distribution

The approach is based on the idea that, no matter how well a process is designed, there exists a certain amount of natural variability in output measurements. When the variation in process quality is due to random causes alone, the process is said to be statistically in-control. If the process variation includes both random and special causes of variation, the process is said to be statistically out-of-control.

All test results have variation that comes from measurement (system) variation and process performance variation. There are two types of variation; Common Cause variation, inherent noise, and Special Cause variation owing to, for example, a process shift, drift or excessive noise .

The control chart is designed to detect the presence of special causes of variation. The normal distribution may be characterised by two particular parameters; a measure of location (the arithmetic mean or average) and a measure of dispersion (the standard deviation). If a process is unstable it means that either of these parameters are changing in an uncontrolled manner (Figure 4 (a)). This would be apparent from a mean and range control chart for example. The next task would be to bring these two parameters into a state of statistical control. This would entail ensuring that the mean and the standard deviations were not varying significantly. This ideal situation is illustrated in (Figure 4 (b)). This would then said to be under statistical control i.e. no special cause variation and controlled common cause variation. In this state, the process is amenable to the tools of Statistical Process Control (SPC). However, a stable process may not be statistically capable of meeting the specification limits. Figure 4 (c) illustrates this showing that the red process albeit stable is incapable. The desired state is, of course, to arrive at the blue capable state. The method of calculating process capabilities are briefly described later in this guidance.




(a) An unstable process (b) A stable process (c) Stable Processes; Capable and In Capable

**Figure 4 Process stability & capability**[3]

The question is how are we to judge when a process is in a state of statistical control with respect to time?

[3] Redrawn and based on QMS – Process Validation Guidance, GHTF/SG3/N99-10:2004 (Edition 2) Annex A Statistical methods and tools for process validation [http://www.imdrf.org/documents/doc-ghtf-sg3.asp]

The answer lies in the application of SPC decision rules. These are based on the patterns expected from the distribution shown in Figure 3. These rules were developed many years ago and the simplest are the four WECO rules[4].

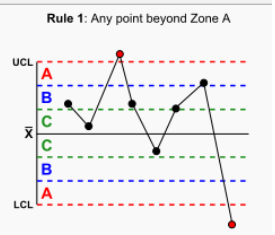
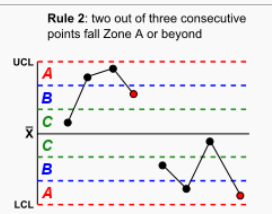
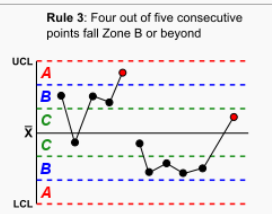
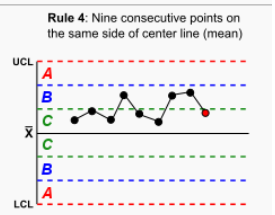
| Rule | Description | Chart Example |
|---|---|---|
| Rule 1 | Any single data point falls outside the 3σ limit from the centerline (i.e., any point that falls outside Zone A, beyond either the upper or lower control limit) |   |
| Rule 2 | Two out of three consecutive points fall beyond the 2σ limit (in zone A or beyond), on the same side of the centerline |   |
| Rule 3 | Four out of five consecutive points fall beyond the 1σ limit (in zone B or beyond), on the same side of the centerline |   |
| Rule 4 | Nine consecutive points fall on the same side of the centerline (in zone C or beyond) |   |

**Figure 5 The 4 basic WECO rules for detecting out of trend (OOT) results**

More recently, an extended set of 8 rules. the Nelson Rules[5], have been proposed. These rules are incorporated within many standard software control charting applications such as Minitab or SAS JMP for example. The choice of rule selection is left to the user. It is not recommended to select all rules as this increases the likelihood of false trends being identified. Quite often, the 4 basic WECO rules are sufficient.

[4] Western Electric Company (1956), Statistical Quality Control handbook. (1 ed.), Indianapolis, Indiana: Western Electric Co or see Montgomery, Douglas C. (2009), Introduction to Statistical Quality Control (6 ed.), Hoboken, New Jersey: John Wiley & Sons

[5] Lloyd S. Nelson, "Technical Aids," Journal of Quality Technology 16(4), 238-239, (October 1984)

| | | |
|---|---|---|
| Rule 1 | One point is more than 3 standard deviations from the mean | One sample is grossly out of control |
| Rule 2 | Nine or more points in a row are on the same side of the mean | Some prolonged bias exists |
| Rule 3 | Six or more points in a row are continually increasing or decreasing | A trend exists. This is directional and the position of the mean and size of the standard deviation do not affect this rule. |
| Rule 4 | Fourteen or more points in a row alternate in direction, increasing then decreasing. | This much oscillation is beyond noise. This is directional and the position of the mean and size of the standard deviation do not affect this rule. |
| Rule 5 | Two or three out of three points in a row are more than 2 standard deviations from the mean in the same direction. | *There is a medium tendency for samples to be out of control.* |
| Rule 6 | Four (or five) out of five points in a row are more than 1 standard deviation from the mean in the same direction | *There is a strong tendency for samples to be slightly out of control.* |
| Rule 7 | Fifteen points in a row are all within 1 standard deviation of the mean on either side of the mean | *With 1 standard deviation, greater variation would be expected* |
| Rule 8 | Eight points in a row exist with none within 1 standard deviation of the mean and the points are in both directions from the mean. | *Jumping from above to below whilst missing the first standard deviation band is rarely random* |

**Table 1 Nelson Rules for Trend Detection**

# Detecting and Managing OOE results

## Introduction

The terms “Out of Specification Result” and “Out of Trend Result” are well defined, e.g. in the UK Medicines and Healthcare Products Regulatory Agency (MHRA Guidance “Out Of Specification Investigations”, detailing the MHRA expectations, second version, issued 2013:

- **Out-of-Specification (OOS) Result** – test result that does not comply with the pre-determined acceptance criteria (i.e. for example, filed applications, drug master files, approved marketing submissions, or official compendia or internal acceptance criteria)
- **Out of Trend (OOT) Result** – a stability result that does not follow the expected trend, either in comparison with other stability batches or with respect to previous results collected during a stability study. However, trends of starting materials and in-process samples may also yield out of trend data. The result is not necessarily OOS but does not look like a typical data point. Should be considered for environmental trend analysis such as for viable and non viable data action limit or warning limit trends.

This definition is extremely focused on stability studies, however, mentioning environmental trend analysis indicates that OOT results may also be observed during trend analysis for statistical process control.

However, no formal definition is given for the term “Out of Expectation Result”. In contrast to OOS results this is not linked to a violation of a formal specification and in contrast to OOT results this is not statistically deducible from a data base comprehensive enough to allow calculation whether the result belongs to a population to be expected from the analytical procedure’s uncertainty or not. This might be possible starting from a number of 30 independent tests.

To be considered an "Out of Expectation Result” or to be "discordant" there must be an expectation based on some evidence what would be the most likely outcome of the analytical process performed. This excludes any unusual result derived from analysing a sample with a totally unknown assay or content of the analyte in question.

Two different cases might therefore be considered "Out of Expectation Results”:

## Unexpected Variation in Replicate Determinations

Usual analytical practice will use a specific number of replicates - that is several discrete measurements - to provide more accurate results. These may be either replicate injections from the same HPLC sample preparation, replicate readings or other multiple determinations. This procedure has to be specified in the written, approved test procedure together with the limits for variability (range and/or RSD) among the replicates. These could be based upon the process capability of the method as determined during the method development and its subsequent validation. However, usually companies use a general limit of the range of $\Delta \leq 2.0$ % for assays. In case of replicate series of complete tests (full run-throughs of the test procedure) wider limits for variability among the replicates may be defined.

Any unexpected variation in replicate determinations - either derived by multiple measurements of one sample preparation or replicate series of complete tests - disqualifies this data set from being used for further result calculation. E.g. if the range between replicates is limited to $\Delta \leq 2.0$ % and the two replicates differ by 2.2 %, data generated from the analysis cannot be used. It is very important that the documentation accompanying the analysis is very clear about why the data sets have been rejected.
If only one set of data within a bigger data pool is affected - e.g. one out of several samples and reference samples tested over night using an automated HPLC system - only the directly affected replicates are considered disqualified, all other data in the series may be further processed to calculate the results of the other samples.

When unexpected variation in replicate determinations occurs, investigation into the cause is required similar to an investigation in the case of a non-compliant system suitability (SST) test. Usually this is reported as a laboratory deviation. The flow of the investigation may follow the proven approach of investigating an OOS result on a lab scale.

Repeating the test or measurement- preferable using the same sample preparation if appropriate - should not be performed prior to identifying a hypothesis why the replicates range was higher than expected and having taken corresponding actions.

## Unexpected Results in a Single Test or a Small Set of Tests

Analytical results from one single performance of one test or from a small number of tests obtained over a short period of time may be considered "Out of Expectation" if

- The test result does not fit into the other results of that series, but the number of tests and data points is not comprehensive enough to allow statistical calculation whether the result belongs to a population to be expected from the mean and the variability of the overall data set.
- The result does not violate a given specification
- There is enough evidence and information allowing to anticipate the "expected" result and thus to allow judgement that the result does not represent the expectations.

This anticipation may be based on

- Analytical results of the same sample or the same material using another, validated analytical procedure (e.g. IPC testing of a compounded bulk product, using an UV assay procedure and a later testing of the filled product using HPLC)
- Knowing the theoretical composition of the sample (e.g. samples prepared during galenic development)
- Results of tests of other samples/batches within a campaign or series of experiments(e.g. results of three out of four batches in one campaign are close to the theoretical assay, one is close to a specification limit)

To decide, whether a result is really out of expectation or may be considered representing the typical variability of the procedure applied, data of the analytical validation of the procedure used should be used.

According to the concept of analytical uncertainty usually applied in chemical analysis, the combined standard uncertainty of the result would be the appropriate performance indicator to help deciding, whether the result in question really is "unexpected" or simply represents a rare, but still probable value.

As analytical uncertainties of pharmaceutical test procedures are rarely established, a common way to estimate this range may be used.

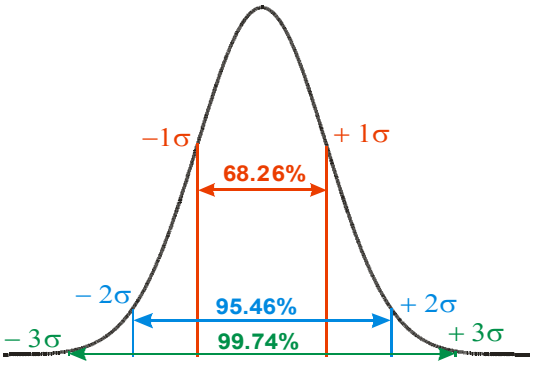
- Expanded analytical uncertainty = 1.5 x RSD intermediate precision[6]

In case an assay procedure based on HPLC has a reported (and correctly determined) intermediate precision of 0.8 %, the expanded analytical uncertainty to be expected in later routine application of the procedure is 1.2 % RSD.

To determine the limits (based on a 95 % confidence level) within which analytical results are representing the analytical variability of the procedure to be expected and accepted, the following calculation has to be performed;

- 95 % confidence interval = 2 x expanded analytical uncertainty

In the example, any analytical result falling within a range of ± 2 x 1.2 % = ± 2.4 % of the anticipated result are representing analytical variability of the procedure on a 95 % confidence level and have to be accepted as is.




Only results falling outside this range are to be considered "out of Expectation".

In this case, data should not be used and accepted without previous investigation to determine the cause for the unexpected discrepancy from the anticipated result. This investigation should follow the well established process of laboratory investigations in case of OOS results.

[6] B Renger, Journal Chromatography B, 745 (2000), 167 - 176

# Trend Analysis for Statistical Process Control

## Overview

A control chart provides the simplest means of visually tracking a process to identify trends. It consists of a horizontal plot of an ongoing performance characteristic -- for example, analytical result for a particular parameter -- with a new data point added for each new measurement. Overlaid lines show evaluation criteria such as allowed tolerances. The control chart highlights poor quality by showing when a measurement lies outside the expected variation. More importantly, it shows when a process is trending toward failure. There are many different types of control charts, a number of these are discussed in this guideline.

As mentioned earlier, all measurements have variation. There are two types of variation.

1. Common Cause variation or noise
2. Special Cause variation such as process shift, drift or excessive noise.

The purpose of a control chart is to detect Special Cause variation. The expectation for a process is that it is under statistical control i.e. the only component of the variation is the test result noise.

## Control of continuous data

Quality Control (QC) plays an essential role in the Pharmaceutical and Biopharmaceutical industries and associated processes. A large part of QC focuses on tracking the ongoing performance of a process to spot problems or identify opportunities for improvement. An ideal quality control system will highlight the approach of trouble even before it becomes a problem. A number of statistical and graphical techniques exist for tracking ongoing quality performance.

Under certain circumstances, if not investigated and or corrected, an OOT may lead in time to an OOS and therefore an identification of an OOT may be an early indicator of a potential future OOS and so facilitate action being taken designed to reduce the incidence of non-random OOS results.

Thus the generation of trended data is an essential tool in the management of manufacturing and quality issues. These processes may only be effective where there is a suitable control strategy in place.

A control strategy is a planned set of controls, derived from current product and process understanding, that ensures process performance and product quality. These controls can include parameters and attributes related to drug substance and drug product materials and components, facility and equipment operating conditions, in-process controls, finished product specifications, and the associated methods and frequency of monitoring and control.

A typical control strategy for a Product Quality and Process Performance life cycle in the pharmaceutical industry today may consist of the following elements:

- Process mapping and identification of Critical Process Parameters
- In-Process Monitoring and control of Process Performance Attributes
- Monitoring and control of Critical Process Parameters linked to Critical Quality Attributes
- Controls for facility and equipment
- Monitoring the Drug Substance (API) and excipients against purchasing specification
- Monitoring and trending of stability data for product and raw materials including the API

An Out of Trent (OOT) result is a non-random event that is identified as test result or pattern of results that are outside of pre-defined limits. For continuous data evaluation, this guideline recommends using simple Shewhart type control charts in the first instance. These control charts developed in the 1930s have been widely applied in engineering and manufacturing industries.

These control chart use data that is collected in an appropriate manner and then applied to the standard or ideal result based upon historical data. The centre line on any control chart represents the mean (average) of the values collected during a reference period.

One (or more) line(s) is positioned both above and below the centre line to serve as control limits. These limits, the Upper Control Limit and the Lower Control Limit (UCL and LCL), provide a range of what is still acceptable for a result. Control charts are therefore used to determine if the results that are coming in are within the limits of what is acceptable or if the process is out of control. These upper and lower control limits must, wherever possible, be based on the values determined for the Proven Acceptable Range (PAR) and Normal Operating Range (NOR) for a process.

In investigational circumstances it may be required to analyse historical data to see if there have been special cause variations. In this instance a post mortem CuSum approach is to be recommended

CuSum stands for "cumulative sum." A CuSum chart is related to a standard control chart and is made in much the same manner, except that the vertical axis is used to plot the cumulative sum of the variability (differences between successive values) in a process. This CuSum is plotted on the vertical (Y) axis against time on the horizontal (X) axis.

This type of plot is helpful in spotting a biased process, in which the process misses the calculated mean value high or misses it low, since repeated misses on one side of the ideal value will force the cumulative sum away from the ideal value or benchmark value (which may be zero ) which is the ideal low variance (no variance) objective.

The minimum number of data values from which a suitable statistical mean can be calculated for use in a CuSum chart is 10 individual values. The maximum number of values to limit variation in the data set, should be set at 30 to 100 data values.

This technique is discussed in detail in on page 44 with a worked example in Appendix 4.

## Determination of a Trend using Statistical Process Control (SPC)

Statistical Process Control (SPC) is a way of using statistical methods and visual display of data to allow us to understand the variation in a process with respect to time. By understanding the types and magnitudes of variation in the process we can make improvements to the process that we predict will lead to better outcomes. SPC can also then be used to confirm if our predictions were correct. The methods were developed by Walter Shewhart and W Edwards Deming (and others) throughout the first half of the twentieth century.

Measurements of all outcomes and processes will vary over time but variation is often hidden by current practices in data management, where data is aggregated (averaged) and presented over long time periods (e.g. by quarter). Plotting data continuously (weekly or monthly) can be very informative. If we do this we reveal the sources and extent of variation.

## Control of continuous data

When dealing with a quality characteristic that is a variable we want to make sure that the characteristic is under control.
Shewhart identified two sources of process variation: common cause variation (chance variation) that is inherent in process, and stable over time, and special cause variation (assignable, or uncontrolled variation), which is unstable over time - the result of specific events outside the system.
A process that is operating only with common causes of variation is said to be in statistical control. A process that is operating in the presence of assignable causes is said to be out of control. The eventual goal of SPC is the elimination of variability in the process.

The control chart was designed so that one could distinguish between common and special causes of variation within a process and to provide a rule for minimizing the risk of reacting to a special cause when it is in fact a common cause, and not reacting to a special cause when one is present. It allows visualizing variations that occur in the central tendency and dispersion of a set of observations

A typical control chart has control limits set at values such that if the process is in control, nearly all points will lie between the upper control limit (UCL) and the lower control limit (LCL).
A control chart is typically constructed as follows:

$$UCL = \mu_W + L\sigma_W$$
$$\text{Centre Line} = \mu_W \qquad (1.2)$$
$$UCL = \mu_W - L\sigma_W$$

where L = a constant multiplier which will define the distance of the control limit from the centre line
$\mu_W$ = mean of the sample statistic, W.
$\sigma_W$ = standard deviation of the statistic, *W*.

When the assignable causes are eliminated and the points plotted are within the control limits, the process is in state of control. Further improvement can be obtained through changing basic process, system. Depending on the data than can be collected and on the purpose (detect small shift or large shift, investigation or continuous process verification), different control charts can be used. The following flowchart gives an indication of which chart to use when.

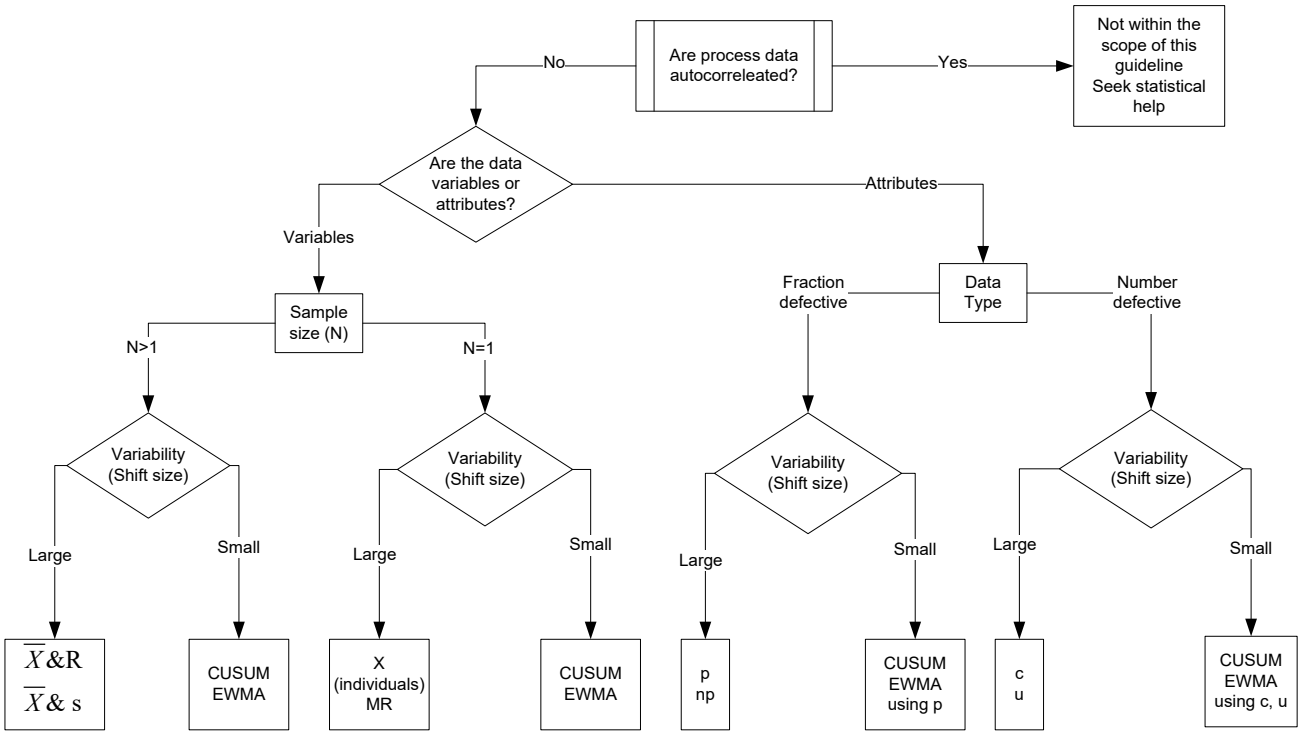



**Figure 6: Control Charting selection process**
**[redrawn & based on frontis illustration in D. C. Montgomery – Introduction to Statistical Quality Control) 6th Edition 2009]**

## I-Moving Range (MR) Control Charts

Individual control charts (or Shewhart control charts) are used whenever the sample size for process monitoring is n=1, for example one observation per batch. The moving range (MR) of two consecutive observations is used as an estimation of process variability:

$$MR_i = |x_i - x_{i-1}| \tag{1.3}$$

The estimator of process average, $\bar{x}$, is:

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{1.4}$$

### The Individuals chart control limits

$$UCL = \bar{x} + 3\frac{\overline{MR}}{d_2}$$
$$\text{Centre Line} = \bar{x} \tag{1.5}$$
$$UCL = \bar{x} - 3\frac{\overline{MR}}{d_2}$$

Where $d_2$ =1.128 (from Table 2) and $\overline{MR} = \frac{1}{m}\sum_{i=1}^{m} MR_i$

### The MR chart control limits

$$UCL = D_4\overline{MR}$$
$$\text{Centre Line} = \overline{MR} \tag{1.6}$$
$$LCL = D_3\overline{MR}$$

Where *$D_3$ and $D_4$* are from Table 2

### X-bar and R/S Control Charts

When data are collected in subgroups (eg, several determinations on the same batch), the X-bar control chart for subgroups means is being used. It is usually presented along with R-charts or S-charts. The R-chart plots subgroup ranges (when subgroup sample size <9), and the S-chart plots subgroup standard deviations (when subgroup sample size >=9).

Suppose m samples are available, each containing n observations. Let $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_m$ be the average of each sample, then the estimator of the process average is

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m}{m}$$

Let $R_i = |\max(group\ i) - \min(group\ i)|$, the range for group i, i=1, …, m. Then the average range is:

$$\bar{R} = \frac{R_1 + R_2 + \cdots + R_m}{m}$$

The X-bar chart control limits

$$UCL = \bar{\bar{x}} + A_2 \bar{R}$$
$$\text{Centre Line} = \bar{\bar{x}} \qquad (1.7)$$
$$LCL = \bar{\bar{x}} - A_2 \bar{R}$$

where the constant $A_2$ is tabulated for various sample sizes in Table 2.

### The R Chart control limits

$$UCL = D_4 \bar{R}$$
$$\text{Centre Line} = \bar{R} \qquad (1.8)$$
$$LCL = D_3 \bar{R}$$

Where $\bar{R}$ is the sample average range and the constants $D_3$ and $D_4$ are tabulated for various sample sizes in Table 2

### The S Chart control limits

The average of the m standard deviations is

$$\bar{s} = \frac{s_1 + s_2 + \cdots + s_m}{m}$$

The limits of the S-Chart are

$$UCL = B_4 \bar{s}$$
$$\text{Centre Line} = \bar{s} \qquad (1.9)$$
$$LCL = B_3 \bar{s}$$

Where the constants $B_3$ and $B_4$ are tabulated for various sample sizes in Table 2

Also the parameters of the X-bar chart can be adapted to include $\bar{s}$, instead of $\bar{R}$.

### The X-bar chart control limits

$$UCL = \bar{\bar{x}} + A_3 \bar{s}$$
$$\text{Centre Line} = \bar{\bar{x}} \qquad (1.10)$$
$$LCL = \bar{\bar{x}} - A_3 \bar{s}$$

Where the constant $A_3$ is tabulated for various sample sizes in Table 2.

| | AVERAGES | | STANDARD DEVIATIONS | | MEAN and RANGE | | |
|---|---|---|---|---|---|---|---|
| # of Observations (n) | $A_2$ | $A_3$ | $B_3$ | $B_4$ | $d_2$ | $D_3$ | $D_4$ |
| 2 | 1.880 | 2.659 | 0 | 3.267 | 1.128 | 0 | 3.267 |
| 3 | 1.023 | 1.954 | 0 | 2.568 | 1.693 | 0 | 2.574 |
| 4 | 0.729 | 1.628 | 0 | 2.266 | 2.059 | 0 | 2.282 |
| 5 | 0.577 | 1.427 | 0 | 2.089 | 2.326 | 0 | 2.114 |
| 6 | 0.483 | 1.287 | 0.030 | 1.970 | 2.534 | 0 | 2.004 |
| 7 | 0.419 | 1.182 | 0.118 | 1.882 | 2.704 | 0.076 | 1.924 |
| 8 | 0.373 | 1.099 | 0.185 | 1.815 | 2.847 | 0.136 | 1.864 |
| 9 | 0.337 | 1.032 | 0.239 | 1.761 | 2.970 | 0.184 | 1.816 |
| 10 | 0.308 | 0.975 | 0.284 | 1.716 | 3.078 | 0.223 | 1.777 |
| 11 | 0.285 | 0.927 | 0.321 | 1.679 | 3.173 | 0.256 | 1.744 |
| 12 | 0.266 | 0.886 | 0.354 | 1.646 | 3.258 | 0.283 | 1.717 |
| 13 | 0.249 | 0.850 | 0.382 | 1.618 | 3.336 | 0.307 | 1.693 |
| 14 | 0.235 | 0.817 | 0.405 | 1.594 | 3.407 | 0.328 | 1.672 |
| 15 | 0.223 | 0.789 | 0.428 | 1.572 | 3.472 | 0.347 | 1.653 |
| 16 | 0.212 | 0.763 | 0.448 | 1.552 | 3.532 | 0.363 | 1.637 |
| 17 | 0.303 | 0.739 | 0.466 | 1.534 | 3.588 | 0.378 | 1.622 |
| 18 | 0.194 | 0.718 | 0.482 | 1.518 | 3.640 | 0.391 | 1.608 |
| 19 | 0.187 | 0.698 | 0.497 | 1.503 | 3.689 | 0.403 | 1.597 |
| 20 | 0.180 | 0.680 | 0.51 | 1.490 | 3.735 | 0.415 | 1.585 |
| 21 | 0.173 | 0.663 | 0.523 | 1.477 | 3.778 | 0.425 | 1.575 |
| 22 | 0.167 | 0.647 | 0.534 | 1.466 | 3.819 | 0.434 | 1.566 |
| 23 | 0.162 | 0.633 | 0.545 | 1.455 | 3.858 | 0.443 | 1.577 |
| 24 | 0.157 | 0.619 | 0.555 | 1.445 | 3.895 | 0.451 | 1.548 |
| 25 | 0.153 | 0.606 | 0.565 | 1.435 | 3.931 | 0.459 | 1.541 |

**Table 2: Factors for constructing variable control charts**

**[based on values from D. C. Montgomery – Introduction to Statistical Quality Control) 6th Edition 2009 Appendix VI]**

## Normality assumption

A common assumption when constructing control charts for continuous data (individuals or X-bar) is that data follows a normal distribution. The normality should be tested before using these charts. A common way to check for normality is to visually inspect the histogram and the quantile-quantile plot, as well as to conduct a normality test. The most used normality test is the Shapiro-Wilk test. If data are not normally distributed, a deeper understanding of the non-normality is necessary: are there outliers, are there trends in the data, are there two populations or is it another distribution? Often, data might be log-normally distributed, in which case a logarithmic transformation is necessary in order to normalize the data. Another common transformation is the reciprocal one, 1/X. The control charts should be constructed on the transformed data.

## CuSum & EWMA charts

### CuSum charts

Although the Variables/Shewhart chart is sensitive to sudden and large changes in measurement, it is ineffective in detecting small but persistent departure from the target or predefined value (bench mark). For this task, the CuSum chart is more appropriate.

CuSum is short for Cumulative Sums. As measurements are taken, the difference between each measurement and the bench mark value/process target ($\mu_0$) is calculated, and this is cumulatively summed up (thus CuSum):

$$c_i = \sum_{j=1}^{i} (x_j - \mu_0)$$

If the processes are in control, measurements do not deviate significantly from the bench mark, so measurements greater than the bench mark and those less than the bench mark averaged each other out, and the CuSum value should vary narrowly around the bench mark level. If the processes are out of control, measurements will more likely to be on one side of the bench mark, so the CuSum value will progressively depart from that of the bench mark.

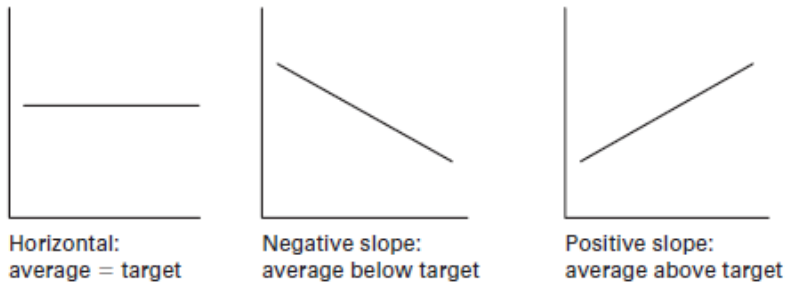



**Figure 7: Interpretation of CuSum charts**

CuSum can be used as a 'post-mortem' analysis of historical data, that may allow to determine the cause of unexpected changes in result.

## EWMA

Exponentially Weighted Moving Average (EWMA) chart, also referred to as a Geometric Moving Average (GMA) chart are a good alternative to the Shewart control chart when we want to detect small shifts. It acts in the same way as a CuSum chart.

Each point on a EWMA chart is the weighted average of all the previous subgroup means, including the mean of the present subgroup sample. The weights decrease exponentially going backward in time.

$$z_i = \lambda x_i + (1-\lambda) z_{i-1}$$

Where $0 < \lambda \leq 1$ is a constant and the starting value is the process target:

$$z_0 = \mu_0$$

If $\boldsymbol{\lambda}$ is close to 0, more weight is given to past observations. If $\boldsymbol{\lambda}$ is close to 1, more weight is given to present information. When $\boldsymbol{\lambda}$=1, the EWMA becomes the Individuals control chart. Typical values for $\boldsymbol{\lambda}$ are less than 0.25.

## The EWMA chart control limits

$$UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]}$$

$$\text{Center Line} = \mu_0 \qquad (1.11)$$

$$UCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]}$$

EWMA with $\boldsymbol{\lambda}$=0.05 or $\boldsymbol{\lambda}$=0.10 and an appropriately chosen control limit will perform very well against both normal and non-normal distributions, in contrast with individual charts that are very sensitive to non-normality.

### Process Capability Indices

Specification limits are used to evaluate process capability enabling a measure of how well the product meets customer expectations. Control limits are used to evaluate process stability. Unstable processes generally lead to failure to meet customer expectations.

Process capability refers to the performance of the process when it is operating in control. Two capability indices are usually computed: $C_p$ and $C_{pk}$. $C_p$ measures the potential capability in the process, if the process was centred (it does not take into account where the process mean is located relative to the specifications), while $C_{pk}$ measures the actual capability in a process (process can be off-centre). If a process is centred, then $C_p = C_{pk}$.

$$C_p = \frac{(USL - LSL)}{6\sigma}$$

$$C_{pk} = \min\left(C_{pu}, C_{pl}\right)$$

$$C_{pu} = \frac{(USL-\mu)}{3\sigma} \text{ and } C_{pl} = \frac{(\mu-LSL)}{3\sigma}$$

Where σ is estimated either by $\frac{\bar{R}}{d_2}$ when variables control charts are used in the capability studies or by the sample standard deviation s.

Typical values for $C_p$ and $C_{pk}$ are 0.5 or 1 for not capable processes, 1.33 and 1.67 for capable processes and >2 for highly capable processes.

An important assumption underlying the interpretation of $C_p$ and $C_{pk}$ is that the process output follows a normal distribution. If data are not normally distributed, one can transform the data to normalize it. Then work with the transformed data (and specifications!) to compute the indices. Commonly used transformations are logarithmic, ln(X), or reciprocal, 1/X.

## Control of discrete data SPC charts

Whenever the measured quantities for one item are not continuous but rather quality characteristics or count data, control chart for discrete data should be used. Usually, one would classify the inspected item into “conforming item” or “nonconforming item”. A nonconforming item is a unit of product that does not satisfy one or more of the specifications of the product (it contains at least one nonconformity). If more than one defect can be observed on the same unit, one can be interested in the number of nonconformities (defects) per unit, instead of the fraction nonconforming for a single nonconformity (defect).

## Control charts for single nonconformity: p-chart and np-chart

One can construct control charts for either fraction of non-conforming (p-chart) or total number of non-conforming units, if sample size is the same (np-chart)

### P-Charts – control chart for fraction nonconforming

Suppose m samples of sample size $n_i$ are available , let $\bar{n}$ be the average sample size:

$$\bar{n} = \frac{1}{m}\sum_{i=1}^{m} n_i$$

If the sample size is the same for each group, then $\bar{n} = n$.

The sample fraction nonconforming for sample *i* is defined as the ratio of the number of non-conforming units in the sample *i*, $D_i$, to the sample size $n_i$.

$$p_i = \frac{D_i}{n_i}$$

Suppose m samples are available, than the average fraction nonconforming is:

$$\bar{p} = \frac{\sum_{i=1}^{m} p_i}{m}$$

The distribution of the random variable $p$ can be obtained from the binomial distribution.

### The P-chart control limits

$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$
$$\text{Centre Line} = \bar{p} \qquad (1.12)$$
$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

Depending on the values of $\bar{p}$ and $n_i$, sometimes the lower control limit LCL<0. In these cases, we set LCL=0 and assume the control chart only has an upper limit.

## nP-charts

If the sample sizes for all samples are equal, one can also construct the control chart for the number nonconforming (*np*-control chart) instead of the fraction non-conforming.

The nP-chart control limits are

$$UCL = n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$$
$$\text{Centre Line} = n\bar{p} \qquad (1.13)$$
$$LCL = n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})}$$

Many commercial statistical programmes will produce *np* control charts with variable control limits based upon *n*.

## Discussion

If sample size is too large with respect to the number of nonconforming units (eg., 20 nonconforming units out of 500000), than the p-chart will not work properly because the control limits are inversely proportional to the sample size. Therefore they became very small and process will look out of control, as data plotted on the control chart will be out of control limits. If the sample size is the same (or approximately the same), one could use the individuals control charts instead, where one would plot the number of nonconforming units. If the sample size is significantly different from one sample point to another, then one could use a Laney p-chart[7].

Over dispersion exists when there is more variation in your data than you would expect based on a binomial distribution (for defectives) or a Poisson distribution (for defects). Traditional P charts and U charts assume that your rate of defectives or defects remains constant over time. However, external noise factors, which are not special causes, normally cause some variation in the rate of defectives or defects over time. Under dispersion is the opposite of over dispersion. Under dispersion occurs when there is less variation in your data than you would expect based on a binomial distribution (for defectives) or a Poisson distribution (for defects). Under dispersion can occur when adjacent subgroups are correlated with each other, also known as autocorrelation. For example, as a tool wears out, the number of defects may increase. The increase in defect counts across subgroups can make the subgroups more similar than they would be by chance. When data exhibit under dispersion, the control limits on a traditional P chart or U chart may be too wide. If the control limits are too wide, you can overlook special cause variation and mistake it for common cause variation.

[7] David B. Laney Quality Engineering, 14(4), 531–537 (2002) and see also, for example, Chin-liang Hung, M.S dissertation from Iowa State University, *Control Charts for Attributes: Some Variations*, 1997, http://www.public.iastate.edu/~wrstephe/HungCC.pdf

## Discrete data SPC charts: C and U charts

When more than one defect can be observed on the inspected unit, one will then be interested in the number of nonconformities per sample or average number of nonconformities per unit, instead of fraction of non-conforming or number of nonconforming units in a sample.

One can construct control charts for either the total number of nonconformities in a unit (c-chart) or the average number of nonconformities per unit (u-chart)

### C-Charts– control chart for number nonconforming

When we have a constant sample size, *n*, of inspection units from one sample to another, one can work with the total number of nonconformities per sample and construct the c-chart. The total number of nonconformities in a unit is represented on the chart:

$$\bar{c}_1 = \sum_{j=1}^{n} x_{ij}$$

where $x_{ij}$ is the number of defects for inspection unit *i* in sample *j*. The total nonconformities in a sample follow a Poisson distribution.

### The C-chart control limits

$$\begin{aligned} UCL &= \bar{c} + 3\sqrt{\bar{c}} \\ \text{Centre Line} &= \bar{c} \\ LCL &= \bar{c} - 3\sqrt{\bar{c}} \end{aligned} \qquad (1.14)$$

where $\bar{c}$ is the observed average number of non-conformities in a preliminary sample of *m* inspection units, n is the constant sample size and $x_{ij}$ is the number of defects for inspection unit *i*:

$$\bar{c} = \frac{\sum_{i=1}^{n} \bar{c}_i}{m}$$

If LCL yields a negative value, than LCL is fixed to 0.

### U-Charts

If the sample size is not constant and can vary from one sample to another, then one should work with the average number of nonconformities per unit of product instead of total number of nonconformities per sample and the u-chart is to be used, instead of a c-chart. Let the average number of nonconformities per unit be

$$\bar{u}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$$

Where $x_{ij}$ is the total nonconformities in a sample of $n_i$.

### U-chart control limits

$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}}$$
$$\text{Center Line} = \bar{u} \qquad (1.15)$$
$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}}$$

Where $\bar{u}$ represents the observed average number of nonconformities per unit in a preliminary data set of m inspection units, n is the sample size of the current inspected sample:

$$\bar{u} = \frac{\sum_{i=1}^{m} \bar{u}_i}{m}$$

# Trend Analysis for Stability Testing

## Overview

The purpose of this section is to provide guidance for generating, maintaining and monitoring trends of stability data by establishing trend limits calculated from existing historical stability data for pharmaceutical products stored at the recommended storage condition.

The purpose of trend analysis for stability data should be to detect if;

- a batch is out of trend with respect to historical batches

and

- one or more observations is out of trend within a batch

Although there are numerous approaches to trending stability data, two different models are presented here to generate stability trends. These two approaches that may be used are a simple linear regression model and the more sophisticated Random Coefficients Regression Model. These models are used to understand the degradation rates over time to support the expiration dating of the product.

To see if a specific batch is out of trend, a comparison of the slope of the batch under study with the slopes of the historical batches should be performed. A poolability test[8] may be used for this comparison, or improved statistical description of the historical behaviour and detection of an OOT batch can be obtained by estimating the slope of the historical batches and the new batch via the Random Coefficients Regression Model (with a fixed effect being the type of batch: historical or under study) and then use contrasts to make a test whether the difference between the slope of the historical batches and the slope of the new batch is different from zero or not. If the difference is significantly (at 0.05) different from zero, then the new batch is considered to be OOT. A minimum number of observations per batch needed for this analysis should be defined (e.g. 3 or 4 observations to determine a meaningful statistical trend based on product history and measurement variability, as 2 may not be sufficient; however, 2 time points may be sufficient to highlight an OOE).

To see if one observation is out-of-trend, a prediction interval for the batch under study should be constructed (without the observation under study), taking into account the variability from the historical batches (via a common error model between historical+batch under study). If the observation is outside the prediction interval, than it is considered as OOT. The data set must include a minimum of 3 lots with at least 4 time points per lot to start this analysis.

[8] For example as described in ICH Q1E

Data from multiple configurations maybe combined if there is a technical rationale or if equivalency of the configurations can be demonstrated.

Generate a preliminary graph of the test result versus storage time. Any data point that appears atypical or discrepant ***might*** be removed from the data set. Any data point removed must be identified and its removal justified in a final report.

The principles discussed here are in accordance with WHO TRS No. 953; Annex 2 Stability testing of active pharmaceutical ingredients and finished pharmaceutical products, 2009 and the ICH Guidance Q1A(R2) Stability Testing of New Drug Substances and Product.

The following papers should also be consulted for more detailed information;


1. Identification of Out-of-Trend Stability Results, A Review of Potential Regulatory Issue and Various Approaches, Pharma CMC Statistics and Stability Expert Teams, Mary Ann Gorko, Pharmaceutical Technology, 27 (4), 38–52, 2003
2. Identification of Out-of-Trend Stability Results, Part II PhRMA CMC Statistics, Stability Expert Teams, Pharmaceutical Technology, 2006
3. Methods for Identifying Out-of-Trend Results in Ongoing Stability Data, Adrijana Torbovska and Suzana Trajkovic-Jolevska, Pharmaceutical Technology Europe, June, 2013
4. Carter, R. L. and Yang, M. C. K. (1986). "Large Sample Inference in Random Coefficient Regression Models." Communication in Statistics Theory and Methods 15(8), 2507-2525
5. Chow, Shein-Chung, Statistical Design and Analysis of Stability Studies, Chapman & Hall/CRC Biostatistics Series, Boca Raton Fl, 2007
6. Dempster, A. P., Rubin, D. B. and Tsutakawa, R. K. (1981). "Estimation in Covariance Component Models." Journal of the American Statistical Association 76, 341 – 353
7. Laird, N. M. and Ware, J. H. (1982). "Random Effects Model for Longitudinal Data." Biometrics 38, 963 – 974
8. Searle, R. Shayle, Casella, G., McCulloch, C. E., Variance Components, John Wiley & Sons, Inc., New York, 1992


It is recommended that a Stability Subject Matter Expert (SME) advises on steps to be taken in case of insufficient and/or inconclusive stability data. SME must have sufficient education, training, and specific experience to provide such advice. The SMEs are required to have a good understanding of the stability data, analytical methods, as well as the strength, quality, identity, and purity of the product. In addition, a professional statistician can provide specific information and advice on statistical problems that arise in execution of procedures discussed in this Guideline.

## General principles of data selection and evaluation

The quality of any evaluation is only as good as the data and the appropriateness of the technique employed. Good data collection and selection practices are essential.

The following factors should be considered;

- For data selection, use stability data from historic product lots to calculate trend limits and conduct periodic review of stability data for new lots.
- Select validated quantitative stability indicating assays of each configuration of Drug Product or API/Drug Substance at a single real-time storage condition.
- Stability critical quality attributes may be selected for the statistical stability trending program using a risk based approach.
- Quantitative assays from the stability program may be justified and excluded from the statistical trending program.
- A stability test of a drug product/drug substance at given storage conditions intended for trending must have a minimum of three lots with at least four time points. Historical product knowledge, including development knowledge, should be considered. More data may be advisable (for example, in cases of high method variability, lot-to-lot variability, etc).
- Base trend assessment on all available time points for the selected lots.
- Use data values with more digits than reported in the product specifications (e.g. if the specification is greater than or equal to 90%, use stability data values to at minimum to, one significant figure more than your specification is recommended)).

### Establishing Trend Limits from Stability Data - Simplified Approach Using the Linear Regression Model

The basic procedure is as follows;

1. Plot the assay test data vs. storage time and fit a regression line using the simple linear least-squares-regression model. The unit for time is usually months.
2. Consult an SME or a statistician if there are unusual patterns or shifts in the stability graph. The statistician advises if an investigation is required.
3. Any observation that appears atypical or discrepant may be removed from the data set if it has an identified root cause. The removal of data must be justified in the trend report.
4. If the graph is obviously not linear, transform the X-axis, for example by taking a square or square root of the X-axis values. If the graph cannot be linearised, consult a statistician or an SME.
5. Fit a linear regression to the data, and plot the 99% regression and prediction curves for the stability trend limits.
6. In addition, calculate and plot the 99.5% confidence trend limits

## The model

The stability data consists of responses from some method collected over multiple lots over a period of time. Linear regression is used to analyse the stability data (y-axis) versus time (x-axis). The analysis is used to understand the relationship between the stability data and time and can be used to predict an expected stability result over time.

The data consists of *j* pairs of numbers; the stability response and the time when the data were collected. Denote this pair as $R_j, T_j$ , where and $R_j$ is the response and $T_j$ is associated the time point.

Note that the data may be from all lots on test simply put together and there is no identification of a response and a time point to a lot during the analysis.
Hence, if we have L lots, there will most likely be L pairs of numbers have time point = 0, one for each of the L lots. The corresponding response, $R_j$ , will be different depending on which lot was analysed.

The simple linear model we seek to fit is;

$$\hat{R}_j = b + mT_j \tag{1.16}$$

where $\hat{R}_j$ is the best fit estimate of the regression line, *b* is the intercept and *m* is the slope.

Let us assume that there are *N* data pairs. We can calculate the mean response, $\overline{R}$ , and the mean time $\overline{T}$ from;

$$\overline{R} = \frac{1}{N}\sum_{j=1}^{N} R_j \tag{1.17}$$

and

$$\overline{T} = \frac{1}{N}\sum_{j=1}^{N} T_j \tag{1.18}$$

The sum of squares of the differences from these means from the actual values, $S_R$ and $S_T$ are then readily calculated from

$$S_R = \sum_{j}^{N}\left(R_j - \overline{R}\right)^2 \tag{1.19}$$

and similarly

$$S_T = \sum_{j}^{N}\left(T_j - \overline{T}\right)^2 \tag{1.20}$$

the cross product term, $S_{RT}$ , is found from

$$S_{RT} = \sum_{j}^{N}\left(R_j - \overline{R}\right)\left(T_j - \overline{T}\right) \tag{1.21}$$

The slope of the regression line, *m*, is simply the ratio of the two sums of squares

$$m = \frac{S_{RT}}{S_T} \quad (1.22)$$

The intercept of the regression line, *b*, where the x axis is zero is calculated using *m* and the two mean values $\overline{R}$ and $\overline{T}$

$$b = \overline{R} - m\overline{T} \quad (1.23)$$

The degree of correlation, *r*, is found from

$$r = \sqrt{\frac{S_{RT}^2}{S_R S_T}} \quad (1.24)$$

Note that this is not a measure of linearity but of correlation.

The errors associated with the slope and the intercept can now be calculated from the mean square error, *MSE*,

$$MSE = \frac{S_R - mS_{RT}}{N-2} \quad (1.25)$$

giving the standard error of the slope, $SE_m$, and the standard error of the intercept, $SE_b$, in equations (1.26) and (1.27)

$$SE_m = \sqrt{\frac{MSE}{S_T}} \quad (1.26)$$

$$SE_b = \sqrt{MSE\left(\frac{\overline{T}^2}{S_t} + \frac{1}{N}\right)} \quad (1.27)$$

The confidence intervals at 99% confidence for the slope, $CI_m$ , and the intercept, $CI_b$ , can be calculated from equations (1.28) and (1.29).

$$CI_m = \pm t_{(0.01,N-2)} SE_m \quad (1.28)$$

and

$$CI_b = \pm t_{(0.01,N-2)} SE_b \quad (1.29)$$

The root mean square error (standard deviation), *RMSE*, is found by taking the square root of *MSE* from equation (1.25).

The confidence intervals for both regression and prediction are calculated from

$$CI_{\text{REG}} = \pm t_{(0.01,N-2)} RMSE \sqrt{\frac{1}{N} + \frac{\left(T_j - \overline{T}\right)^2}{S_T}} \quad (1.30)$$

and

$$CI_{\text{PRE}} = \pm t_{(0.01,N-2)} RMSE \sqrt{1 + \frac{1}{N} + \frac{\left(T_j - \overline{T}\right)^2}{S_T}} \quad (1.31)$$

Using these values the confidence contours for regression and prediction are calculated for each of the $T_j$ time points.

Upper and lower 99.5% confidence acceptance trend limits (*TL*) can be calculated from

$$TL = mT + b \pm t_{(0.005,N-2)} RMSE \sqrt{1 + \frac{1}{N}} \quad (1.32)$$

## Establishing Trend Limits from Stability Data -; a more advanced Random Coefficients Regression model approach

### Overview

The general random coefficients regression model is a flexible model that allows for multivariate inputs and covariates. The discussion below applies the random coefficients regression model to stability data in which for each lot there exists a single response at each time point. Thus, a simplified version of the general Random Coefficients Regression Model (RCRM) is considered in which only an intercept and slope are present in the model.

### The model

Assume that the trend limits are to be established based on stability data performed on N lots of product.

Lot $\ell$ is tested at the $n_\ell$ time points $t_{\ell,1},\cdots,t_{\ell,n_\ell}$ with corresponding responses $y_{\ell,1},\cdots,y_{\ell,n_\ell}$ The Random Coefficients Regression (RCR) model can be written (Carter & Yang 1986 reference 4 on page 34)

$$y_{\ell,j} = a_\ell + b_\ell \times t_{\ell,j} + \varepsilon_{\ell,j} \text{ where } \ell = 1,\cdots,N \quad j = 1,\cdots,n_\ell \quad (1.33)$$

The coefficients $a_\ell$ and $b_\ell$ are the lot-specific intercept and slope for the degradation rate of lot $\ell$. It is assumed that the coefficients have a bivariate normal distribution:

$$\begin{pmatrix} a_j \\ b_j \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \Sigma\right) \quad (1.34)$$

The error terms $\varepsilon_{\ell,j}$ are assumed to be independent, identically distribution from a normal distribution with mean 0 and variance $\sigma^2$. It is further assumed that the error terms are independent of the coefficients $a_\ell$ and $b_\ell$.

Several properties of the RCR Model make it a suitable model for stability trend data:

- *The parameters in the model are easily interpreted*.
  The intercept, α, represents the response at product release averaged over the manufacturing process. When testing for homogeneity of slopes, the slopes of the individual lots are compared to a common slope. The slope, β, represents this common slope. The error variance, $\sigma^2$, estimates the variance of the method. The component of the covariance matrix, Σ, corresponding to the variance of the intercept represents the variance of the response observed at product release due solely to the manufacturing process. (note: this variance excludes any variation due to the method.) Finally, the component of the covariance matrix, Σ, corresponding to the variance of the slope represents the variance of the degradation rates among the stability lots under consideration. The variance equals 0 when all lots degrade at the same rate and corresponds approximately to passing a homogeneity of slopes test.
- *The intercept and slope are allowed to vary between lots*.
  There can be differences in the release response at product release between different lots due to manufacturing variability. Additionally, examples have been observed in which the degradation rate varies between product lots. The RCR model (1) assumes both the intercept and slope to be random effects, thus allowing for different intercepts and slopes for each lot.
- *There are few restrictions on the design space.*

  On-going stability studies are usually designed to have data collected at fixed time points. However, it is possible to for time points to be missing, or for certain lots to be still under study, in which case later time points have not been collected. The model (1) allows complete flexibility in the collection of data, subject to the minor constraint that at least 3 time points must be collected per lot.

There are two constraints however.

- *The degradation rate is assumed to be linear.*
  Most degradation rates observed are sufficiently linear to permit fitting a linear model. For those degradation rates that are not linear, it may be possible to linearize the data by applying a transformation to the time variable. Detecting and remedying non-linearity is required prior to applying the RCR Model. Additionally, the scales for the response and time axis often differ by several orders of magnitude. Disparate ranges in the time and response axes can result in numerical instability. The time and response variables may be normalized prior to analysis, the trend limits determined, and then results re-expressed in the original scale.

- *The error terms, $\varepsilon_{i,j}^2$, are assumed to have constant variance across lots and time*.
  The error terms, $\varepsilon_{i,j}^2$, represent the variability of the method. There is no reason to suspect a priori that the variability of the method depends on the lot tested. Thus, it is not unreasonable to assume that the method variance is the same across all time points and stability lots.

## Parameter estimation

There are multiple approaches to estimating the parameters of the RCR model. See references 4, 6 and 7 on page 34. Trending analysis may be performed across multiple analysis platforms for example, SAS, R, JMP, Minitab.

An approach that requires only algebraic calculations and not numerical optimization or iterative re-weighting schemes is a modified version of the estimation scheme described by Carter and Yang (4 on page 34).

The estimation of the parameters of the RCR model is performed in three steps:

Step 1) A simple linear regression is fitted to each individual lot of stability data as used in the simple approach

Step 2) The covariance matrix, $\Sigma$, and error variance, $\sigma^2$ are estimated using the regression results obtained in Step 1)

Step 3) The mean vector, $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, is estimated as a weighted average of the individual slopes and intercepts obtained in Step 1, with weights depending on the estimates obtained in Step 2

Once the parameters of the RCR model have been estimated, an approximate prediction interval can be constructed at any time point.

### Step 1) Fitting a simple linear regression

Select a lot, k. The data associated with this lot is $\left(t_{k,j}, y_{k,j}\right),\ j = 1, \cdots, n_k.$

Let $X_k$ be the design matrix for lot k

$$X_k = \begin{pmatrix} 1 & t_{k,1} \\ \vdots & \vdots \\ 1 & t_{k,n_k} \end{pmatrix} \tag{1.35}$$

Define a normalised matrix $M_k$ such that;

$$M_k = \left(X_k' X_k\right)^{-1} \tag{1.36}$$

Fit a simple linear regression to the data to obtain:

$a_k$; the estimated intercept for lot *k*
$b_k$; the estimated slope for lot *k*
$MSE_k$; the Mean Square Error of Regression for lot *k*
$df_k (= n_k - 2)$; the degrees of freedom associated with $MSE_k$

**Step 2) Estimating the error variance and covariance matrix**
Estimate the pooled mean square error by

$$\hat{\sigma}^2 = \frac{\sum df_j \cdot MSE_j}{\sum df_j} \tag{1.37}$$

Let S be the sample covariance matrix of the estimated intercepts and slopes.
Define

$$\bar{M} = \frac{1}{N}\sum M_k \tag{1.38}$$

The estimated covariance matrix, $\hat{\Sigma}$, for the intercept and slope is given by:

$$\hat{\Sigma} = S - \hat{\sigma}^2 \cdot \bar{M} \tag{1.39}$$

The estimated covariance matrix, $\hat{\Sigma}$, given by equation (1.39) may not be positive definite. Carter and Yang provide a modification to equation (1.39) to insure that the estimate $\hat{\Sigma}$ is positive definite.

An alternate approach is that if either the slope or intercept variance is negative, the estimate along with the estimated covariance are replaced with 0. This is a standard approach for negative variance estimates [4.10] and is equivalent to converting a random effect into a fixed effect in the model.

**Step 3) Estimating the mean vector**
Define

$$W_k = \left\{\hat{\Sigma} + \hat{\sigma}^2 \cdot M_k\right\}^{-1} \text{ and} \tag{1.40}$$

$$\Omega = \left(\sum W_k\right)^{-1} \tag{1.41}$$

The estimated mean vector is given by:

$$\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix} = \Omega\left(\sum W_k \cdot \begin{pmatrix}a_k\\ b_k\end{pmatrix}\right) \tag{1.42}$$

## Constructing the approximate 99% Prediction Interval

For any time point, t, an approximate 99% prediction interval for some constant k is given by [4.6]:

$$\hat{a} + \hat{b}t \pm k \cdot \sqrt{\begin{pmatrix}1\\ t\end{pmatrix}' \left\{\hat{\Sigma} + \frac{1}{N}\Omega\right\}\begin{pmatrix}1\\ t\end{pmatrix} + \hat{\sigma}^2} \tag{1.43}$$

Carter & Yang 1986 (reference 4 on page 34) use a *t* statistic with degrees of freedom estimated by Satterthwaite's approximation for *k*. An alternate approach which provides conservative trend limits is to use the 99.5 percentile of the standard normal distribution for the constant *k*; replacing the *t* percentile with a normal percentile results in more stringent trend limits and hence reduces the risk of not detecting out-of-trend data.

## Process flow for evaluating trending of stability data

1. Compare a stability test result with the trend limits
2. If a stability test result is out of trend limits, evaluate the cause for the out of trend result as defined within the quality system. For example see a process flow in Figure 8.
3. The level of the investigation for out of trend results depends on the frequency (single out of trend point, multiple out of trend results), risk of future out of specification result, precision of stability test, product history and known characteristics (consider a risk based assessment), and potential impact to patient safety and product efficacy. Test results for other parameters should be considered. Be alert to process improvements and manufacturing changes.
4. An out of trend result should not automatically require a new stability time point.
5. Within a single stability lot, if the value is significantly different from the time zero (degradation), compare the value to the previous time point(s). If the value is significantly different from the expected value (OOE) and the method performance, the value is suspect and should be evaluated as an out of trend value.
6. In cases where there are no established stability trend limits, evaluate the suspect value by comparing to known historic stability data. The result may be out of trend based on the historic pattern.
7. Periodic reassessment of trend limits is required. This reassessment will help detect drifts or other changes over time. Additional data will likely change the trend limits.
8. Assess prediction intervals according to a defined interval (annually, or at a minimum of every 3 years, for example) to confirm stability trend limits. Include appropriate graphs, investigations, and/or supporting documentation in the annual evaluation.
9. Assessment of trend limits may also be used to evaluate site or post-change differences

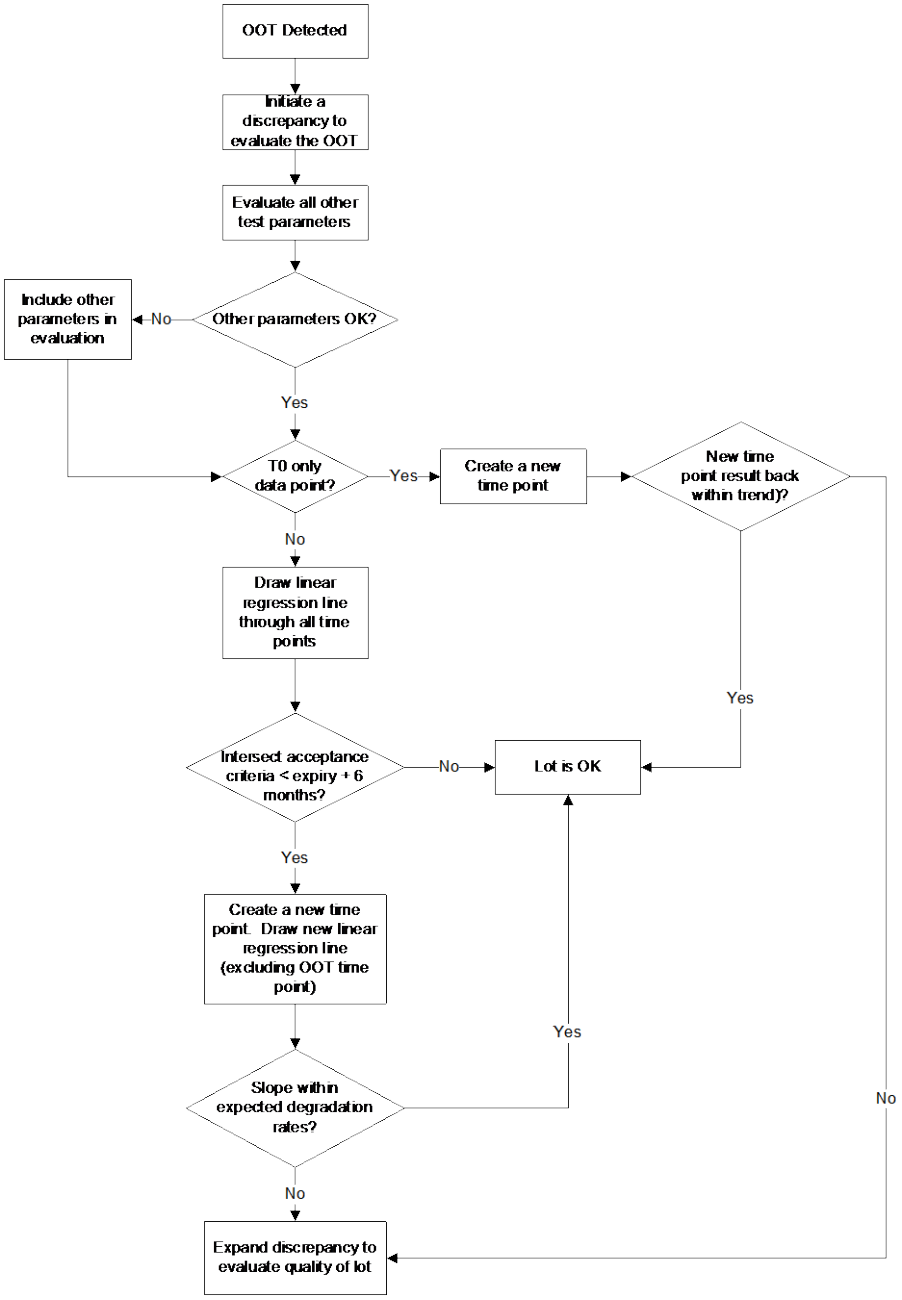



**Figure 8 Example of a Process Flow for OOT stability test results**

# Trend Analysis for Investigations

On many occasions, laboratories are faced with historical data requiring analysis (post mortem) after the discovery of a problem. One of the questions asked of the data is usually has anything changed and, if so, when did it happen.

Although the Shewhart chart is sensitive to sudden and/or large changes in a measurement sequence, it is ineffective in detecting small but persistent departure from a bench mark. This bench mark may be a target or specification value or, more commonly, in post mortem investigations the mean of the data set. The method of choice in this situation is to employ post mortem CuSum analysis. This technique was developed in the 1950s by Imperial Chemical Industries Ltd[9] and this technique is also described in an obsolescent British Standard, BS5703 Part 2 recently replaced an ISO norm[10].

This is a simple but powerful technique which is not as widely known as it should be. As the name CuSum implies it is merely the cumulative sum of differences from a bench mark.

The objective of this technique is to;

- detect changes from successive differences
- estimate when the change occurred
- estimate the average value before and after the change.

It is important to note that this technique attempts to identify if a special cause variation has occurred and when it happened not why it happened.

## Theory of post mortem CuSum analysis

The CuSum is calculated from the successive differences from a bench mark. Assuming this bench mark is the mean of the data set, $\bar{X}$, then, for i data points, the value of the CuSum for the i$^{th}$ data point is given by

$$S_i = S_{i-1} + (X_i - \bar{X}) \quad (1.44)$$

The last value of the CuSum is always zero.
If a process is under statistical control ie contains no special cause variation, the CuSum from the mean will only have common cause variation ie noise. Therefore, a plot of this CuSum with respect to time (or batch) will be a straight line parallel to the X axis.

However if there is a downward slope this would indicate that the process average was below the benchmark and conversely an upward or positive slope would indicate that process average was above the benchmark. The steeper

[9] Cumulative Sum Techniques, ICI Monograph N° 3, Oliver and Boyd, 1964

[10] BS ISO 7870-4:2011; Control charts. Cumulative sum charts

the slope the greater is the difference. Hence the objective is to detect changes in slope thereby partitioning the data into segments. The key aspect is to determine if a slope change is due to a real effect or merely chance through noise. The distance between the successive real turning points is called the span.

In an ideal noise free world, the interpretation of the CuSum plot would be trivial as illustrated in Figure 9.

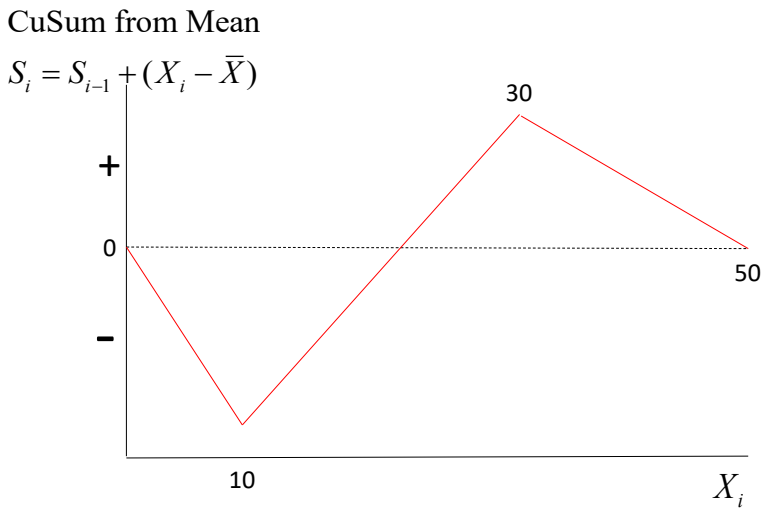



**Figure 9 Idealised CuSum plot**

The start and endpoints on a CuSum from the mean are always zero. between the 1st point and the 10th point the slope is negative indicating that the process average is less than the mean and also between 30th point and the 50th point. Between the 10th point and the 30th point the reverse is true.

It is important to recognise that this post mortem technique is not an exact statistical evaluation but rather a method of indicating where to look for change.

One method used for post mortem analysis is as follows;

1. Calculate the sum of the squares of the differences between successive values and dividing it by 2(n-1). The localised standard deviation of the data is the square root of this value. The reason for calculating this localised standard deviation is to minimise the effects of any special cause variation which would increase the value and make detection of these special causes less sensitive.

   The successive differences in values are given by $\Delta_i = X_i - X_{i+1}$ and therefore $s_L = \sqrt{\frac{\sum_{i=1}^{i=n} \Delta_i^2}{2(n-1)}}$
2. Find by inspection the absolute value of the maximum CuSum for the data set and note the index number
3. Calculate the test statistic

   $$\frac{|CuSum_{\max}|}{s_L}$$
4. Compare this value with the critical value for the span (Table 3). The span is given by the number of data points within each region. The first value for the Span is the total number of data points in the CuSum.
5. If this change point is statistically significant, divide the CuSum plot into two groups by drawing two lines from the maximum CuSum to the extremities of the plot. These are the new baselines.
6. Inspect these regions for the next largest CuSum to be tested.
7. If appropriate, recalculate the CuSum and the localised standard deviation for each region.
8. Repeat steps 1 to 7 until no significant statistically turning points remain.
9. Draw the Manhattan plot for each of the regions identified. There will be n+1 regions from n turning points. The Manhattan plot value is based upon the mean value for each region identified in the CuSum analysis.

This process will be illustrated by example in Appendix 4.

| | Critical value | | | Critical value | |
|---|---|---|---|---|---|
| **Span** | **95%** | **99%** | **Span** | **95%** | **99%** |
| 2 | 1.6 | 2.1 | 14 | 4.6 | 5.6 |
| 3 | 2.0 | 2.5 | 15 | 4.8 | 5.8 |
| 4 | 2.3 | 2.9 | 20 | 5.6 | 6.8 |
| 5 | 2.7 | 3.3 | 25 | 6.0 | 7.3 |
| 6 | 3.0 | 3.6 | 30 | 6.7 | 8.0 |
| 7 | 3.2 | 4.0 | 40 | 7.8 | 9.3 |
| 8 | 3.5 | 4.3 | 50 | 8.6 | 10.4 |
| 9 | 3.7 | 4.6 | 60 | 9.5 | 11.3 |
| 10 | 3.9 | 4.9 | 70 | 10.3 | 12.2 |
| 11 | 4.1 | 5.1 | 80 | 10.8 | 12.9 |
| 12 | 4.3 | 5.3 | 90 | 11.3 | 13.6 |
| 13 | 4.5 | 5.5 | 100 | 11.8 | 14.3 |

**Table 3 Critical values for Post Mortem CuSum analysis. Values derived from the nomogram (Figure 12) of British Standard BS 5703 Part 2 (1980) which was generated using numerical simulation.**

## Appendix 1: Technical Glossary

| TERM | DEFINITION |
|---|---|
| Acceptance Criteria | Numerical limits, ranges, or other suitable measures for acceptance of test results . |
| Acceptance Limit | The maximum amount of carryover of one product or cleaning agent allowed in a batch or dose. |
| Acceptance Sampling | Inspection used to determine whether a batch conforms or not to visual inspection acceptance criteria. |
| Accuracy | The closeness of agreement between the value which is accepted either as a conventional true value or an accepted reference value and the value found. |
| Action Limit/Action Level | A level that, when exceeded, indicates a drift from normal operating conditions. Action limits are based on design criteria, regulatory/industry standards, and intended use of the area. |
| Adverse Trend (AT) | A continuing deviation from normal "expected" process, product or quality performance characteristic, that has potential severity impact on safety, purity, efficacy or quality of the intended product function. |
| AQL (Acceptance Quality Limit) | Quality level that is the worst tolerable process average when a continuing series of lots is submitted for acceptance sampling. |
| AQL Inspection | Statistical inspection by attributes based on AQL. |
| Attribute Data | Data that consist of counts (i.e. number of defectives in a lot, pass or fail, yes or no) of defects or defectives in a lot. Typically counts of defective lots or of defects within lots are used. |
| Calibration | The set of operations which establish, under specified conditions, the relationship between values indicated by a measuring instrument or measuring system, or values represented by a material measure, and the corresponding known values of a reference standard. |
| Centre Line (CL) | Mean value of the control chart statistic |
| Control Charts | Control charts are a graphical method for comparing information from samples representing the current state of a process against limits established after consideration of inherent process variability. Their primary use is to provide a means of evaluating if a process is or is not in a "state of statistical control". |
| Control Limits | Control limits are used as criteria for signaling the need for assessment or for judging whether a set of data does or does not indicate a "state of statistical control". · Lower Control Limit (LCL) – Minimum value of the control chart statistic that indicates statistical control · Centre Line (CL) – Mean value of the control chart statistic · Upper Control Limit (UCL) – Maximum value of the control chart statistic that indicates statistical control |
| Critical Process Parameter (CPP) | A process parameter whose variability has an impact on a critical quality attribute and therefore should be monitored or controlled to ensure the process produces the desired quality. |
| Critical Quality Attribute (CQA) | A physical, chemical, biological or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality. |

| TERM | DEFINITION |
|---|---|
| Defect | A departure of a quality characteristic that results in a product, process or service not satisfying its intended usage requirements. |
| Failure Mode | The manner by which a failure is observed; it generally describes the way the failure occurs. |
| Histogram | A histogram is a graphic display (bar chart) of the frequency distribution (showing shape, location and variation on a number line) of a data set. The x-axis is divided into equal size "bins" or segments of the numerical data. The y-axis is the count of occurrences of data in each segment of the data. The height of the bar for each segment is proportional to the frequency of occurrence of sample data falling into that category. Categories must not overlap and must be equal in size. |
| Individual (I) Chart | The individual chart plots each measurement as a separate data point |
| Individual Moving Range (I-MR) Chart | The I-MR chart is made up of an individual chart and moving range chart. |
| In-Process Controls | Checks performed during production in order to monitor and, if appropriate, to adjust the process and/or to ensure that the intermediate or API conforms to its specifications. |
| Inspection by Attributes | Inspection where the unit is classified as conforming or nonconforming with respect to a specified requirement or set of specified requirements. |
| Inspection Level (AQL based inspections only) | Relationship between lot or batch size and the sample size. |
| Intermediate precision | Variation in the measurements made by different operators and/or instruments and/or times in one laboratory from the same sample. |
| Key Performance Indicator (KPI) | A measured or calculated attribute (typically characterizing the output of a process step) that is indicative of whether a process step has met its goal. KPIs are routinely measured and should be trended. |
| Lifecycle | All phases in the life of a product from the initial development through marketing until the product's discontinuation |
| Lower Control Limit (LCL) | Minimum value of the control chart statistic that indicates statistical control. |
| Moving Range (MR) Chart | The MR chart plots the difference between two consecutive data points as they come from the process in a sequential order. |
| Out of Control Limit | Single result that is markedly different from others in a series as confirmed by statistical evaluation, i.e., one point beyond established control limits |
| Out of Expectation (OOE) Result | OOE results for the purposes of this document are anomalous, unexpected or unusual findings, that have not been classified as either OOS or OOT, for both qualitative or quantitative testing. |
| Out of Specification (OOS) | All reportable results not meeting established specifications. A confirmed reportable value that is outside the acceptable specification criteria as stated in the product or analyte specification (e.g. CoAs, USP). |
| Out of Trend (OOT) | A test result or pattern of results that are outside of pre-defined limits. |
| Outlier | A single result which is markedly different from others in a series as confirmed by statistical evaluation. |

| TERM | DEFINITION |
|---|---|
| Prediction Interval (PI) | An interval that contains a future observation with a predetermined confidence usually 95% or 99%. |
| Process Capability Index | A statistical measure of process capability. Process Capability Indices are measured in terms of proportion of the process output that is within product specification or tolerance. |
| Process Capability | Process capability can be defined as the natural or inherent behaviour of a stable process that is in a state of statistical control. It is a statistical estimate of the outcome of a characteristic from a process that has been demonstrated to be in a state of statistical control. |
| Process Parameters | The variables or conditions of the manufacturing process or in an in-process test on which a processing decision is based. |
| Random Sample | A random sample is defined as one in which the entire population has an equal chance of being selected to ensure representative sampling across the batch. |
| Range | Interval comprising an upper and lower range of an attribute or variable |
| Real-Time Trend Analysis | Trend analysis that is performed on applicable tests prior to lot release and as soon as practically possible after each test result is produced and analyzed. |
| Repeatability | Variation in measurements obtained with one measuring instrument when used several times by an operator while measuring the identical characteristic on the same part or parts from the same sample. |
| Reproducibility | Variation in the measurements made by different operators and/or instruments and/or times and in different laboratories from the same sample. |
| Robustness | The measure of a method's capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage. |
| Sample | A portion of supply chain material collected according to a defined sampling procedure. |
| Sampling Inspection | The inspection of products using one or more samples (e.g. simple, double, multiple, sequential, skip-lot, etc.). |
| Sampling Plan | Combination of sample size(s) and associated lot/batch acceptability criteria. |
| Sampling Size | The number of items statistically selected from a defined population. Sample size is always an integer value. |
| State of Control | A condition in which the set of controls consistently provides assurance of continued process performance and product quality. |
| State of Statistical Control | A process is considered to be in a "state of statistical control" if it is affected by random (common) causes only, i.e. if no extraordinary, unexpected, special causes have entered the system. |
| Tolerance Intervals (TI) | An interval within which a stated proportion of a population will lie at a stated confidence level. |

| TERM | DEFINITION |
|---|---|
| Trend Drift | Six consecutive points all increasing or decreasing |
| Trend Limits | Upper and lower limits for evaluating potential trends; calculated with prediction intervals. |
| Trend Shift | Nine consecutive points on the same side of the centre line |
| Trend | An evaluation for significant changes in the data over time. |
| Trending | Trending is the search for significant changes in the data over time. Unplanned or unexplained trends indicate lack of consistency and stability. |
| Upper Control Limit (UCL) | Maximum value of the control chart statistic that indicates statistical control. |
| Western Electric/Nelson Rules | Decision rules for detecting "out-of-control" or non-random conditions on control charts. |

# Appendix 2: Example of SPC for Continuous Data; a Moving Range (MR) Shewhart Chart for individual data points

Suppose we have a run of 84 individual data points from an in-process control. These data are shown in Table 4

| # | Value | # | Value | # | Value | # | Value |
|---|---|---|---|---|---|---|---|
| 1 | 99.88% | 22 | 102.52% | 43 | 98.33% | 64 | 99.95% |
| 2 | 101.19% | 23 | 95.34% | 44 | 103.19% | 65 | 106.64% |
| 3 | 98.58% | 24 | 98.49% | 45 | 96.29% | 66 | 104.26% |
| 4 | 97.93% | 25 | 96.82% | 46 | 107.18% | 67 | 105.18% |
| 5 | 101.93% | 26 | 100.58% | 47 | 94.38% | 68 | 105.10% |
| 6 | 106.44% | 27 | 98.40% | 48 | 96.11% | 69 | 100.95% |
| 7 | 100.66% | 28 | 100.53% | 49 | 109.52% | 70 | 101.70% |
| 8 | 99.77% | 29 | 99.67% | 50 | 105.73% | 71 | 102.23% |
| 9 | 103.25% | 30 | 98.82% | 51 | 100.15% | 72 | 97.86% |
| 10 | 105.89% | 31 | 95.38% | 52 | 101.01% | 73 | 100.32% |
| 11 | 100.41% | 32 | 98.07% | 53 | 103.33% | 74 | 98.84% |
| 12 | 102.38% | 33 | 98.67% | 54 | 98.41% | 75 | 105.25% |
| 13 | 103.75% | 34 | 96.70% | 55 | 98.99% | 76 | 98.60% |
| 14 | 101.52% | 35 | 99.41% | 56 | 100.70% | 77 | 104.68% |
| 15 | 96.78% | 36 | 97.70% | 57 | 98.99% | 78 | 105.12% |
| 16 | 95.26% | 37 | 96.79% | 58 | 101.85% | 79 | 104.73% |
| 17 | 101.68% | 38 | 99.85% | 59 | 101.86% | 80 | 99.01% |
| 18 | 99.45% | 39 | 96.73% | 60 | 106.11% | 81 | 102.41% |
| 19 | 103.26% | 40 | 99.08% | 61 | 99.29% | 82 | 100.36% |
| 20 | 102.58% | 41 | 97.45% | 62 | 103.84% | 83 | 104.18% |
| 21 | 99.99% | 42 | 100.64% | 63 | 97.36% | 84 | 106.30% |

**Table 4 Data for MR Shewhart Chart (one more significant figure is given for statistical purposes)**

The mean value is 100.79. Each moving range is calculated from equation (1.3) and the average moving range calculated $\overline{MR}$. The control limits and the centre line are found from Equation (1.5) as shown below.

$$\begin{aligned} UCL &= \bar{x} + 3\frac{\overline{MR}}{d_2} = 100.79 + 3(\frac{3.29}{1.128}) = 100.79 + 8.76 = 109.55 \\ \text{Centre Line} &= \bar{x} = 100.79 \\ UCL &= \bar{x} - 3\frac{\overline{MR}}{d_2} = 100.79 - 3(\frac{3.29}{1.128}) = 100.79 - 8.76 = 92.03 \end{aligned} \qquad (1.45)$$

The value of $d_2$=1.128 is from Table 2 and $\overline{MR}$ =3.29 from Figure 11.

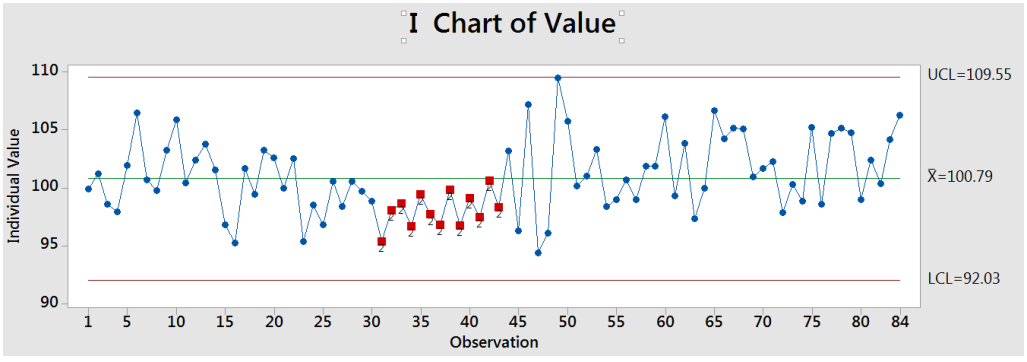



**Figure 10 Individual data from Table 4 (Minitab 17)**

It is apparent that WECO Rule 4 failures were noted at point 31 on onwards to point 43 indicating a process shift.

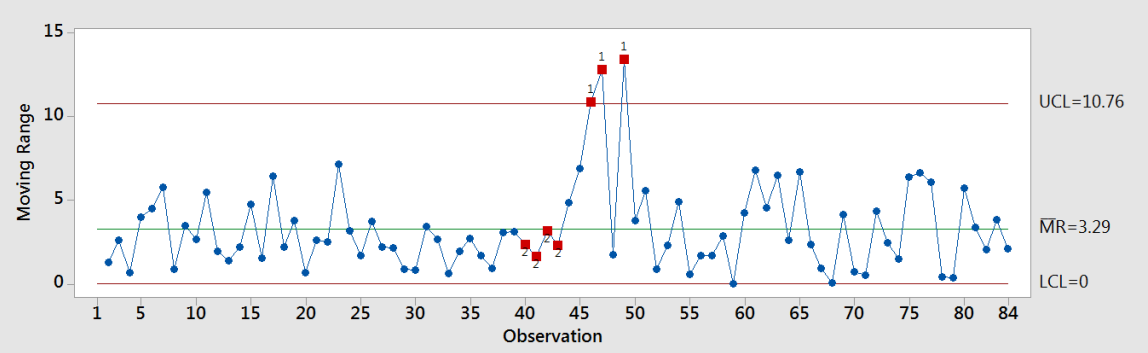



**Figure 11 Moving Range plot for data from Table 4 (Minitab 17)**

A similar feature is noted in Figure 11 for the Moving Range but at a slightly later time, observations 40 to 43, and, in addition 3 points, which exceed WECO rule 1 at 46, 47 and 49 indicating excessive dispersion. This would indicate that significant trends were apparent between observations 31 to 49 before the process return

## Appendix 3: Example of SPC for continuous data Xbar and R

Suppose that the data shown in Table 4 were actually collected in 21 sub groups of n=4 at each time point Allowing the mean and range to be calculated for each subgroup. It is now possible to plot the mean and range charts

| Subgroup | X Bar | R |
|---|---|---|
| 1 | 99.39% | 3.26% |
| 2 | 102.20% | 6.67% |
| 3 | 102.98% | 5.48% |
| 4 | 99.33% | 8.49% |
| 5 | 101.74% | 3.81% |
| 6 | 99.09% | 7.18% |
| 7 | 99.08% | 3.75% |
| 8 | 97.99% | 4.29% |
| 9 | 98.12% | 2.71% |
| 10 | 98.11% | 3.12% |
| 11 | 99.90% | 5.74% |
| 12 | 98.49% | 12.79% |
| 13 | 104.10% | 9.37% |
| 14 | 100.36% | 4.92% |
| 15 | 102.20% | 7.12% |
| 16 | 100.11% | 6.48% |
| 17 | 105.29% | 2.38% |
| 18 | 100.68% | 4.37% |
| 19 | 100.75% | 6.64% |
| 20 | 103.39% | 6.11% |
| 21 | 103.31% | 5.95% |

**Table 5 Mean & range subgroup data from Table 4**

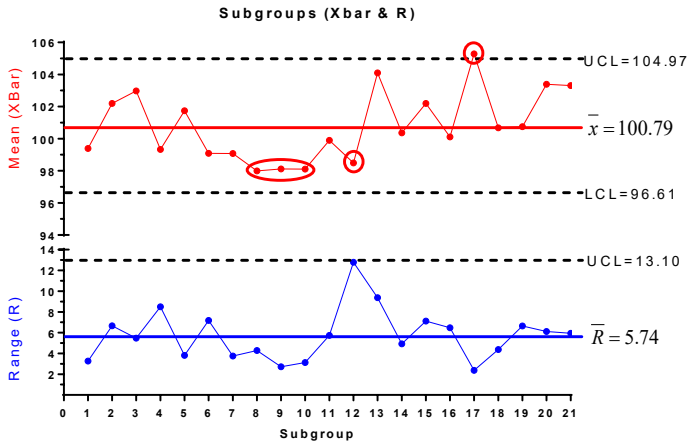



**Figure 12 Mean and Range Chart from Table 5**

It is now readily apparent that the range is in control but the mean is not as a subgroups 8, 9, 10 & 12 there are failures of WECO Rule 3 and WECO Rule 1 at point 17. These findings are consistent with those of the MR chart for individuals in Appendix 2.

The control limits for the mean are calculated from;

$$UCL = \bar{\bar{x}} + A_2\bar{R} = 100.79 + 0.729(5.74) = 100.79 + 4.18 = 104.97$$
$$\text{Centre Line} = \bar{\bar{x}} = 100.79 \qquad (1.46)$$
$$LCL = \bar{\bar{x}} - A_2\bar{R} = 100.79 - 0.729(5.74) = 100.79 - 4.18 = 96.61$$

and for the range (Note that the LCL for the range is not calculated as the subgroup size n is too small;

$$UCL = D_4\bar{R} = 2.282(5.74) = 13.10$$
$$\text{Centre Line} = \bar{R} = 5.74 \qquad (1.47)$$
$$LCL = D_3\bar{R} = 0 \text{ as n<7}$$

These are from equations (1.7) and (1.8) with the values for n=4 of the constants $A_2$ and $D_4$ from Table 2.

# Appendix 4: Example of investigation of continuous data; Post mortem CuSum analysis

By way of example, assume that we have 50 assay results from batches of a drug substance for an impurity and we wish to see if there have been any changes. The initial plot and the CuSum from the mean is shown in Figure 13. The CuSum is calculated from equation (1.48). The mean value for the impurity data is 0.45(4).

## Data Set

| Batch *i* | % impurity | Difference from mean $\delta_i = X_i - \bar{X}$ | CuSum $S_i$ $S_i = S_{i-1} + \delta_i$ |
|---|---|---|---|
| 1 | 0.61 | 0.15 | 0.15 |
| 2 | 0.80 | 0.35 | 0.50 |
| 3 | 0.61 | 0.15 | 0.65 |
| 4 | 0.41 | -0.05 | 0.61 |
| 5 | 0.71 | 0.25 | 0.86 |
| 6 | 0.61 | 0.15 | 1.01 |
| 7 | 0.51 | 0.05 | 1.07 |
| 8 | 0.70 | 0.25 | 1.31 |
| 9 | 0.60 | 0.15 | 1.46 |
| 10 | 0.61 | 0.16 | 1.62 |
| 11 | 0.80 | 0.35 | 1.97 |
| 12 | 0.51 | 0.05 | 2.02 |
| 13 | 0.41 | -0.05 | 1.97 |
| 14 | 0.35 | -0.11 | 1.87 |
| 15 | 0.30 | -0.15 | 1.72 |
| 16 | 0.41 | -0.05 | 1.67 |
| 17 | 0.64 | 0.19 | 1.86 |
| 18 | 0.42 | -0.04 | 1.82 |
| 19 | 0.27 | -0.19 | 1.63 |
| 20 | 0.18 | -0.27 | 1.36 |
| 21 | 0.53 | 0.08 | 1.44 |
| 22 | 0.48 | 0.03 | 1.47 |
| 23 | 0.26 | -0.19 | 1.28 |
| 24 | 0.45 | 0.00 | 1.27 |
| 25 | 0.57 | 0.12 | 1.39 |
| 26 | 0.20 | -0.25 | 1.14 |
| 27 | 0.61 | 0.16 | 1.30 |
| 28 | 0.11 | -0.35 | 0.95 |
| 29 | 0.43 | -0.03 | 0.93 |
| 30 | 0.21 | -0.25 | 0.68 |
| 31 | 0.31 | -0.15 | 0.53 |
| 32 | 0.11 | -0.35 | 0.19 |
| 33 | 0.59 | 0.13 | 0.32 |
| 34 | 0.46 | 0.00 | 0.32 |
| 35 | 0.60 | 0.15 | 0.47 |
| 36 | 0.31 | -0.14 | 0.32 |
| 37 | 0.22 | -0.24 | 0.09 |
| 38 | 0.41 | -0.05 | 0.04 |
| 39 | 0.51 | 0.06 | 0.10 |
| 40 | 0.50 | 0.05 | 0.14 |
| 41 | 0.20 | -0.25 | -0.11 |
| 42 | 0.51 | 0.05 | -0.05 |
| 43 | 0.75 | 0.30 | 0.24 |
| 44 | 0.61 | 0.15 | 0.39 |
| 45 | 0.35 | -0.10 | 0.29 |
| 46 | 0.51 | 0.05 | 0.34 |
| 47 | 0.41 | -0.05 | 0.30 |
| 48 | 0.20 | -0.25 | 0.05 |
| 49 | 0.30 | -0.15 | -0.10 |
| 50 | 0.56 | 0.10 | 0.00 |

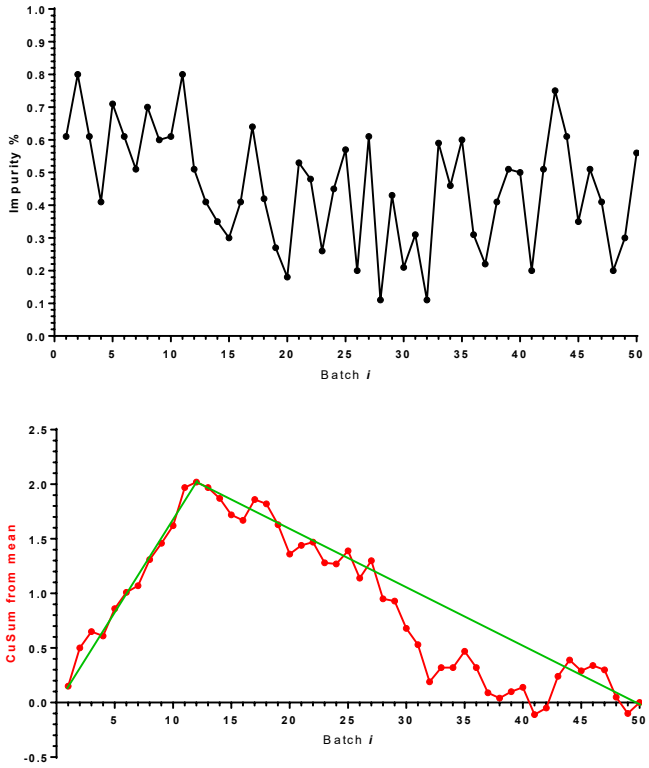



**Figure 13 Assay data scatter plot & CuSum from the mean**

The scatter plot of the assay data shows broad scatter but no immediately obvious trend. On the other hand, the CuSum plot shows a maximum at batch 12. The question to be asked is 'Is this maximum statistically significantly different from the noise?'

Firstly we need to calculate the localised standard deviation from step 1 on page 46. This is illustrated on the next page.

| Batch *i* | % impurity | Difference from previous batch $\Delta_i = X_i - X_{i+1}$ | $\Delta_i^2$ |
|---|---|---|---|
| 1 | 0.61 | -0.20 | 0.039 |
| 2 | 0.80 | 0.20 | 0.039 |
| 3 | 0.61 | 0.20 | 0.041 |
| 4 | 0.41 | -0.30 | 0.092 |
| 5 | 0.71 | 0.10 | 0.011 |
| 6 | 0.61 | 0.10 | 0.010 |
| 7 | 0.51 | -0.19 | 0.037 |
| 8 | 0.70 | 0.10 | 0.010 |
| 9 | 0.60 | -0.01 | 0.000 |
| 10 | 0.61 | -0.19 | 0.038 |
| 11 | 0.80 | 0.30 | 0.087 |
| 12 | 0.51 | 0.10 | 0.010 |
| 13 | 0.41 | 0.06 | 0.004 |
| 14 | 0.35 | 0.04 | 0.002 |
| 15 | 0.30 | -0.10 | 0.010 |
| 16 | 0.41 | -0.24 | 0.057 |
| 17 | 0.64 | 0.23 | 0.051 |
| 18 | 0.42 | 0.15 | 0.023 |
| 19 | 0.27 | 0.09 | 0.007 |
| 20 | 0.18 | -0.35 | 0.122 |
| 21 | 0.53 | 0.05 | 0.003 |
| 22 | 0.48 | 0.22 | 0.048 |
| 23 | 0.26 | -0.19 | 0.036 |
| 24 | 0.45 | -0.12 | 0.015 |
| 25 | 0.57 | 0.37 | 0.139 |
| 26 | 0.20 | -0.41 | 0.171 |
| 27 | 0.61 | 0.51 | 0.256 |
| 28 | 0.11 | -0.32 | 0.102 |
| 29 | 0.43 | 0.22 | 0.049 |
| 30 | 0.21 | -0.10 | 0.010 |
| 31 | 0.31 | 0.20 | 0.040 |
| 32 | 0.11 | -0.48 | 0.232 |
| 33 | 0.59 | 0.13 | 0.018 |
| 34 | 0.46 | -0.15 | 0.021 |
| 35 | 0.60 | 0.29 | 0.085 |
| 36 | 0.31 | 0.09 | 0.008 |
| 37 | 0.22 | -0.19 | 0.035 |
| 38 | 0.41 | -0.11 | 0.011 |
| 39 | 0.51 | 0.01 | 0.000 |
| 40 | 0.50 | 0.30 | 0.090 |
| 41 | 0.20 | -0.30 | 0.092 |
| 42 | 0.51 | -0.24 | 0.060 |
| 43 | 0.75 | 0.14 | 0.021 |
| 44 | 0.61 | 0.25 | 0.064 |
| 45 | 0.35 | -0.15 | 0.023 |
| 46 | 0.51 | 0.10 | 0.010 |
| 47 | 0.41 | 0.20 | 0.041 |
| 48 | 0.20 | -0.10 | 0.010 |
| 49 | 0.30 | -0.25 | 0.065 |
| 50 | 0.56 | | |

Thus $s_L$ is calculated from;

$$s_L = \sqrt{\frac{\sum_{i=1}^{i=n} \Delta_i^2}{2(n-1)}} = \sqrt{\frac{2.404}{2(50-1)}} = 0.157$$

and the maximum CuSum occurs at Batch 12 with a value of 2.02.
The test statistic is hence

$$\frac{|CuSum_{\max}|}{s_L} = \frac{10.44}{0.823} = 12.7$$

From Table 3, and a span of 50, the critical values for 95% and 99% confidence are 8.6 and 10.4 respectively. Therefore this turning point is highly statistically significant and partitions the data set into two parts. The mean for the first part (Batches 1 to 12) is 0.62% with a standard deviation of 0.12 and the mean for the second part (Batches 13 to 50) is 0.40% with a standard deviation of 0.16. This result is shown in Figure 14.

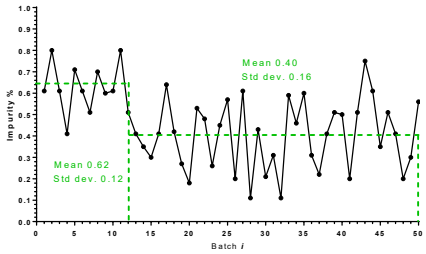



**Figure 14 Assay data scatter plot & CuSum from the mean with Manhattan plot.**

Inspection of Figure 13 shows that if there is another statistically significant turning point it is likely to lie in the second part.

Therefore a second CuSum analysis is performed on the second 38 batches only 13 to 50). The results are shown on the next page.

| Batch $i$ | % impurity | Difference from mean $\delta_i = X_i - \bar{X}$ | CuSum $S_i$ $S_i = S_{i-1} + \delta_i$ | Difference from previous batch $\Delta_i = X_i - X_{i+1}$ | $\Delta_i^2$ |
|---|---|---|---|---|---|
| 13 | 0.41 | 0.01 | 0.01 | 0.06 | 0.004 |
| 14 | 0.35 | -0.06 | -0.05 | 0.04 | 0.002 |
| 15 | 0.30 | -0.10 | -0.15 | -0.10 | 0.010 |
| 16 | 0.41 | 0.00 | -0.14 | -0.24 | 0.057 |
| 17 | 0.64 | 0.24 | 0.10 | 0.23 | 0.051 |
| 18 | 0.42 | 0.02 | 0.12 | 0.15 | 0.023 |
| 19 | 0.27 | -0.13 | -0.02 | 0.09 | 0.007 |
| 20 | 0.18 | -0.22 | -0.23 | -0.35 | 0.122 |
| 21 | 0.53 | 0.13 | -0.10 | 0.05 | 0.003 |
| 22 | 0.48 | 0.08 | -0.02 | 0.22 | 0.048 |
| 23 | 0.26 | -0.14 | -0.16 | -0.19 | 0.036 |
| 24 | 0.45 | 0.05 | -0.11 | -0.12 | 0.015 |
| 25 | 0.57 | 0.17 | 0.06 | 0.37 | 0.139 |
| 26 | 0.20 | -0.20 | -0.14 | -0.41 | 0.171 |
| 27 | 0.61 | 0.21 | 0.08 | 0.51 | 0.256 |
| 28 | 0.11 | -0.29 | -0.22 | -0.32 | 0.102 |
| 29 | 0.43 | 0.03 | -0.19 | 0.22 | 0.049 |
| 30 | 0.21 | -0.19 | -0.39 | -0.10 | 0.010 |
| 31 | 0.31 | -0.09 | -0.48 | 0.20 | 0.040 |
| 32 | 0.11 | -0.29 | -0.77 | -0.48 | 0.232 |
| 33 | 0.59 | 0.19 | -0.58 | 0.13 | 0.018 |
| 34 | 0.46 | 0.05 | -0.53 | -0.15 | 0.021 |
| 35 | 0.60 | 0.20 | -0.33 | 0.29 | 0.085 |
| 36 | 0.31 | -0.09 | -0.42 | 0.09 | 0.008 |
| 37 | 0.22 | -0.18 | -0.60 | -0.19 | 0.035 |
| 38 | 0.41 | 0.00 | -0.60 | -0.11 | 0.011 |
| 39 | 0.51 | 0.11 | -0.49 | 0.01 | 0.000 |
| 40 | 0.50 | 0.10 | -0.39 | 0.30 | 0.090 |
| 41 | 0.20 | -0.20 | -0.58 | -0.30 | 0.092 |
| 42 | 0.51 | 0.11 | -0.48 | -0.24 | 0.060 |
| 43 | 0.75 | 0.35 | -0.13 | 0.14 | 0.021 |
| 44 | 0.61 | 0.21 | 0.08 | 0.25 | 0.064 |
| 45 | 0.35 | -0.05 | 0.03 | -0.15 | 0.023 |
| 46 | 0.51 | 0.10 | 0.13 | 0.10 | 0.010 |
| 47 | 0.41 | 0.01 | 0.14 | 0.20 | 0.041 |
| 48 | 0.20 | -0.20 | -0.06 | -0.10 | 0.010 |
| 49 | 0.30 | -0.10 | -0.16 | -0.25 | 0.065 |
| 50 | 0.56 | 0.16 | 0.00 | | |

As before $s_L$ is calculated from;

$$s_L = \sqrt{\frac{\sum_{i=1}^{i=n} \Delta_i^2}{2(n-1)}} = \sqrt{\frac{2.030}{2(38-1)}} = 0.166$$

and the maximum CuSum occurs at Batch 32 with a value of -0.77 (Figure 15).

The test statistic is hence

$$\frac{|CuSum_{\max}|}{s_L} = \frac{0.77}{0.166} = 4.7$$

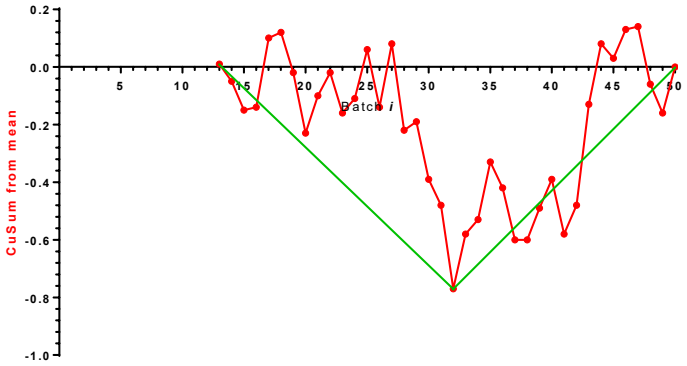



**Figure 15 CuSum from the mean for batches 13 to 50**

From Table 3, and a span of 38, the interpolated critical values for 95% and 99% confidence are 7.8 and 9.1 respectively.

Therefore this turning point is not statistically significant at 95% confidence.

It may be concluded that there is a statistically significant shift in the process mean at around batch 12 which results in a lower impurity content.

## Appendix 5: Example of SPC for discrete data; p and np charts

Similar principles are followed for discrete data trending as for continuous variables but the control limits are calculated from different equations. Let us assume that we are carrying out an inspection of incoming materials for defects and we take a sample size of n=50 units. For each subgroup of 50 items we list the number of defects found and calculate the fraction defective. There is no advantage in calculating the fraction defective p chart over the np chart as will be seen.

| Subgroup | # of defectives | Fraction defective |
|---|---|---|
| 1 | 3 | 0.06 |
| 2 | 8 | 0.16 |
| 3 | 3 | 0.06 |
| 4 | 5 | 0.10 |
| 5 | 4 | 0.08 |
| 6 | 10 | 0.20 |
| 7 | 10 | 0.20 |
| 8 | 9 | 0.18 |
| 9 | 4 | 0.08 |
| 10 | 6 | 0.12 |
| 11 | 9 | 0.18 |
| 12 | 8 | 0.16 |
| 13 | 12 | 0.24 |
| 14 | 6 | 0.12 |
| 15 | 8 | 0.16 |
| 16 | 8 | 0.16 |
| 17 | 10 | 0.20 |
| 18 | 13 | 0.26 |
| 19 | 9 | 0.18 |
| 20 | 5 | 0.10 |
| 21 | 7 | 0.14 |
| 22 | 9 | 0.18 |
| 23 | 5 | 0.10 |
| 24 | 3 | 0.06 |
| 25 | 13 | 0.26 |

**Table 6 Discrete data for p & np control charts**

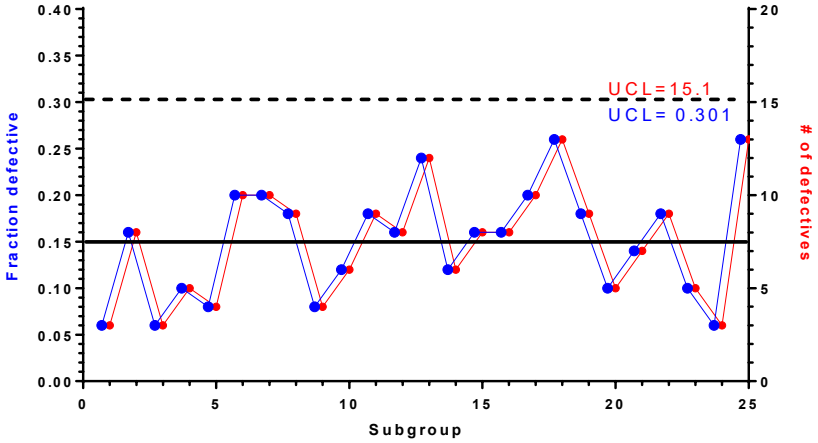



**Figure 16 p and np control charts for data from Table 6 (Note the fraction defective data points have been offset as they are overlap the # of defectives data points)**

The p charts limits for fraction defective are calculated from equation (1.12)

$$UCL = \overline{p} + 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.150 + 3\sqrt{\frac{0.150(1-0.150)}{50}} = 0.150 + 0.151 = 0.301 \qquad (1.49)$$

$$\text{Centre Line} = \overline{p} = 0.150$$

and for the number defective from equation (1.13);

$$UCL = n\overline{p} + 3\sqrt{n\overline{p}(1-\overline{p})} = 7.48 + 3\sqrt{7.48(1-0.150)} = 7.48 + 7.57 = 15.05$$

$$\text{Centre Line} = \text{n}\overline{p} = 7.48 \qquad (1.50)$$

In this example there are no obvious trends detected.

# Appendix 6: Example of setting Stability Trend Limits using a simple linear regression approach

The calculations may be performed manually in Excel using the equations detailed below or by Statistical Software Packages (for example, Minitab, JMP, SAS, GraphPad Prism amongst others). Whatever method is used it should be validated or verified under actual conditions of use..

The example given is illustrated using Excel and JMP. Not all the parameters described in the theory section on page 35 are illustrated but may be calculated from the tables.

The data consist of 8 time points (T*j*) at 0, 3, 6, 9, 12, 18, 24 and 36 months. At each time point the response, (R*j*) is determined in triplicate ie a total (*N*) of 24 data points. The model we are fitting is which was define in (1.16)

$$\hat{R}_j = b + mT_j \tag{1.51}$$

The data and the preliminary calculations are shown in Table 7. The relevant equations are (1.17) to (1.23).

| | Response | Time | | | |
|---|---|---|---|---|---|
| $j$ | $R_j$ | $T_j$ | $(R_j-\overline{R})^2$ | $(T_j-\overline{T})^2$ | $(R_j-\overline{R})(T_j-\overline{T})$ |
| 1 | 99.5 | 0 | 2.2375 | 182.25 | -20.1938 |
| 2 | 99.2 | 0 | 1.4300 | 182.25 | -16.1438 |
| 3 | 99.7 | 0 | 2.8759 | 182.25 | -22.8938 |
| 4 | 99.0 | 3 | 0.9917 | 110.25 | -10.4563 |
| 5 | 99.2 | 3 | 1.4300 | 110.25 | -12.5563 |
| 6 | 99.5 | 3 | 2.2375 | 110.25 | -15.7063 |
| 7 | 98.5 | 6 | 0.2459 | 56.25 | -3.7188 |
| 8 | 97.8 | 6 | 0.0417 | 56.25 | 1.5312 |
| 9 | 99.2 | 6 | 1.4300 | 56.25 | -8.9688 |
| 10 | 98.0 | 9 | 0.0000 | 20.25 | 0.0187 |
| 11 | 98.5 | 9 | 0.2459 | 20.25 | -2.2313 |
| 12 | 99.0 | 9 | 0.9917 | 20.25 | -4.4813 |
| 13 | 98.3 | 12 | 0.0875 | 2.25 | -0.4438 |
| 14 | 98.5 | 12 | 0.2459 | 2.25 | -0.7438 |
| 15 | 98.7 | 12 | 0.4842 | 2.25 | -1.0438 |
| 16 | 97.2 | 18 | 0.6467 | 20.25 | -3.6187 |
| 17 | 97.0 | 18 | 1.0084 | 20.25 | -4.5187 |
| 18 | 97.5 | 18 | 0.2542 | 20.25 | -2.2687 |
| 19 | 96.5 | 24 | 2.2625 | 110.25 | -15.7938 |
| 20 | 95.9 | 24 | 4.4275 | 110.25 | -22.0937 |
| 21 | 97.5 | 24 | 0.2542 | 110.25 | -5.2937 |
| 22 | 95.4 | 36 | 6.7817 | 506.25 | -58.5937 |
| 23 | 96.0 | 36 | 4.0167 | 506.25 | -45.0937 |
| 24 | 96.5 | 36 | 2.2625 | 506.25 | -33.8437 |
| **SUMs** | **2352.1** | **324** | **36.8896** | **3024.00** | **-309.1500** |
| **N** | $\overline{R}$ | $\overline{T}$ | | | |
| **24** | **98.00** | **13.5** | | | |

**Table 7 Stability data and initial summations**

Hence;

$$S_R = \sum_j^N \left(R_j - \overline{R}\right)^2 = 36.8896 \tag{1.52}$$

$$S_T = \sum_j^N \left(T_j - \overline{T}\right)^2 = 3024.00 \tag{1.53}$$

and

$$S_{RT} = \sum_j^N \left(R_j - \overline{R}\right)\left(T_j - \overline{T}\right) = -309.1500 \tag{1.54}$$

The slope of the regression line, *m*, and the intercept, *b*, are obtained by substitution in equations (1.24) and (1.23)

$$m = \frac{S_{RT}}{S_T} = \frac{-309.1500}{3024.00} = -0.1022 \tag{1.55}$$

and

$$b = \overline{R} - m\overline{T} = 98.00 - (-0.1022 * 13.5) = 99.3843 \tag{1.56}$$

Hence (1.51) becomes the fitted regression model

$$\hat{R}_j = 99.3843 - 0.1022T_j \tag{1.57}$$

We can now calculate the errors associated with this regression model from the mean square error, *MSE*, equation (1.25);

$$MSE = \frac{S_R - mS_{RT}}{N-2} = \frac{36.8896 - (-0.1022(-309.1500)}{(24-2)} = 0.2402 \tag{1.58}$$

hence the root mean square error, *RMSE*, is found from

$$RMSE = \sqrt{MSE} = \sqrt{0.2402} = 0.4901 \tag{1.59}$$

The sums of squares required for the ANOVA table is now calculated from the fitted values, $\hat{R}_j$, and shown in Table 8

| | Response | Time | Best Fit Line | Residuals | Squares for ANOVA | |
|---|---|---|---|---|---|---|
| $j$ | $R_j$ | $T_j$ | $\hat{R}_j = b + mT_j$ | $R_j - \hat{R}_j$ | $\left(R_j - \hat{R}_j\right)^2$ | $\left(\hat{R}_j - \overline{R}\right)^2$ |
| 1 | 99.50 | 0 | 99.3843 | 0.1157 | 0.0134 | 1.9048 |
| 2 | 99.20 | 0 | 99.3843 | -0.1843 | 0.0340 | 1.9048 |
| 3 | 99.70 | 0 | 99.3843 | 0.3157 | 0.0997 | 1.9048 |
| 4 | 99.00 | 3 | 99.0776 | -0.0776 | 0.0060 | 1.1523 |
| 5 | 99.20 | 3 | 99.0776 | 0.1224 | 0.0150 | 1.1523 |
| 6 | 99.50 | 3 | 99.0776 | 0.4224 | 0.1784 | 1.1523 |
| 7 | 98.50 | 6 | 98.7709 | -0.2709 | 0.0734 | 0.5879 |
| 8 | 97.80 | 6 | 98.7709 | -0.9709 | 0.9427 | 0.5879 |
| 9 | 99.20 | 6 | 98.7709 | 0.4291 | 0.1841 | 0.5879 |
| 10 | 98.00 | 9 | 98.4642 | -0.4642 | 0.2155 | 0.2116 |
| 11 | 98.50 | 9 | 98.4642 | 0.0358 | 0.0013 | 0.2116 |
| 12 | 99.00 | 9 | 98.4642 | 0.5358 | 0.2871 | 0.2116 |
| 13 | 98.30 | 12 | 98.1575 | 0.1425 | 0.0203 | 0.0235 |
| 14 | 98.50 | 12 | 98.1575 | 0.3425 | 0.1173 | 0.0235 |
| 15 | 98.70 | 12 | 98.1575 | 0.5425 | 0.2943 | 0.0235 |
| 16 | 97.20 | 18 | 97.5441 | -0.3441 | 0.1184 | 0.2116 |
| 17 | 97.00 | 18 | 97.5441 | -0.5441 | 0.2961 | 0.2116 |
| 18 | 97.50 | 18 | 97.5441 | -0.0441 | 0.0019 | 0.2116 |
| 19 | 96.50 | 24 | 96.9307 | -0.4307 | 0.1855 | 1.1523 |
| 20 | 95.90 | 24 | 96.9307 | -1.0307 | 1.0624 | 1.1523 |
| 21 | 97.50 | 24 | 96.9307 | 0.5693 | 0.3241 | 1.1523 |
| 22 | 95.40 | 36 | 95.7039 | -0.3039 | 0.0924 | 5.2910 |
| 23 | 96.00 | 36 | 95.7039 | 0.2961 | 0.0876 | 5.2910 |
| 24 | 96.50 | 36 | 95.7039 | 0.7961 | 0.6337 | 5.2910 |
| **SUMs** | **2352.1** | **324** | | **0.0000** | **5.2845** | **31.6051** |
| | | | | | **Sums of Squares for ANOVA** | |

**Table 8 Sums of squares required for the ANOVA table**

**The sums of squares required are those due to the regression model, $SS_{Reg}$, and the sum of squares attributable to the residual error, $SS_{Error}$. Taking the values from Table 8 we can construct the ANOVA table shown in Table 9. The large *F* ratio indicates a highly significant regression model.**

| SOURCE OF VARIATION | DF | | SUM OF SQUARES (SS) | | MEAN SQUARE (MS) | | F ratio | | Probability |
|---|---|---|---|---|---|---|---|---|---|
| Due to regression | 1 | 1 | $SS_{\text{Reg}} = \sum_{j=1}^{N}\left(\hat{R}_j - \overline{R}\right)^2$ | 31.6051 | $MS_{\text{Reg}} = \frac{SS_{\text{Reg}}}{dof}$ | 31.6051 | $F = \frac{MS_{\text{Reg}}}{MSE}$ | 131.5752 | 0.0000 |
| Residual Error | N-2 | 22 | $SS_{\text{Error}} = \sum_{j=1}^{N}\left(R_j - \hat{R}_j\right)^2$ | 5.2845 | $MSE = \frac{SS_{\text{Res}}}{dof}$ | 0.2402 | | | |
| TOTAL CORRECTED for $\overline{R}$ | N-1 | 23 | | 36.8896 | | | | | |

**Table 9 ANOVA table for the simple linear model**

The 99% confidence contours for regression and prediction are calculated for each of the $T_j$ time points from equations (1.30) and (1.31). The 99.5% confidence acceptance trend limits are calculated from equation (1.32). The results of these calculations are shown in Table 10 and plotted in Figure 17.

| | | | Calculated on actual values | | | | | | Calculated on mean values for plotting limits | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Response | Time | REGRESSION | | | PREDICTION | | | *REGRESSION* | | | *PREDICTION* | | | *TREND LIMITS* | |
| *j* | $R_j$ | $T_j$ | 99% CI of mean Regression | 99% LCL | 99% UCL | 99% CI of individual Prediction | 99% LCL | 99% UCL | *99% CI of mean Regression* | *99% LCL* | *99% UCL* | *99% CI of individual Prediction* | *99% LCL* | *99% UCL* | *99% LTL Individual* | *99% UTL Individual* |
| 1 | 99.50 | 0 | 0.4411 | 99.0589 | 99.9411 | 1.4321 | 98.0679 | 100.9321 | 0.4411 | 98.9432 | 99.8254 | 1.4321 | 97.9522 | 100.8164 | 97.8242 | 100.9444 |
| 2 | 99.20 | 0 | 0.4411 | 98.7589 | 99.6411 | 1.4320 | 97.7680 | 100.6320 | 0.4411 | 98.9432 | 99.8254 | 1.4320 | 97.9523 | 100.8163 | 97.8242 | 100.9444 |
| 3 | 99.70 | 0 | 0.4411 | 99.2589 | 100.1411 | 1.4321 | 98.2679 | 101.1321 | 0.4411 | 98.9432 | 99.8254 | 1.4321 | 97.9522 | 100.8164 | 97.8242 | 100.9444 |
| 4 | 99.00 | 3 | 0.3861 | 98.6139 | 99.3861 | 1.4320 | 97.5680 | 100.4320 | 0.3861 | 98.6915 | 99.4637 | 1.4320 | 97.6456 | 100.5096 | 97.5175 | 100.6377 |
| 5 | 99.20 | 3 | 0.3861 | 98.8139 | 99.5861 | 1.4320 | 97.7680 | 100.6320 | 0.3861 | 98.6915 | 99.4637 | 1.4320 | 97.6456 | 100.5096 | 97.5175 | 100.6377 |
| 6 | 99.50 | 3 | 0.3861 | 99.1139 | 99.8861 | 1.4321 | 98.0679 | 100.9321 | 0.3861 | 98.6915 | 99.4637 | 1.4321 | 97.6455 | 100.5097 | 97.5175 | 100.6377 |
| 7 | 98.50 | 6 | 0.3391 | 98.1609 | 98.8391 | 1.4319 | 97.0681 | 99.9319 | 0.3391 | 98.4318 | 99.1101 | 1.4319 | 97.3391 | 100.2028 | 97.2108 | 100.3310 |
| 8 | 97.80 | 6 | 0.3391 | 97.4609 | 98.1391 | 1.4317 | 96.3683 | 99.2317 | 0.3391 | 98.4318 | 99.1101 | 1.4317 | 97.3392 | 100.2026 | 97.2108 | 100.3310 |
| 9 | 99.20 | 6 | 0.3391 | 98.8609 | 99.5391 | 1.4320 | 97.7680 | 100.6320 | 0.3391 | 98.4318 | 99.1101 | 1.4320 | 97.3389 | 100.2029 | 97.2108 | 100.3310 |
| 10 | 98.00 | 9 | 0.3038 | 97.6962 | 98.3038 | 1.4317 | 96.5683 | 99.4317 | 0.3038 | 98.1604 | 98.7680 | 1.4317 | 97.0325 | 99.8960 | 96.9041 | 100.0243 |
| 11 | 98.50 | 9 | 0.3038 | 98.1962 | 98.8038 | 1.4319 | 97.0681 | 99.9319 | 0.3038 | 98.1604 | 98.7680 | 1.4319 | 97.0324 | 99.8961 | 96.9041 | 100.0243 |
| 12 | 99.00 | 9 | 0.3038 | 98.6962 | 99.3038 | 1.4320 | 97.5680 | 100.4320 | 0.3038 | 98.1604 | 98.7680 | 1.4320 | 97.0322 | 99.8962 | 96.9041 | 100.0243 |
| 13 | 98.30 | 12 | 0.2845 | 98.0155 | 98.5845 | 1.4318 | 96.8682 | 99.7318 | 0.2845 | 97.8730 | 98.4420 | 1.4318 | 96.7257 | 99.5893 | 96.5974 | 99.7176 |
| 14 | 98.50 | 12 | 0.2845 | 98.2155 | 98.7845 | 1.4319 | 97.0681 | 99.9319 | 0.2845 | 97.8730 | 98.4420 | 1.4319 | 96.7257 | 99.5894 | 96.5974 | 99.7176 |
| 15 | 98.70 | 12 | 0.2845 | 98.4155 | 98.9845 | 1.4319 | 97.2681 | 100.1319 | 0.2845 | 97.8730 | 98.4420 | 1.4319 | 96.7256 | 99.5894 | 96.5974 | 99.7176 |
| 16 | 97.20 | 18 | 0.3038 | 96.8962 | 97.5038 | 1.4316 | 95.7684 | 98.6316 | 0.3038 | 97.2403 | 97.8479 | 1.4316 | 96.1126 | 98.9757 | 95.9840 | 99.1042 |
| 17 | 97.00 | 18 | 0.3038 | 96.6962 | 97.3038 | 1.4315 | 95.5685 | 98.4315 | 0.3038 | 97.2403 | 97.8479 | 1.4315 | 96.1126 | 98.9756 | 95.9840 | 99.1042 |
| 18 | 97.50 | 18 | 0.3038 | 97.1962 | 97.8038 | 1.4316 | 96.0684 | 98.9316 | 0.3038 | 97.2403 | 97.8479 | 1.4316 | 96.1125 | 98.9758 | 95.9840 | 99.1042 |
| 19 | 96.50 | 24 | 0.3861 | 96.1139 | 96.8861 | 1.4314 | 95.0686 | 97.9314 | 0.3861 | 96.5446 | 97.3169 | 1.4314 | 95.4993 | 98.3621 | 95.3707 | 98.4908 |
| 20 | 95.90 | 24 | 0.3861 | 95.5139 | 96.2861 | 1.4313 | 94.4687 | 97.3313 | 0.3861 | 96.5446 | 97.3169 | 1.4313 | 95.4994 | 98.3620 | 95.3707 | 98.4908 |
| 21 | 97.50 | 24 | 0.3861 | 97.1139 | 97.8861 | 1.4316 | 96.0684 | 98.9316 | 0.3861 | 96.5446 | 97.3169 | 1.4316 | 95.4991 | 98.3624 | 95.3707 | 98.4908 |
| 22 | 95.40 | 36 | 0.6317 | 94.7683 | 96.0317 | 1.4312 | 93.9688 | 96.8312 | 0.6317 | 95.0723 | 96.3356 | 1.4312 | 94.2728 | 97.1351 | 94.1439 | 97.2640 |
| 23 | 96.00 | 36 | 0.6317 | 95.3683 | 96.6317 | 1.4313 | 94.5687 | 97.4313 | 0.6317 | 95.0723 | 96.3356 | 1.4313 | 94.2726 | 97.1352 | 94.1439 | 97.2640 |
| 24 | 96.50 | 36 | 0.6317 | 95.8683 | 97.1317 | 1.4314 | 95.0686 | 97.9314 | 0.6317 | 95.0723 | 96.3356 | 1.4314 | 94.2725 | 97.1354 | 94.1439 | 97.2640 |

**Table 10 Confidence contour and acceptance trend limit calculations**

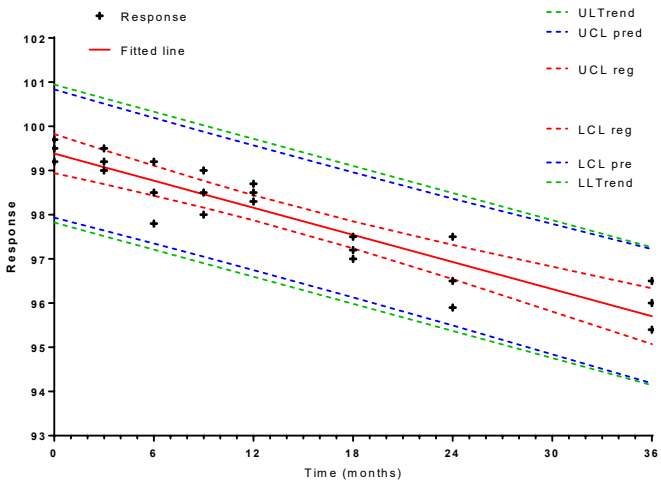



**Figure 17 Regression line with confidence contours of regression, prediction and trend limits**

The output from JMP shows very similar results. They are not identical because of differences in rounding and algorithms

## Bivariate Fit of Result By Time

### Linear Fit

Result = 99.384301 - 0.1022321*Time

### Summary of Fit

| | |
|---|---|
| RSquare | 0.856748 |
| RSquare Adj | 0.850236 |
| Root Mean Square Error | 0.490107 |
| Mean of Response | 98.00417 |
| Observations (or Sum Wgts) | 24 |

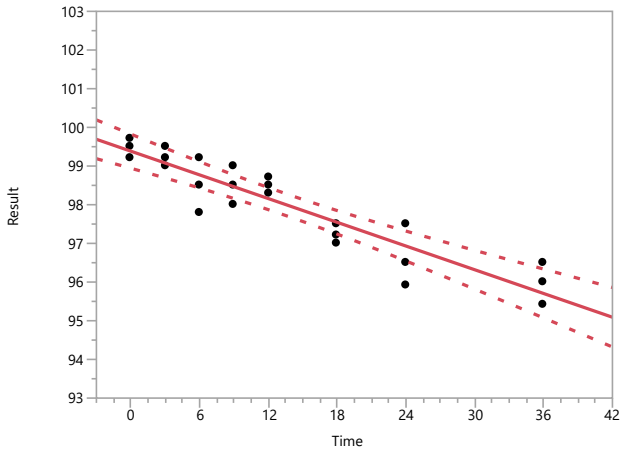
### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 1 | 31.605067 | 31.6051 | 131.5752 |
| Error | 22 | 5.284516 | 0.2402 | Prob > F |
| C. Total | 23 | 36.889583 | | <.0001* |

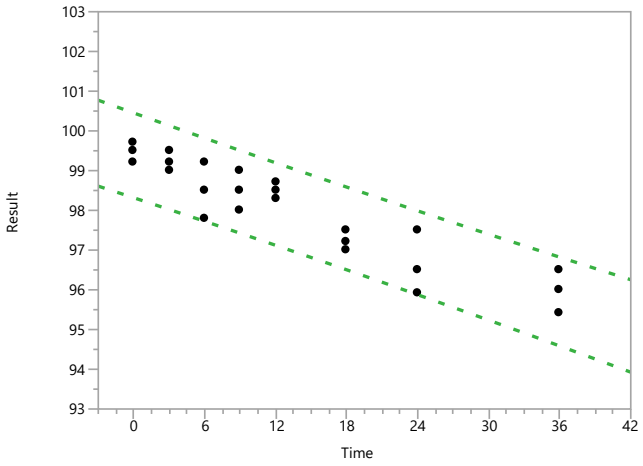
### Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 99.384301 | 0.156478 | 635.13 | <.0001* |
| Time | -0.102232 | 0.008913 | -11.47 | <.0001* |

**Linear Regression with Confidence Intervals**




**Linear Regression with Prediction Intervals**




| DATA | | Linear Regression | Confidence Intervals | | Prediction Intervals | |
|---|---|---|---|---|---|---|
| Time | Result | Predicted Result | Lower 99% Mean Result | Upper 99% Mean Result | Lower 99% Individual Result | Upper 99% Individual Result |
| 0 | 99.5 | 99.3843 | 98.9432 | 99.8254 | 97.9341 | 100.8345 |
| 0 | 99.2 | 99.3843 | 98.9432 | 99.8254 | 97.9341 | 100.8345 |
| 0 | 99.7 | 99.3843 | 98.9432 | 99.8254 | 97.9341 | 100.8345 |
| 3 | 99.0 | 99.0776 | 98.6915 | 99.4637 | 97.6432 | 100.5120 |
| 3 | 99.2 | 99.0776 | 98.6915 | 99.4637 | 97.6432 | 100.5120 |
| 3 | 99.5 | 99.0776 | 98.6915 | 99.4637 | 97.6432 | 100.5120 |
| 6 | 98.5 | 98.7709 | 98.4318 | 99.1101 | 97.3484 | 100.1934 |
| 6 | 97.8 | 98.7709 | 98.4318 | 99.1101 | 97.3484 | 100.1934 |
| 6 | 99.2 | 98.7709 | 98.4318 | 99.1101 | 97.3484 | 100.1934 |
| 9 | 98.0 | 98.4642 | 98.1604 | 98.7680 | 97.0497 | 99.8787 |
| 9 | 98.5 | 98.4642 | 98.1604 | 98.7680 | 97.0497 | 99.8787 |
| 9 | 99.0 | 98.4642 | 98.1604 | 98.7680 | 97.0497 | 99.8787 |
| 12 | 98.3 | 98.1575 | 97.8730 | 98.4420 | 96.7470 | 99.5680 |
| 12 | 98.5 | 98.1575 | 97.8730 | 98.4420 | 96.7470 | 99.5680 |
| 12 | 98.7 | 98.1575 | 97.8730 | 98.4420 | 96.7470 | 99.5680 |
| 18 | 97.2 | 97.5441 | 97.2403 | 97.8479 | 96.1296 | 98.9586 |
| 18 | 97.0 | 97.5441 | 97.2403 | 97.8479 | 96.1296 | 98.9586 |
| 18 | 97.5 | 97.5441 | 97.2403 | 97.8479 | 96.1296 | 98.9586 |
| 24 | 96.5 | 96.9307 | 96.5446 | 97.3169 | 95.4963 | 98.3652 |
| 24 | 95.9 | 96.9307 | 96.5446 | 97.3169 | 95.4963 | 98.3652 |
| 24 | 97.5 | 96.9307 | 96.5446 | 97.3169 | 95.4963 | 98.3652 |
| 36 | 95.4 | 95.7039 | 95.0723 | 96.3356 | 94.1849 | 97.2230 |
| 36 | 96.0 | 95.7039 | 95.0723 | 96.3356 | 94.1849 | 97.2230 |
| 36 | 96.5 | 95.7039 | 95.0723 | 96.3356 | 94.1849 | 97.2230 |

Evaluate the trend limits:

- To review possible changes to trend limits in order to detect any drifts or other changes over time. Changes may indicate potential changes in testing or the production process
- To detect differences between production sites
- To detect product or process changes
- To determine the need for follow up or corrective actions

Note: Trend limit based on limited data may significantly change after additional data is collected.

# Appendix 7: Examples of determining parameters and Stability Trend Limits using a Random Coefficients Regression (RCR)Model

The trends generated by the RCR model and by a more simplified regression model both originate with a linear model:

$$y_t = a + b \cdot t + \varepsilon_t \quad (1.60)$$

where $y_t$ is the response at time t
$a$ is the initial response (t = 0)
$b$ is the rate of degradation and
$\varepsilon_t$ is an error term associated with $y_t$

The two models differ in the interpretation and assumptions of the parameters appearing in the equation.

The RCR model does not assume that each lot has the same intercept and slope, but rather that the intercept and slope come from a bivariate normal distribution. The trend limits are generated using the estimates determined by fitting the stability data to the bivariate normal distribution.

There is no association of a response to a specific lot in the model assumed by a simple regression. The data is combined and assumed to come from a single population. There is an implicit assumption that all lots have the same intercept and slope. A single (constant) intercept and slope are estimated from the pooled data set. The trend limits are generated using the estimates determined by fitting the stability data to this simple linear model.

To consider the difference between the two trend limit approaches it is necessary to consider if there is variability in degradation rates between lots, and if so if there is a correlation between the degradation rate and the response at product release.

Generally, the two trend limits are similar when all lots have a common degradation rate. The RCR trend limits tend to better reflect the distribution of the stability data over time when there are differences in slopes between the lots.

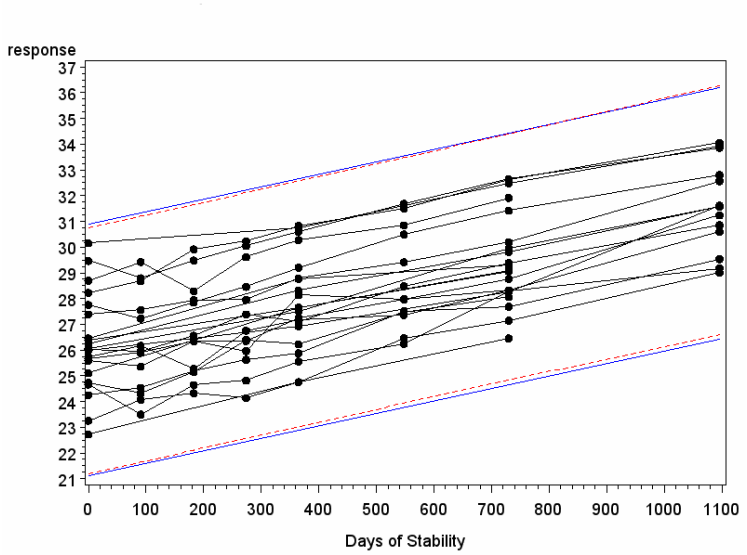
## Case 1: $\sigma^2_{\text{slope}} = 0$

The components of Ω are small and as a first order approximation can be set to zero, additionally there is no variability between the degradation rates of the lots under study.
The approximate RCR trend limits simplify to

$$a_{RCR} + b_{RCR} \cdot t \pm z_{0.995} \cdot \sqrt{\sigma^2_{\text{int}} + \sigma^2_{\text{method}}} \quad (1.61)$$

On the other hand, if there is no variability in degradation rates, the only sources of variability in the pooled data set are method variability, $\sigma^2_{\text{method}}$ and variability in the intercepts, $\sigma^2_{\text{int}}$.

Provided that the RCR model and simple regression model provide similar estimates of slope and intercept, it is expected that the two trend limits will be nearly identical as in the example below:




## Case 2: $\sigma^2_{\text{int,slope}} \geq 0$

Differences in degradation rates exist between the various stability lots under study. Consider two lots that follow a linear model, the difference in response between the two lots at time t is given by the differences in their slopes and intercepts and whether the differences in intercepts and slopes have the same or different signs.
The behaviour of the absolute difference in response for small values of t depends on whether the differences in intercepts and slopes have the same or different signs; the absolute difference in responses between the two lots may diverge for sufficiently large t and the range of the individual data points from multiple stability lots to increase with time.

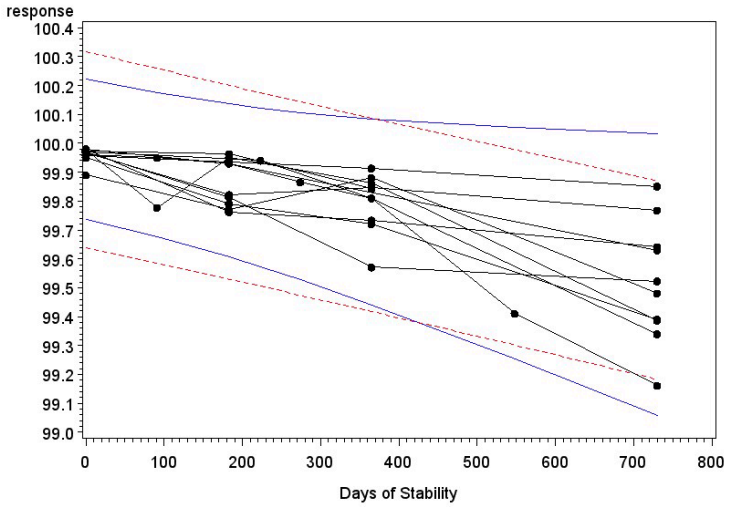
In case 2, the covariance of the intercept and slope is non-negative. There is a probability of 0.5 or greater that differences in intercepts and slopes have the same sign. Consequently, when examining many lots of data, it is expected that the range of the individual data points from multiple stability lots to increase with time.

The width of the RCR trend limits depend on the term

$$\sqrt{\sigma^2_{\text{int}} + 2 \cdot \sigma^2_{\text{int,slope}} \cdot t + \sigma^2_{\text{slope}} \cdot t^2 + \sigma^2_{\text{method}}} \qquad (1.62)$$

The coefficient of the $t^2$ term is positive and the coefficient of the $t$ term is non-negative. Hence, the width of the trend limits increases monotonically with increasing t.

A statistical model that assumes a constant variance in the data points at each time point tends to over-estimate the variability at early time points and under-estimates the variability at later time points.

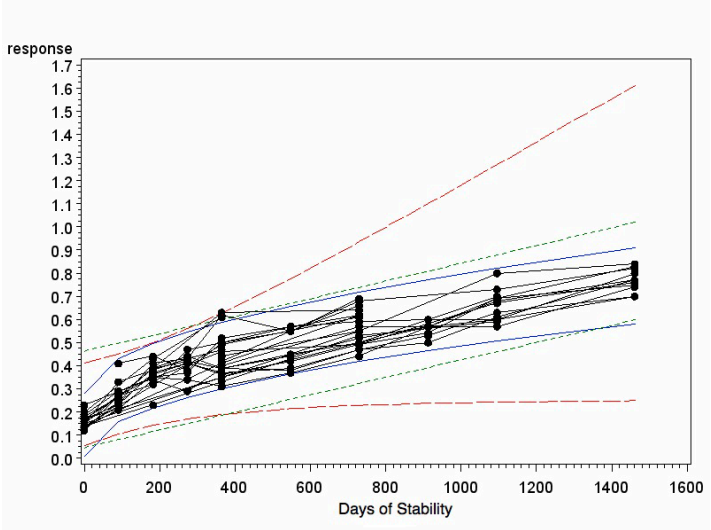



Since the trend limits in that case is a constant multiple of the standard estimate of error, the range of the trend limits is expected to be large relative to the data points for T = 0 and small relative to the range of the data points for later time points. In comparison, RCR trend limits may have a smaller width than simple regression trend limits at T=0 and to have a wider width for later time points, as in the example below:

## Case 3: Non-linearity

RCR and simple regression approaches of establishing trends assume an underlying linear degradation rate. When all data is pooled prior to analysis, small departures from non-linearity in the individual lots can be obscured by the method variability and the variability of the response at time of release due to the manufacturing process. As the RCR model fits a regression to each lot, variability of the response at time of release due to the manufacturing process does not contribute to the variability of a single lot and consequently can not obscure non-linearity in the degradation rates.

Consequently, the RCR model is more sensitive to departures from linearity. In this case, a square root transformation of the time axis was sufficient and necessary to linearize the data and provide trend limits. The graph below shows the RCR trend limits applied to the original data (red line), the trend RCR trend limits applied to the time-transformed data (blue line), and the simple regression trend limits (green line).


response
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
200
400
600
800
1000
1200
1400
1600
Days of Stability

## Data sets for RCR Examples

### Case 1

| Case 1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Lot | Days | Response | Lot | Days | Response | Lot | Days | Response |
| 1 | 0 | 25.75 | 7 | 0 | 30.16 | 14 | 0 | 22.73 |
| 1 | 91 | 26.17 | 7 | 365 | 30.74 | 14 | 365 | 24.76 |
| 1 | 183 | 25.28 | 7 | 730 | 32.49 | 14 | 730 | 26.46 |
| 1 | 274 | 26.76 | 7 | 1096 | 33.91 | 15 | 0 | 26.46 |
| 1 | 365 | 27.18 | 8 | 0 | 28.23 | 15 | 365 | 28.77 |
| 1 | 548 | 27.37 | 8 | 91 | 28.68 | 15 | 730 | 29.29 |
| 1 | 730 | 28.32 | 8 | 183 | 29.92 | 16 | 0 | 28.70 |
| 1 | 1096 | 30.60 | 8 | 274 | 30.23 | 16 | 91 | 29.43 |
| 2 | 0 | 27.76 | 8 | 365 | 30.83 | 16 | 183 | 28.29 |
| 2 | 91 | 27.23 | 8 | 548 | 31.50 | 16 | 274 | 29.62 |
| 2 | 183 | 27.85 | 8 | 730 | 32.61 | 16 | 365 | 30.28 |
| 2 | 274 | 28.46 | 8 | 1096 | 34.06 | 16 | 548 | 30.85 |
| 2 | 365 | 29.21 | 9 | 0 | 25.99 | 16 | 730 | 31.92 |
| 2 | 548 | 30.50 | 9 | 365 | 26.93 | 17 | 0 | 25.59 |
| 2 | 730 | 31.42 | 9 | 730 | 28.26 | 17 | 91 | 25.36 |
| 2 | 1096 | 32.81 | 9 | 1096 | 31.62 | 17 | 183 | 26.36 |
| 3 | 0 | 29.47 | 10 | 0 | 26.04 | 17 | 274 | 25.97 |
| 3 | 91 | 28.80 | 10 | 91 | 25.95 | 17 | 365 | 28.14 |
| 3 | 183 | 29.49 | 10 | 183 | 26.57 | 17 | 548 | 27.97 |
| 3 | 274 | 30.05 | 10 | 274 | 27.39 | 17 | 730 | 28.33 |
| 3 | 365 | 30.60 | 10 | 365 | 27.12 | 18 | 0 | 25.11 |
| 3 | 548 | 31.69 | 10 | 548 | 28.49 | 18 | 365 | 27.66 |
| 3 | 730 | 32.65 | 10 | 730 | 29.39 | 18 | 730 | 29.10 |
| 3 | 1096 | 33.87 | 10 | 1096 | 30.86 | 18 | 0 | 26.38 |
| 4 | 0 | 23.26 | 11 | 0 | 24.25 | 18 | 365 | 27.64 |
| 4 | 91 | 24.07 | 11 | 91 | 24.54 | 18 | 730 | 29.06 |
| 4 | 183 | 24.33 | 11 | 183 | 25.18 | 19 | 0 | 25.66 |
| 4 | 274 | 24.14 | 11 | 274 | 25.63 | 19 | 91 | 25.90 |
| 4 | 365 | 24.76 | 11 | 365 | 25.88 | 19 | 183 | 26.35 |
| 4 | 548 | 26.47 | 11 | 548 | 27.52 | 19 | 274 | 26.36 |
| 4 | 730 | 27.15 | 11 | 730 | 27.69 | 19 | 365 | 27.25 |
| 4 | 1096 | 29.01 | 11 | 1096 | 29.54 | 19 | 548 | 27.99 |
| 5 | 0 | 26.06 | 12 | 0 | 24.74 | 19 | 730 | 28.77 |
| 5 | 365 | 27.52 | 12 | 91 | 24.32 | 19 | 1096 | 31.24 |
| 5 | 730 | 29.95 | 12 | 183 | 25.15 | 20 | 0 | 27.39 |
| 5 | 1096 | 31.57 | 12 | 274 | 26.40 | 20 | 91 | 27.56 |
| 6 | 0 | 24.67 | 12 | 365 | 26.23 | 20 | 183 | 27.95 |
| 6 | 91 | 23.50 | 12 | 548 | 27.47 | 20 | 274 | 27.96 |
| 6 | 183 | 24.65 | 12 | 730 | 28.07 | 20 | 365 | 28.79 |
| 6 | 274 | 24.82 | 13 | 0 | 26.25 | 20 | 548 | 29.42 |
| 6 | 365 | 25.56 | 13 | 365 | 28.32 | 20 | 730 | 30.20 |
| 6 | 548 | 26.25 | 13 | 730 | 29.82 | 20 | 1096 | 32.57 |
| 6 | 730 | 28.31 | 13 | 1096 | 31.58 | | | |
| 6 | 1096 | 29.18 | | | | | | |

## Case 2

| Case 2 | | |
|---|---|---|
| Lot | Days | Response |
| 1 | 0 | 99.97 |
| 1 | 183 | 99.96 |
| 1 | 365 | 99.84 |
| 1 | 0 | 99.95 |
| 1 | 183 | 99.79 |
| 1 | 365 | 99.72 |
| 1 | 730 | 99.39 |
| 2 | 0 | 99.97 |
| 2 | 183 | 99.81 |
| 2 | 365 | 99.57 |
| 2 | 730 | 99.52 |
| 3 | 0 | 99.96 |
| 3 | 91 | 99.95 |
| 3 | 183 | 99.93 |
| 3 | 274 | 99.87 |
| 3 | 365 | 99.81 |
| 3 | 548 | 99.41 |
| 3 | 730 | 99.16 |
| 4 | 0 | 99.97 |
| 4 | 223 | 99.94 |
| 4 | 365 | 99.81 |
| 4 | 730 | 99.34 |
| 5 | 0 | 99.98 |
| 5 | 183 | 99.76 |
| 5 | 365 | 99.73 |
| 5 | 730 | 99.64 |
| 6 | 0 | 99.97 |
| 6 | 183 | 99.82 |
| 6 | 365 | 99.85 |
| 6 | 730 | 99.77 |
| 7 | 0 | 99.96 |
| 7 | 183 | 99.94 |
| 7 | 365 | 99.91 |
| 7 | 730 | 99.85 |
| 8 | 0 | 99.98 |
| 8 | 91 | 99.78 |
| 8 | 183 | 99.95 |
| 8 | 365 | 99.86 |
| 8 | 730 | 99.39 |
| 9 | 0 | 99.98 |
| 9 | 183 | 99.93 |
| 9 | 730 | 99.63 |
| 10 | 0 | 99.89 |
| 10 | 183 | 99.77 |
| 10 | 365 | 99.88 |
| 10 | 730 | 99.48 |

## Case 3

| Case 3 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Lot | Days | Response | Lot | Days | Response | Lot | Days | Response |
| 1 | 0 | 0.20 | 7 | 91 | 0.41 | 13 | 1096 | 0.60 |
| 1 | 183 | 0.43 | 7 | 183 | 0.44 | 13 | 1461 | 0.70 |
| 1 | 365 | 0.61 | 7 | 274 | 0.38 | 14 | 0 | 0.13 |
| 1 | 0 | 0.15 | 7 | 365 | 0.50 | 14 | 91 | 0.24 |
| 1 | 91 | 0.33 | 7 | 548 | 0.57 | 14 | 183 | 0.36 |
| 1 | 183 | 0.38 | 7 | 730 | 0.59 | 14 | 274 | 0.41 |
| 1 | 274 | 0.43 | 7 | 913 | 0.60 | 14 | 365 | 0.37 |
| 1 | 365 | 0.47 | 7 | 1096 | 0.60 | 14 | 548 | 0.39 |
| 2 | 0 | 0.17 | 7 | 1461 | 0.80 | 14 | 730 | 0.47 |
| 2 | 91 | 0.26 | 8 | 0 | 0.14 | 14 | 913 | 0.53 |
| 2 | 183 | 0.39 | 8 | 365 | 0.41 | 14 | 1096 | 0.63 |
| 2 | 274 | 0.44 | 8 | 730 | 0.62 | 14 | 1461 | 0.70 |
| 2 | 365 | 0.39 | 9 | 0 | 0.19 | 15 | 0 | 0.13 |
| 2 | 548 | 0.45 | 9 | 91 | 0.29 | 15 | 91 | 0.21 |
| 2 | 730 | 0.50 | 9 | 183 | 0.35 | 15 | 183 | 0.34 |
| 2 | 913 | 0.57 | 9 | 274 | 0.37 | 15 | 274 | 0.29 |
| 2 | 1096 | 0.69 | 9 | 365 | 0.52 | 15 | 365 | 0.39 |
| 2 | 1461 | 0.83 | 9 | 548 | 0.57 | 15 | 548 | 0.38 |
| 3 | 0 | 0.17 | 9 | 730 | 0.62 | 15 | 730 | 0.47 |
| 3 | 91 | 0.23 | 10 | 0 | 0.17 | 15 | 913 | 0.50 |
| 3 | 183 | 0.37 | 10 | 91 | 0.27 | 15 | 1096 | 0.61 |
| 3 | 274 | 0.41 | 10 | 183 | 0.32 | 15 | 1461 | 0.77 |
| 3 | 365 | 0.36 | 10 | 274 | 0.47 | 16 | 0 | 0.14 |
| 3 | 548 | 0.42 | 10 | 365 | 0.50 | 16 | 183 | 0.23 |
| 3 | 730 | 0.49 | 10 | 548 | 0.55 | 16 | 365 | 0.33 |
| 3 | 913 | 0.54 | 10 | 730 | 0.66 | 16 | 730 | 0.53 |
| 3 | 1096 | 0.69 | 11 | 0 | 0.16 | 16 | 1096 | 0.59 |
| 3 | 1461 | 0.75 | 11 | 91 | 0.24 | 17 | 0 | 0.15 |
| 4 | 0 | 0.18 | 11 | 183 | 0.41 | 17 | 365 | 0.35 |
| 4 | 91 | 0.26 | 11 | 274 | 0.43 | 17 | 730 | 0.55 |
| 4 | 183 | 0.39 | 11 | 365 | 0.62 | 17 | 0 | 0.15 |
| 4 | 274 | 0.42 | 11 | 548 | 0.55 | 17 | 365 | 0.63 |
| 4 | 365 | 0.39 | 11 | 730 | 0.69 | 17 | 730 | 0.64 |
| 4 | 548 | 0.44 | 12 | 0 | 0.23 | 18 | 0 | 0.14 |
| 4 | 730 | 0.52 | 12 | 365 | 0.46 | 18 | 365 | 0.49 |
| 4 | 913 | 0.57 | 12 | 730 | 0.49 | 18 | 730 | 0.61 |
| 4 | 1096 | 0.68 | 12 | 1096 | 0.67 | 18 | 365 | 0.44 |
| 4 | 1461 | 0.77 | 12 | 1461 | 0.76 | 18 | 730 | 0.55 |
| 5 | 0 | 0.17 | 13 | 0 | 0.12 | 18 | 1096 | 0.80 |
| 5 | 365 | 0.35 | 13 | 91 | 0.27 | 18 | 1461 | 0.84 |
| 5 | 730 | 0.50 | 13 | 183 | 0.35 | 19 | 365 | 0.42 |
| 5 | 1096 | 0.70 | 13 | 274 | 0.34 | 19 | 730 | 0.68 |
| 6 | 365 | 0.40 | 13 | 365 | 0.31 | 19 | 1096 | 0.73 |
| 6 | 730 | 0.57 | 13 | 548 | 0.37 | 19 | 1461 | 0.82 |
| 6 | 1096 | 0.57 | 13 | 730 | 0.44 | | | |
| 6 | 1461 | 0.74 | 13 | 913 | 0.57 | | | |