Analytical Science

A course (in 15 chapters) developed as an Open Educational Resource designed for use at 2nd year England & Wales undergraduate level and as a CPD training resource

https://edocs.hull.ac.uk/muradora/objectView.action?parentId=hull%3A2199&type=1&start=10&pid=hull%3A2351

Author	Brian W Woodget
Owner	Royal Society of Chemistry
Title	Chapter 15 – Quality Assurance, Method Development and Method Validation
Classification	F180, Analytical Chemistry
Keywords	ukoer, sfsoer, oer, open educational resources, metadata, analytical science, cpd training resource, analytical chemistry, measurement science, quality management, quality assurance, qa, vam principles, quqlity control, qc, method validation, method performance characteristics, measurement uncertainty
Description	This chapter considers the important regulatory parameters associate with 21st century analytical measurement.
Creative Commons licence	http://creativecommons.org/licenses/by-nc-nd/2.0/uk/
Language	English
File size	1.5 Mbytes
File format	Microsoft PowerPoint (1997 – 2003)

RSC | Advancing the Chemical Sciences

Chapter 14 - Quality Assurance, Method Development and Method Validation

Contents

Introduction to Quality

The value of chemical measurements is dependent on the level of confidence that can be placed on the results.
An analyst has a reason to make a measurement. The result of the measurement is used to inform the customer. This data is interpreted as information which contributes to knowledge. It can be either an estimate or with a degree of precision, it can be quantitative or qualitative, but whatever other attributes it may have it should be **fit for purpose**. If it is not fit for purpose, then there was no point in making the measurement in the first place, the data will be unreliable.
In the process unit 7 of Element 1 (see Chapter 5) of this teaching and learning programme, you learned how to attribute value to a result, in the form of confidence limits. The result can be tested for accuracy and/or precision, these are **performance characteristics** and will be discussed later with **method validation**.

Quality has to be planned and managed, often referred to as **Quality Management**.

Quality

'Fit for purpose', is a useful definition of quality, others include 'conformance to requirements' and yet others include a time element extending the definition to 'needs of the customer present and future'. In the context of analytical results, the customer is the user of the data, whether it be the production team in a pharmaceutical plant, or a research chemist using the result of a single analysis to inform the next stage of the research process.

The **ISO** definition of quality, covers most important points that should be considered by the analyst:

The ISO definition of quality

'The totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs'

ISO 8402

Quality Management

Quality management encompasses all aspects of managing quality in an analytical laboratory or a company. It includes **Quality Assurance (QA), Good Laboratory Practise (GLP)** and **Method Validation**.

A quality management system is set up to manage the quality policy, identify quality objectives, plan the processes and procedures required to achieve the standards of quality expected by the customer or **regulator**. It is a systematic management approach involving continuous improvement and assessment.

A quality management system may be large, including all aspects of the company, or it might be small, limited to one project. When written down the quality management procedures of the company or laboratory are referred to as the **Quality Manual**.

Scientists today work in a global community so a harmonised vocabulary is increasingly important. In this unit all standards referred to are international standards, and the **International Conference on Harmonisation** is referred to for validation terminology.

Benefits of Quality Management

Quality management ensures the infrastructure exits within which the analyst can supply meaningful data. The need for quality procedures in the analytical lab cannot be overstated. Poor quality data is a major cost to both business and society, an example of three quality drivers are:

- **Ethical:** E.g. pharmaceutical products are taken completely on trust by the patient. The manufacturer has a duty of care to supply safe and efficacious products to the consumer. Analysts want to contribute to good science.
- **Commercial:** In addition to due care, recognised quality standards allow mutual recognition of products across the globe, and are vital for free trade relationships. It is also in a business's interest to be recognised as a quality provider
- **Regulatory:** There are legal requirements to ensure the safety of products, international standards also allows comparability between regions.

Employee focus. A well motivated and trained workforce have greater job satisfaction and are also more mobile in a global market.

Need for quality in analytical measurements and consequences of failure

6% of a ‘first-world’ nations gross national product (GNP) is spent on making measurements, in the case of the UK this represents approx £7Bn per annum.

Wrongful convictions are sometimes due to unreliable science. Recently the Jill Dando case was reopened and the man originally convicted of her murder was acquitted by the Court of Appeal. A single particle of gunshot residue was given greater significance, or ‘evidential weighting’, than it should have and has lead to a successful appeal against the conviction.

The main suspect of the Omagh bombings was cleared and another murder trial adjourned when doubts were raised against the method of low copy number DNA, a method no longer deemed reliable in the UK.

Continued on the next slide

The pharmaceutical industry is a very well regulated industry, recent examples of unreliable data are thankfully rare, but tragic examples such as the Thalidomide case of the early 1960s are reminders of the horrific consequences of failure. Not all cases of bad data are due to incompetence or error, fraudulent data is also a possibility.

Fraud is not only faking or distorting data, but also omitting data.

Simply ignoring data because it does not support a hypothesis or the less confident scientist repeating an experiment until they ‘get it right’ and ignoring the many ‘wrong’ results obtained on their search for truth, must never be allowed. Whether intentional or by ignorance and neglect, fraudulent data exists where it should not.

Where two experts disagree, it should be a simple matter to test the evidence, but it is not surprising to find scientists with vested interests interpreting data to suit their own interests. The debates on the fluoridation of tap water, the tobacco industry and global warming, are good examples of two bodies of conflicting evidence.

Quality Assurance

Quality assurance (QA) is the over arching concept in quality management. It encompasses GMP (Good manufacturing practise), GLP (Good laboratory practise) and QC (Quality control). QA should influence original product research and development, through manufacturing, to analysis, data analysis and archiving. It can be described as **the planned and systematic control, which provides confidence that the data obtained from the analysis is of the quality required by the client.**

- Internal QA gives confidence to the management that the procedures are working.
- External QA gives confidence to the consumer and regulator.

The ***planned activities*** designed to ensure **Quality Control (QC)** procedures are being properly implemented, includes; documentation of procedures, policies, organisational relationships, responsibilities, resources, records etc. These are all included in the **Quality Manual** required for regulatory bodies such as **ISO**, or **UKAS** when applying for **Laboratory Accreditation**

Principles of Quality Assurance

The principles of QA are formalised in a number of published protocols:

- **ISO 17025**, the regulatory standard for chemical testing laboratories. Considers the technical competence of laboratories to carry out specific tests and calibrations. It is a core requirement for laboratory accreditation.
- **ISO 9001:2000** this is the international standard covering quality management for companies involved in production or services such as chemical analysis.
- Principles of **Good Laboratory Practise** (GLP) is concerned with organisational process and conditions under which laboratory studies related to regulatory work is carried out.

Quality assessment, audit and review.

Quality **audits** can be internal or external and can be **horizontal** or **vertical**. A horizontal audit looks at the overall quality management, whilst a vertical audit, trails individual samples from the point when they enter the laboratory to the reporting and archiving of data.

An external audit carried out by an independent external body, such as a regulatory body, is referred to as an assessment. The **audit** is a systematic independent and documented process for evaluating if specified requirements have been fulfilled. **Laboratory accreditation** is an important external systems audit. Accreditation is voluntary and maybe carried out by public or private national organisations. **Proficiency testing** is a form of external audit, based on inter-laboratory comparisons.

An internal audit checks if the quality procedures are in place and are being implemented. A **review**, also carried out internally, checks that the quality procedures are effective and achieves the stated objectives (i.e. the **quality policy** stated in the **quality manual**).

Regulation, Certification and Accreditation

- ISO guide 2 defines **accreditation** as the '**procedure by which an authoritative body gives formal recognition that a body or person is competent to carry out specific tasks**.' For an analytical laboratory in the UK the 'General Requirements for the Competence of Testing and Calibration Laboratories' are defined in **ISO/IEC 17025:2005**.

- In the UK, the United Kingdom Accreditation Service (UKAS) is the competent body that manages and awards accreditation in accordance with the ISO 17025 standard. This standard is recognised by other countries through an international agreement with accrediting bodies in the other countries.

- Laboratories that conduct regulatory studies in the Organisation for Economic Co-operation and Development (OECD) member states, must comply with the **Principles of good laboratory practise (GLP).** These are harmonised principles designed to promote the development of quality test data. In the UK, GLP is monitored by the Department of Health. GLP is an alternative system to the ISO, but both have key elements in common.

Good laboratory and manufacturing practise

The aim of **Good laboratory practise, (GLP)** is to encourage scientists to organise their work in such a way as to ensure the production of reliable results. GLP does not cover scientific aspects, it focuses on resources, procedures and protocols, characterisation of test systems, documentation and archiving and independent quality assurance. GLP introduces the **standard operating procedure**.

Good manufacturing practise (GMP) In the UK GMP is a guide, not a directive or regulation. In the US current GMP (cGMP) has a legal status. Manufacturers are required to **comply**, deviation from the code is considered **non-compliance**. GMP is beyond the scope of this unit, but in the pharmaceutical industry it is very important aspect of the quality system, and all analysts should be aware of it.

A Standard Operating Procedure (SOP)

Controlled Document- MTH 20 - Version 3 - 19/01/00

SOP: Sample Preservation & Storage for Testing - Effluent

1. All water samples should be tested as soon as possible after collection.
2. The following steps should be taken to reduce changes to the effluent sample due to chemical and biological activities:
   - Fill containers to the top
   - Store in darkness
   - Store at 4°C
3. Samples may be collected in glass or plastic (polythene) with the following exceptions:
   - Dissolved Oxygen - Glass Only
   - Phosphorous - Glass only, rinsed with 50% v/v HNO_3
4. Maximum storage times are as follows:
   - Conductivity - 28 days
   - pH - 2 hours
   - Turbidity - 24 hours (must be in dark)
   - Temperature - Immediately (done by client)
   - Dissolved Oxygen - fix then 8 hours
   - Nitrogen - 48 hours
   - Phosphorous - 48 hours
5. Tests must be completed within the maximum storage times for the results to be valid.

Comments on SOPs

The **standard operating procedure** is a document with details written instructions designed to achieve uniformity in the performance of a specific function.

The operator is not required to make any judgements and in a properly written SOP, there should be no ambiguities.

For further information on SOPs please read the following web based article:

http://en.wikipedia.org/wiki/Standing_operating_procedure

This article refers to the terminology as applied to fields other than Analytical Science and the following slide contains a copy of the two relevant paragraphs.

Clinical research

In clinical research, SOPs are defined by the *International Conference on Harmonisation* (ICH) as "detailed, written instructions to achieve uniformity of the performance of a specific function". SOPs are necessary for a clinical research organization—whether it concerns a pharmaceutical company, a sponsor, a contract research organization, an investigator site, an Ethics Committee or any other party involved in clinical research—to achieve maximum safety and efficiency of the performed clinical research operations. It is therefore a must that all people and sites involved in clinical studies (both at the sponsor and at the investigative sites) have appropriate SOPs in place in order to conduct clinical research and to ensure compliance with the current regulations. The ICH GCP (good clinical practice) Step 5 Guideline (Section 3.2.2) also suggests that an Institutional Review Board (IRB) have its own SOPs or written standard procedures. This itself proves that presence of SOPs are an integral part of the clinical trial at all levels. The presence of these quality documents is essential when inspections take place since the most frequent reported deficiencies during inspections are the lack of written SOPs and/or the failure to adhere to them. The risk of GCP non-compliance is high at organizations with a poor availability of clinical research specific SOPs and also if at all they are available the staff or the people for whom they were written are not either aware of them or their need. It therefore becomes very important for the staff to train them on these SOPs so that they are actually aware of why and how SOPs can play important role in fulfilling the ICH and other regulatory requirements.

Good manufacturing practice

An SOP is a written document / instruction detailing all steps and activities of a process or procedure. These should be carried out without any deviation or modification to guarantee the expected outcome. Any modification or deviation from a given SOP should be thoroughly investigated and outcomes of the investigation documented according the internal deviation procedure. All quality impacting processes and procedures should be laid out in Standard Operating Procedures (SOPs). These SOPs should be the basis for the routine training program of each employee. SOPs should be regularly updated to assure compliance to the regulatory requirements and the working practice. A minimum review schedule of 3 years is recommended. Changes of SOPs are in general triggered by process or procedural changes / adjustments. These changes should be managed by the internal site change control procedure. Part of the activity list of such changes should be to update the related SOP. SOPs should be in place for all Quality Systems plus the specific operational activities on site. The structure of an SOP System and the total amount of individual SOPs should be carefully taken into consideration. Too many SOPs could lead to a collapse of the SOP System. System SOPs should not be mixed up to keep systems and interaction between quality systems easy. [2] ISO 9000 is essentially simply a directive to document all procedures used in any manufacturing process that could affect the quality of the product.

The information on this slide has been copied from *http://en.wikipedia.org/wiki/Standing_operating_procedure*

Valid Analytical Measurement – the VAM Principles

Towards the end of the last century, comparative results were published which cast doubt on analytical results being reported by some of the major laboratories in the UK. Given the significance of incorrect data to industry, the DTI (Department Of Trade and Industry as it was at the time) funded the setting up of the VAM Executive, with the aim of improving the quality of analytical data. One of the first actions of the Executive was to publish a list of VAM principles, which are shown below:

VAM Principles

1. Analytical measurements should be made to satisfy an agreed requirement.
2. Analytical measurements should be made using methods and equipment which have been tested to ensure they are fit for purpose.
3. Staff making analytical measurements should be both qualified and competent to undertake the task.
4. There should be a regular and independent assessment of the technical performance of a laboratory
5. Analytical measurements made in one location should be consistent with those made elsewhere.
6. Organisations making analytical measurements should have well defined Quality Control and Quality Assurance procedures.

Quality Control

Definition of **Quality Control**
The operational techniques and activities that are used to fulfil the requirements for quality

Quality control (QC) procedures ensure the quality of the analysis. They include:

- Regular analysis of standards and/or reference materials
- Analysis of blind samples
- Quality control samples and control charts
- Blanks, spikes and duplicates.

Quality control should not be confined to the laboratory operations, it should involve all decisions which may concern the quality of the product, or process.

Principles of quality control

Quality control is essentially quantitative; it is based on comparisons of data, including uncertainties with standards and references. Typical activities include:

- Using control charts: Reference materials are routinely analysed and the results recorded on the chart.
- Inspection and adjustment of instruments and equipment to ensure correct performance - equipment calibration with **reference materials** and blanks.
- Studying the purity and stability of reagents to be used in the measurement process.
- Monitoring and maintaining the ambient conditions of the laboratory - temperature, humidity etc.
- Supervising the sample custody chain.
- Quality control checks, such as **blind** samples, are fed through the system to check specific aspects of the analysis process
- Assessing the results of samples passed through the process before and after any changes to the process.

Calibration and performance checks

Calibration is a set of operations that establish, under specified conditions, the relationship between the response obtained by a measuring system (or instrument), and the corresponding values obtained from **standards**.

Detailed coverage of calibration procedures was covered in Process unit 6 –see Chapter 4 of this teaching and learning programme

In physical measurement, such as length, weight, time etc. **primary** or **absolute standards** have been defined an agreed internationally. See the definitions of distance and time given in the text box below.

The standard meter is the distance travelled by light, in a vacuum, in 1/299792258th of a second. In turn the standard second is the duration of 9,192,631,770 cycles of radiation associated with a specified transition of the caesium atom.

Continued on the next slide

It is generally impractical to refer back to absolute standards such as those shown on the previous slide. Instead comparisons are made with more readily available standards, these are known as secondary, transfer or **working standards**. These will have been certified against a higher standard, e.g. weights or thermometers that have been certified in the UK by the National Physical Laboratory in Teddington.

The standard Kg, an Iridium-platinum alloy cast in the UK and held at the BIPM (bureau international des poids et mesures) Paris. Approved in 1889. The British prototype (No18) is held at the National Physics Laboratory in Teddington.

Figure 15.1- reference Kg weight

Every laboratory will have a program for calibration, designed to ensure that all measurements are **traceable** back to a primary measurement standard. This could be a working standard or a national or international standard such as a **certified reference material**. As standards are compared back through the standard hierarchy, with known uncertainty, to the absolute standard, it is possible to relate the result of a measurement directly back to a primary standard. This is referred to as **traceability.**

Measurement traceability

In an absolute sense, the ‘true’ value of an analyte can be defined only as being that value directly traceable to the base system of measurement units (SI) or their derivatives. ISO defines **traceability** ‘as the property or result of a measurement, whereby it can be related to appropriate national or international standards, through an unbroken chain of comparisons’. With chemical analysis, the most important base measurement is the kilogram and every analytical laboratory should have at least one analytical balance, whose calibration is directly traceable to the National standard of measurement as shown in figure (15.1).

However one problem that frequently occurs and must be taken into account, is the problem caused by the sample matrix. In many analytical situations, the matrix of the sample has a major influence on the determination of the analyte, resulting in measurement uncertainties considerably greater than that associated with the base SI unit.

Using the SI base unit as absolute or true, a procedure has been developed whereby traceability can be related to a hierarchy of standard substances, most of which are materials of similar character to the sample being analysed.

Continued on the next slide

Figure (15.2) illustrates the hierarchy, and shows that the role of each component of the hierarchy is to transfer accuracy to the level below it and to provide traceability to the level immediately above it. The definition of a certified reference material is given in the text box below. A photograph of a typical reference material is shown in figure (15.3).

Definition of a **Certified reference material**

A reference material, one or more property values are certified by a technically valid procedure, accompanied by or traceable to documentation issued by the certifying Body.

Figure 15.3 Soil sample CRM

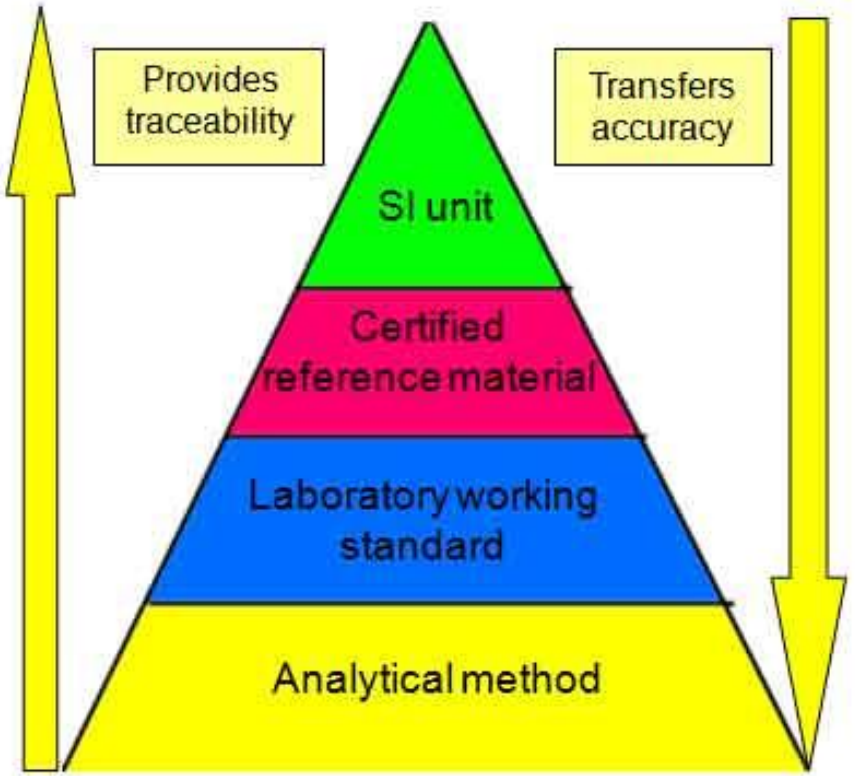
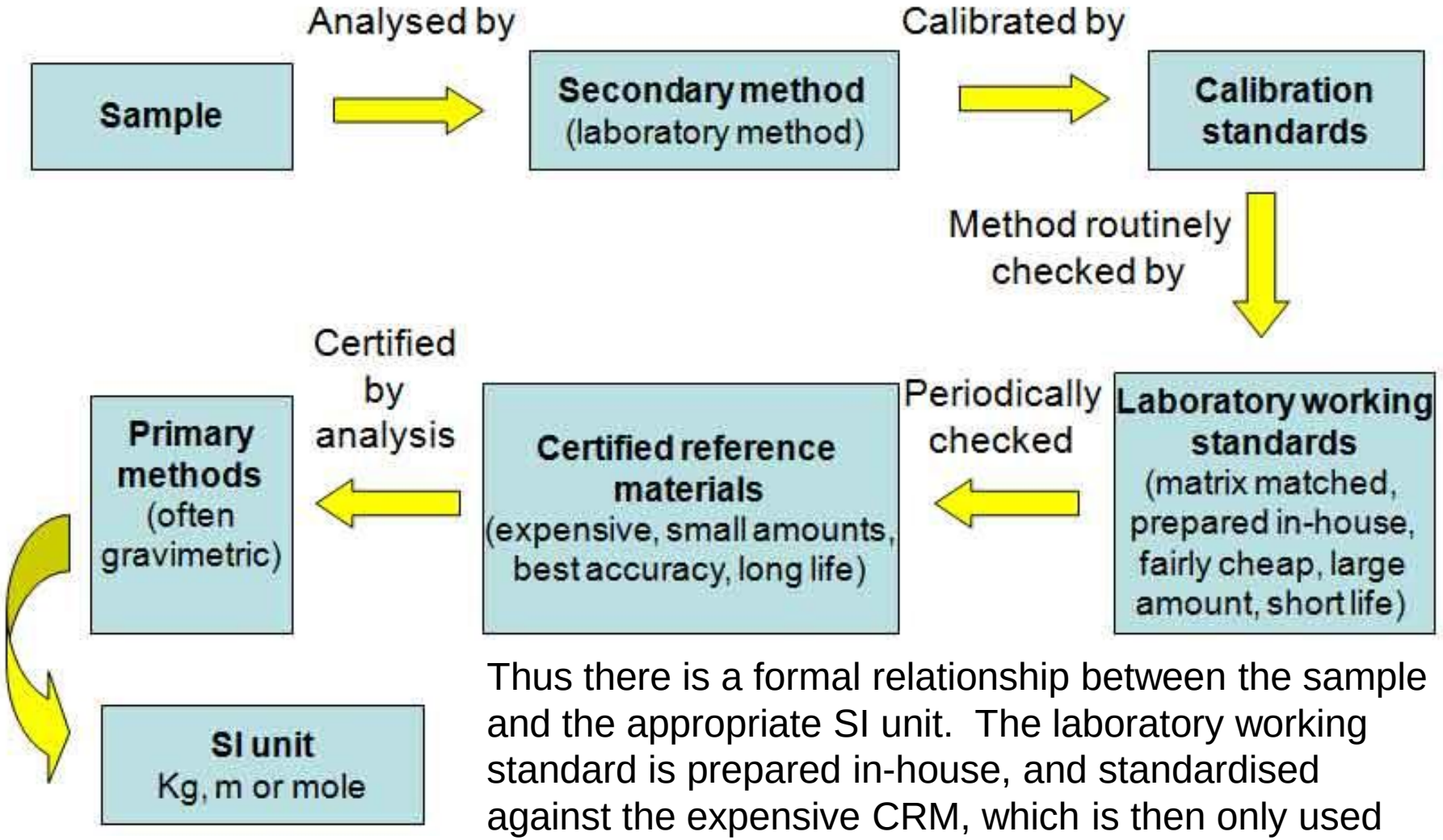


Figure 15.2 - hierarchy of traceability

The whole relationship existing between the sample and the appropriate SI unit can be illustrated diagrammatically, as shown below.



Thus there is a formal relationship between the sample and the appropriate SI unit. The laboratory working standard is prepared in-house, and standardised against the expensive CRM, which is then only used for periodic checks on the continuing validity of the method.

Reference materials & check samples

INTERNAL QC SAMPLES

Internal QC samples are prepared and quantity values of target components are checked against CRMs

CERTIFIED REFERENCE MATERIALS (CRMs)

SAMPLES SUPPLIED FOR PROFICIENCY TESTING

An accredited laboratory has to prove its performance by routinely analysing samples supplied by an independent laboratory

In Chapter 4 of this teaching and learning programme, procedures were described, to calibrate the eventual measurement stage in the analytical process. However none of the procedures described were capable of measuring the efficiency of the initial practical stages of sampling, sample preparation, separation and concentration, all of which can of course have a very major effect on the overall measurement uncertainty. By using CRM's, it is possible to estimate the total efficiency of the whole of the analytical process and consequently a meaningful value for measurement uncertainty.

For instance, if replicate samples of the Certified Reference Soil shown in figure (4b.3), containing known quantities of trace metals, was digested and the subsequent digest analysed by plasma emission spectroscopy, a comparison of the results obtained, compared to the certified values will provide a measure of the method uncertainty appropriate to the laboratory performing the analysis.

By repeating the analysis with soils of similar types, but containing differing abundances of the same trace metals, enables calibration of the whole method to be achieved.

Performance checks

Once a method has been validated, matrix reference materials or spiked samples are used as quality control checks. These are put through the analytical process to check the performance of the method. The results of these checks can be plotted on control charts such as **Shewhart charts**, which are useful in monitoring a system's performance. A typical control chart is shown of the next slide and further information on these charts may be found at http://en.wikipedia.org/wiki/Control_chart

Verification checks can also be carried out. These are simple performance checks to ensure a specified requirement has been fulfilled. E.g. a check weight on a balance verifies that the balance is close enough to the known value to be used [Figure (15.4)]

Figure 15.4 - checking the accuracy of a balance

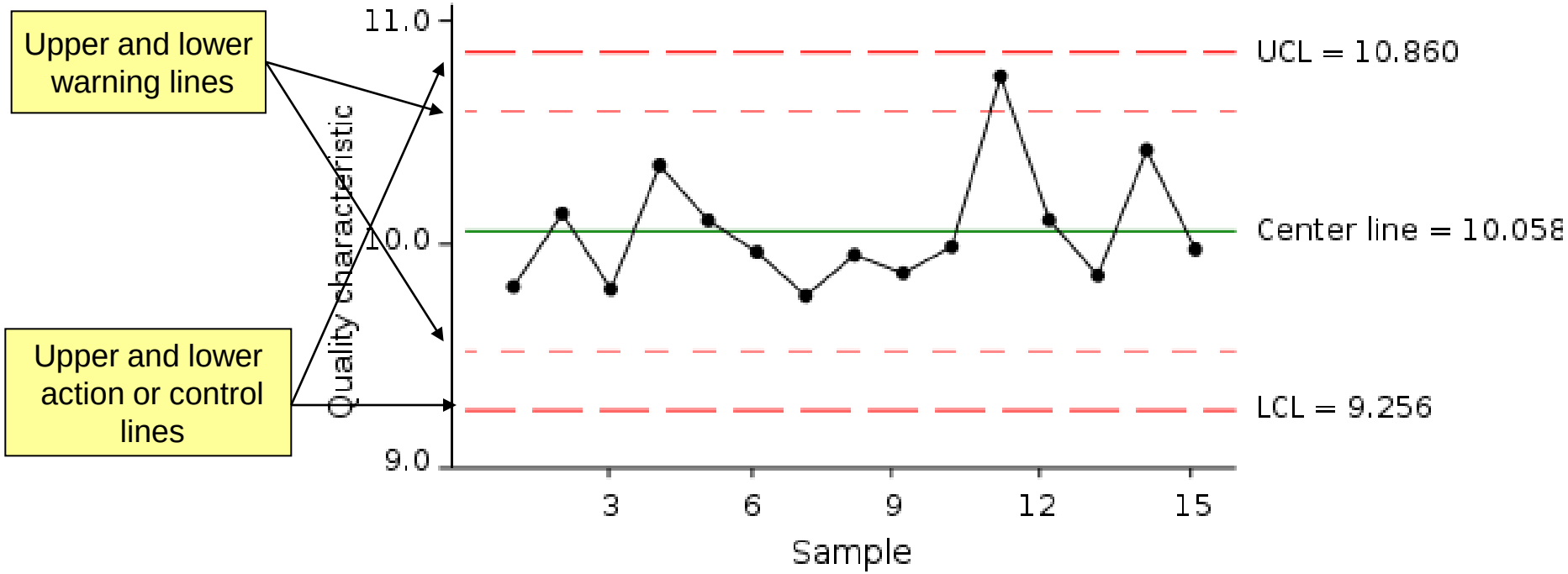


Figure 15.5 – Typical Control (Shewhart) Chart

Figure (15.5) shows a typical control chart with the green centre line as the target value for the product. Analysis of the product at timed intervals should produce results between the two warning lines. Consistent results above the warning lines should indicate that the process needs investigation. Results above the action or control lines could result in a process shut-down in order to identify and rectify the problem. In statistical terms, the positions of the two pairs of lines are governed by the following equations:
μ_0 = target value: $\mu_0 \pm 2\sigma/\sqrt{n}$ (warning lines): $\mu_0 \pm 2\sigma/\sqrt{n}$ (action lines), where ‘σ’ is the standard deviation for the measurement and ‘n’ is the number of samples taken for analysis at any one time point.

Method validation

Before embarking on the development of a new method, always research the chemical literature to see if a suitable one already exists. If a suitable one is found, it will still be necessary however to perform some **method validation** to prove that the method can be successfully adapted to **your laboratory, equipment and personnel**. More extensive validation is required when developing and proving a brand new method.

ISO definition of **Method validation**

The **confirmation** by examination and provision of **objective evidence**, that the particular requirements for a specific **intended use** are fulfilled'.

The three important terms in the definition of method validation, have been highlighted – **confirmation, objective evidence** and **intended use**.

Continued on the next slide

Confirmation is obtained by the comparison of the performance of the method with what is required. The requirement is usually laid out in the '**specific or intended use'** – different application of the same method may well have different validation requirements. The **objective evidence** is obtained by quantitative measure of the performance characteristics of the method

Methods in any field of analysis may be defined in terms of '**Method performance characteristics'** and it is these parameters, that are quantified during a method validation exercise. Remember that the 'method' encompasses all of the practical stages of the analytical process.

Method validation procedures may vary depending on the sector, e.g. pharmaceuticals, the food industry, environmental analysis etc. Useful guidelines are available from the **International Conference on Harmonisation** (**ICH**). The ICH deals with the technical requirements for the registration of pharmaceuticals for human use, but the approach to validation and definitions of validation terms and performance parameters are applicable to all sectors.

Performing a method validation

Method validation should be performed**:**

- When a new method is being developed;
- Before an existing method is adopted and used to analyse any samples;
- If there is any significant change such as a transfer to another laboratory, or change of application, (refer to **robustness** later);
- After a significant period of non-use.

The validation is carried out by an analyst with the support of a standard laboratory **protocol**. Groups of laboratories may collaborate and adopt methods developed in individual laboratories, these will be subjected to a peer review. Method validation is planned, following the strategies outlined on the next two slides. This planning results in a **statement of validation** which is the report asserting the method is fit for purpose, that it will produce meaningful data.

Continued on the next slide

The types of analytical method to be validated include:

- Qualitative methods such as an identification tests, e.g. an infra red of a degradation product – see figure (15.6)
- Quantitative analysis, such as a chromatographic assay for an active pharmaceutical ingredient (API).
- Limit tests for impurities such as for heavy metals in a food or pharmaceutical product.

www.pharmacopoeia.org.uk

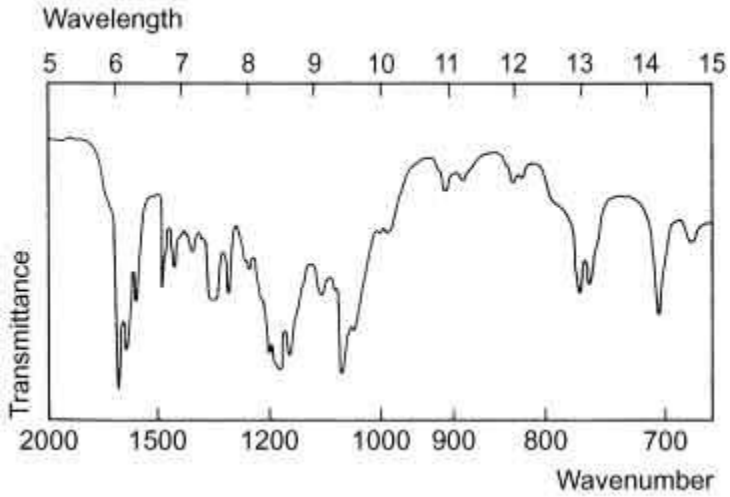


Figure 15.6 - typical IR spectrum to be used for spectral comparison

Examples of validated methods

Methods described in the official **pharmacopoeias** are deemed to be validated when applied to pharmacopoeial materials or products such as those produced by the pharmaceutical industry. Examples include the British Pharmacopoeia BP, the European Pharmacopoeia Ph. Eur. and the United States Pharmacopoeia USP.

Figure (15.7) shown on the next slide illustrates the tests that have to be carried out and passed, before a pharmaceutical can be graded 'BP'. The example highlighted is that for Paracetamol and tests to be carried out include:

- Identification
- Tests for related substances
- Heavy metals
- Loss on drying
- Sulphated ash
- Assay

In other fields such as environmental analysis methods described by Environmental Protection Agency (EPA) are used.

ASTM International is an organisation dedicated to producing standard methods, they are not a national organisation, nor are they involved in regulation. Standards from the Annual Book of ASTM Standards are recognised by many different organisations.

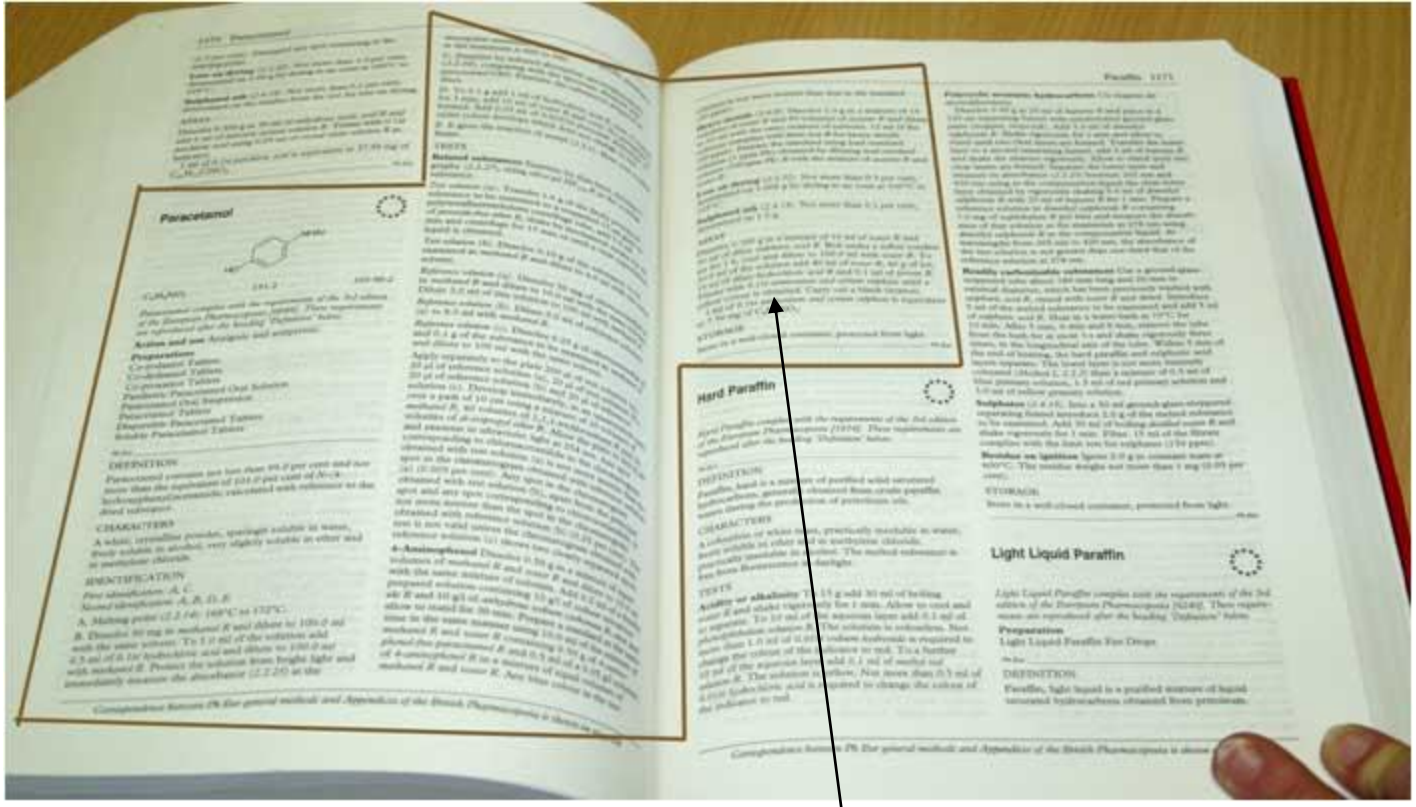


Figure 15.7 - example of the statutory tests to be carried out for a product to comply with BP Specifications

Copied by permission of British Pharmacopoeia

Below is a copy of the procedure to be adopted to **ASSAY** to BP specifications, a sample of paracetamol

"*Dissolve 0.300 g in a mixture of 10 ml of water R and 30 ml of dilute sulphuric acid R. Boil under a reflux condenser for 1 h, cool and dilute to 100.0 ml with water R, 40 g ice, 15 ml of dilute hydrochloric acid R and 0.1ml of ferroin R. Titrate with O.1 M ammonium and cerium sulphate until a yellow colour is obtained. Carry out a blank titration. 1 ml of 0.1M ammonium and cerium sulphate is equivalent to 7.56 mg of* $C_8H_9NO_2$" [Copied from British Pharmacopoeia and quoted as an illustration of the type of test typically carried out to assay a product to BP specifications]

Strategies for method validation

The very first stage of a method validation is to develop a validation **protocol**. This is the operating procedure for the validation. Within this procedure, the application, the purpose and the scope of the analytical method must be clearly described.

The **performance characteristics** and **acceptance criteria**, should be defined, and the validation experiments to be carried out should be described in detail. All materials (e.g. standards and reagents) to be used should be identified and qualified. Instrument validation should also be carried out, although details of these processes are outside the scope of this programme.

Method validation is an iterative process, with the method constantly being re-evaluated. Within this process, method parameters and acceptance criteria may also need to be changed. Finally the full internal (and external) validation experiments can be carried out.

Continued on the next slide

On completion of the validation experiments:

- Develop **standard operating procedures (SOPs)** for routine executing of the method;
- Define the criteria for re-validation;
- Define type and frequency of system-suitability tests and/or analytical quality control (QC) checks;
- Document all of validation experiments and results in the validation report.
- Archive all relevant documentation

Definition of a **Standard operating procedure**

Standard operating procedures are unambiguous instructions on how to carry out a wide range of operations. In this context the SOP would define the setting up and operation of specific instrumentation together with the carrying out of the full analytical method.

Defining a method to be validated

Prior to validation, the method needs to be precisely defined, the formal definition of the analytical method will include:

- Sampling, including sampling method, type of containers, pre-treatment, sample size and number of samples.
- Sample storage requirements.
- Equipment specifications and details of reagents to be used in the method.
- Definitions of reference materials to be used
- Performance tests of the analytical apparatus
- System suitability tests, such as resolving power of a chromatographic column
- Test conditions, an exact description of all reaction conditions.
- Advice on any safety precautions to be taken
- Description of the data analysis required, use of formulae, including instructions on how to carry out any statistical tests.

Method performance characteristics

Throughout this section of the Chapter you may wish to refer to Chapter 5 of this teaching and learning programme

These are also referred to as **performance parameters** and supply the **objective evidence** to support the confirmation that the requirements of the analytical method have been fulfilled. The characteristics to be evaluated are:

- Accuracy [trueness/bias - difference between the measured value and the true value]
- Precision [repeatability and reproducibility]
- Specificity and selectivity
- Linearity of calibration data and analytical range of the method
- Sensitivity [slope of the calibration curve]
- Limit of detection
- Limit of quantitation
- Robustness or ruggedness [sensitivity of the method's response to small changes in analysis parameters]

Trueness, accuracy and bias

In quantitatively analysing a sample, an estimate of the abundance or concentration of a selected analyte is being made. This is often achieved by comparison to a standard or selection of standards. However carefully the standard analyte has been prepared, it is never possible to be absolutely certain exactly how much is in it. Some of the analyte may have been lost, or the measurement technique may not be capable of determining the exact amount of analyte every time. There will always be a random variation and possibly a systematic variation associated with the measurement process.

Method performance characteristics associated with how close the results of the method are to the actual amount or concentration of the analyte include; **accuracy, bias and trueness**. Accuracy is used as a generic term to describe a technique or method which obtains the 'true' value. In modern analytical laboratories where nomenclature is very important, the term 'accuracy' has been replaced by trueness and bias.

An ISO definition of **accuracy** is "**the closeness of agreement of a test result and the accepted reference value**". So the term accuracy should only be used when dealing with an individual result.

Continued on the next slide

Where accuracy is a quality measure of an individual result, in most cases we are dealing with a set of results. To evaluate the performance of a method, **trueness** is more useful. An ISO definition of **trueness** is "**The closeness of agreement between the average value obtained from large set of test results and an accepted reference value**."

Trueness is measured in terms of **bias**, which is defined as "**the difference of the mean value of a set of measurements and the reference value**". Bias can be measured and expressed as:

$(\bar{x} - x_0)$

Where $\bar{x}$ is the mean of a set of measurements and x_0 is the reference value.

It is usually expressed, so that if the observed value is **greater** than the reference value then there is a **positive bias**. If **recovery studies** are used, then bias is expressed as a percentage, ie:

$(\bar{x} / x_0)$ X100

When evaluating bias, it is recommended at least seven replicate measurements are made.

Evaluating bias

The usual method of evaluating bias is to compare the mean of the data set $\bar{x}$ with a reference value. When the data has been collected, it should be inspected for outliers and suspect values. Then the mean and standard deviation should be calculated. To be statistically useful it is suggested that 7 replicates are made (n = 7).

The mean is then compared to the reference value x_0, using the **t-test**. The calculated value for t is obtained from the formula:

$$(\bar{x}-x_0)\ (\sqrt{n}/s) \qquad \text{Equation (15.1)}$$

and compared with the critical value of 't' obtained from tables for (n-1) degrees of freedom and at the 95% confidence level (two-tailed).

Bias represents systematic variation and to detect this, random variation must be minimised. This is achieved by using many replicates - in other words there should be sufficient **precision**.

Acceptable bias, is a value for bias considered small when compared to the method's precision

Recovery studies – example (15.i)

The routine laboratory is expected to quote measurement uncertainties and bias data for its measurements to allow customers judge the laboratories validity. In this example Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES) is used to evaluate major oxides in sediments. Sediment reference materials, 'Slate' and 'Chert' were digested and analysed to evaluate reproducibility and bias over a six year period. Figure (15.8) shows a samples of the reference materials being fused with a borate fusion mix.

Figure 15.8 - borate fusion of the reference materials

Table (15.1), shown on the next slide, represents the reproducibility data and bias data for a six year **reproducibility** study of oxide analysis in sediment following a lithium metaborate-tetraborate fusion and analysis by ICP-AES. **Reproducibility** is expressed as **relative standard deviation RSD**, and the bias as a percentage relative deviation from the reference value. Note bias is always quoted with a sign, to represent positive or negative bias and facilitate correction if the bias is deemed significant.

Continued on the next slide

Oxide	Matrix*	Reproducibility (RSD)	% deviation, control data	% deviation, extreme data
SiO_2	AGV-1	1.58	-0.12	+1.7
SiO_2	BHVO-1	1.68	+0.60	-1.64
Al_2O_3	BIR-1	1.9	+0.20	+2.6
Al_2O_3	G-2	2.0	+0.5	+3.2
Fe_2O_3	AGV-1	1.8	-0.2	+3.3
Fe_2O_3	BIR-1	1.7	+1.9	-2.4
Na_2O	G-2	2.9	-1.7	-2.7
Na_20	BHVO-1	4.0	+0.4	+2.2

Table 15.1 - reproducibility data over a 6 year period

The study used the same method and instrument throughout the six year period. However, there were changes in conditions, such as replacement torches and mirrors in the instrument, changes in solution composition at the fusion step, and variations in instrument set-up. The effect of such changes on a method is evaluated by a study of **robustness**, later. For the this study, the bias was not considered significant, and no correction was required. This study was used to evaluate a laboratories **standard operating procedures**.

Precision

According to the ISO definition **Precision** is **'the closeness of agreement between independent test results obtained under stipulated conditions'.**

As shown in Process Unit 7 (Chapter 5) of this teaching and learning programme, the **standard deviation** gives an estimate of precision - the spread or dispersion of data around a mean. The deviation of each result from the mean, estimates the **random** variation. There are three different levels of precision:

- **Repeatability:** Inter-assay precision, the data is obtained over a short time interval using the same operating conditions. This is the level of precision expected from set of replicate determinations.
- **Intermediate Precision:** This level incorporates variations in conditions such as different analysts, different equipment, over a longer timescale. When measuring precision at this level it is important to design your experiment to monitor the effects of these variations. This is the most appropriate level for setting acceptance limits for routine analysis in a quality control laboratory.
- **Reproducibility:** This is based on collaborative studies where the determinations are carried out in different laboratories, over a much longer timescale. This precision will reflect variation from a wider range of sources.

Continued on the next slide

When measuring precision as part of a validation protocol remember to state the conditions under which the precision has been evaluated.

Once the validation experiment has been agreed, i.e. the experimental conditions, then a number of **independent** determinations are made. Usually not less than 6, or more than 15. From these values a mean and standard deviation is obtained.

NOTE: Each independent determination should represent all stages of the Analytical process, including sampling, sample preparation (such as extraction), as well as the measurement step. Replicate determinations of a extract, for example, is not sufficient.

If there is a range of typical sample means then precision should be measured at 3 levels in that range, with 3 determinations each.

Precision may then be quoted as a **standard deviation** SD, a **relative standard deviation** RSD, or **standard deviation of the mean**, ($s/\sqrt{n}$).

There may be a reference to the level of precision measured such as **'repeatability standard deviation'** or '**repeatability precision**' s_r (%CV).

Continued on the next slide

Precisions are used to compare methods, laboratories, instruments etc.

Precision is independent of trueness. Remember: Precise data does not mean true data.

Precision is related to **sensitivity**. The lower the analyte concentration, the greater the measurement sensitivity required and the lower the precision obtained. If in a trace analysis an additional pre-concentration step is added to compensate for a lack of sensitivity, then as more sources of variability are added the precision is lowered accordingly. **Limit of Detection, LoD** is determined using a precision dependent parameter, the standard deviation of the blank.

For **qualitative analysis**, where there may be just a yes/no answer, as in limit tests, precision cannot be expressed as a standard deviation. It can however, be expressed as true and false positives, (and negatives). Please note in some disciplines, (e.g. microbiology) this terminology may conflict.

Example (15.ii)

A Karl Fischer titration was used to determine the water content of rice. In this experiment two sets of precision data were collected, the aim was to determine the **confidence level** at which an individual result could be accepted. One set of data was collected under **repeatability** conditions (Study A) And the other under conditions of **reproducibility** (Study B).

The data:

A Water content (%): 12.40, 12.39, 12.37,12.42, 12.39 and12.40.
B Water content (%): 12.45, 12.35, 12.39,12.47, 12.32 and 12.39.
Step One: determine mean and standard deviation:

	n	**mean**	**sd**
A	6	12.395	0.016
B	6	12.395	0.057

The standard deviation under conditions of reproducibility is over three times that of the standard deviation of the experiment carried out under repeatability conditions.

Is this surprising?

Continued on the next slide

The value for students-t, using a 'two-tailed' t for 5 degrees of freedom at four different confidence levels is:

% confidence	90%	95%	98%	99%
Significance level (P)	0.1	0.05	0.02	0.01
t	2.02	2.57	3.36	4.03

Step Two: Based on the formula for confidence limits [Mean ± t x ($s/\sqrt{n}$)], the calculated limits at the four confidence levels are:

% confidence	90%	95%	98%	99%
Study A	0.013	0.017	0.022	0.026
Study B	0.047	0.060	0.078	0.094

Thus the confidence interval of the mean values at 95% is:

Study A 12.395 ± 0.017
Study B 12.395 ± 0.060

Discussion of results from example (15.ii)

Note the ranges obtained in the example above. The narrowest limits are for confidence in study A at the 90% level, and the widest limits are for 99% confidence in study B. This simply means that the precision is better in the repeatability study, and the more confidence required the wider your acceptance criteria must be.

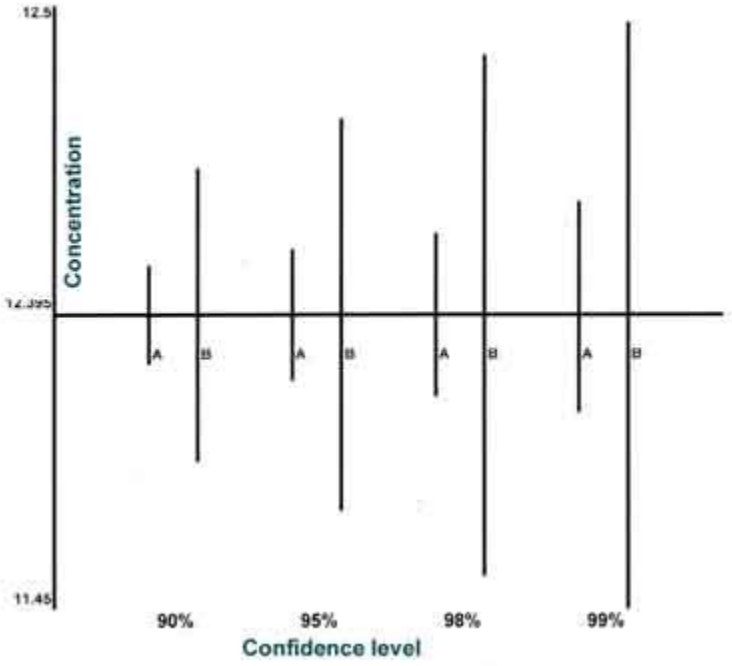


Figure 15.9 - confidence limits

Selectivity and specificity

At the validation stage, tests are carried out to ensure that only the analyte of interest is being measured. Both of these terms, **sensitivity** and **selectivity** describe the extent to which the analytical method responds to the analyte and the analyte only. Both of these terms refer to the ability to measure the analyte in the presence of other potentially interfering substances that may be present in the matrix, such as impurities, degradation products etc.

In quantitative methods the term **selectivity** is often used with another substantive term i.e. ‘coefficient’, ‘index’, factor’ etc. for the quantitative characterisation of interferences.

In a chromatographic separation, **selectivity** is assessed using metrics such as resolution, plate count N, and tailing factor. **Diode array detectors** (**see Chapter 7 of this teaching and learning programme**) are also very useful in identifying potential interferants.

In potentiometric analysis using Ion Selective Electrodes (see Chapter 9), selectivity is assessed by individually measuring **selectivity coefficients** for ions, other than the analyte ion, that are likely to be present in the sample.

Continued on the next slide

ICH uses the term **specificity**, describing it as the ‘ability to assess **unequivocally** the analyte in the presence of components that maybe expected to be present’. This terminology may be found in FDA (Food and Drug Administration) publications and the USP (United States Pharmacopoeia).

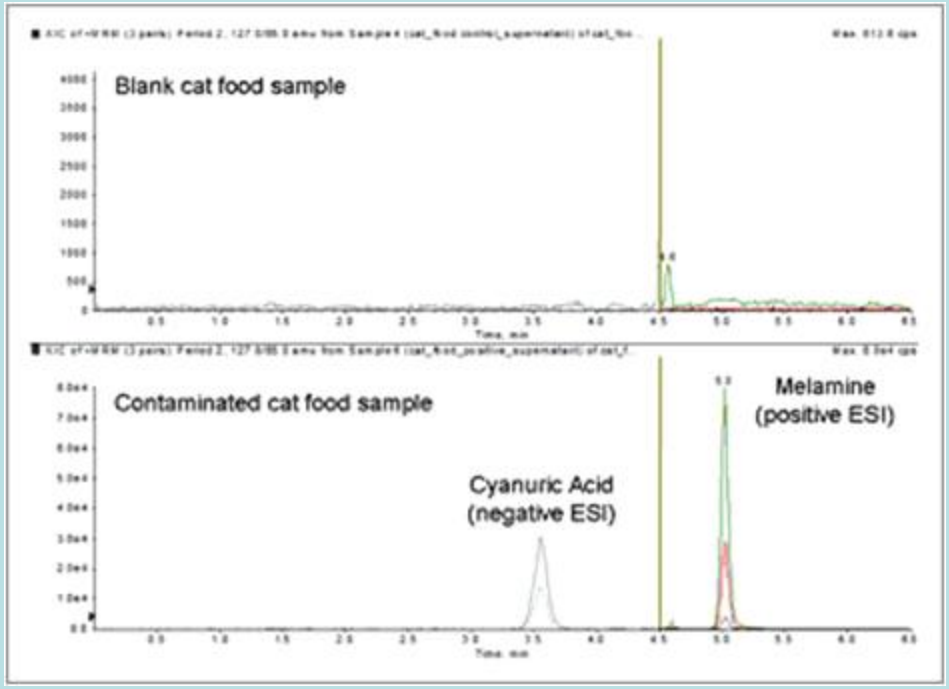
IUPAC promote the term **selectivity** and define it as **the extent to which a method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behaviour**.

IUPAC acknowledge that the term specificity is used throughout the analytical community, but discourage the use of the terms interchangeably and have referred specificity as the ‘ultimate of selectivity’ as few if any analytical techniques are truly specific.

Very selective analytical methods are currently in use and improving all the time, LC-MS-MS is particularly selective and can be used in legal cases, where the requirements for non-biased identification are high.

Example (15.iii)

Recently huge quantities of pet foods were recalled due to the contamination of imported Chinese wheat gluten, by melamine and cyanuric acid. A **selective** analytical technique to determine these compounds in pet food was required. By using LC-MS-MS, a highly selective technique it was possible to show conclusively the presence of these contaminants in the cat food.



LC-MS-MS of Melamine and Cyanuric acid in petfood.

Analyser

Measuring selectivity

Measuring selectivity requires knowledge of your sample matrix and sample chemistry. All possible interferents must be identified, those that contribute to the sample signal as well as those who attenuate it. Individually and in combination.

Many selectivity experiments are qualitative. Such as identification and purity tests to assess all impurities in the matrix.

There are some quantitative examples for instance **matrix spikes**. A known amount of the target analyte is added to a 'field sample' i.e. a real sample matrix. Three analytical measurements are obtained:

- The amount of analyte in the field sample (A_{fs});
- The amount of analytes in the spiked sample (A_{ms});
- The amount of analyte added to the field sample (A_a);

The recovery of the matrix spike can be calculated by:

$(A_{ms} - A_{fs})/ A_a \times 100$ Equation (15.2)

If the matrix recovery does not fall within established acceptance criteria, then it is likely there are interferents present in the sample matrix.

Working range

The **working range** of an analysis, sometimes called the **dynamic range**, is the concentration range over which the method gives test results related to analyte concentration. [**Please refer to figure (4.15) on slide 68 in Chapter 4 of this teaching and learning programme**] The **upper end** of the working range of an analysis is identified by a significant lowering in sensitivity. This could be a tailing or a plateau. The **lower end** of the range is considered to be the **Limit of Detection** or **Limit of Quantitation** depending on whether the analysis is quantitative or qualitative.

During validation, the working range is assessed to check that it corresponds to the concentration range required in the analysis. Throughout the working range there should be an equation to describe the relationship between the method response and concentration, this is referred to as the **calibration function**. This relationship is usually, but not always, linear.

If the calibration curve is non-linear, it should be repeatable from day to day.

Note: within the pharmaceutical sector, the working range should cover ± 20% of the expected analyte concentration.

Establishing a working range

A calibration curve is established by determining the analyte concentration in **blanks**, **reference materials**, or **fortified blanks** across the expected working range.

At least 10 different concentrations should be used, spread evenly over the working range, at least 6 of these should be used within the linear region. When preparing a calibration curve, matrix-matched certified reference materials or spiked samples can be used. The calibration standards need not be independent of one another.

If the calibration response is not linear, then some form of mathematical transformation may be required.

Then linearity is tested. **Refer to Process unit 7 in Chapter 5 of this teaching and learning programme.**

A graph of response *versus* concentration is plotted and visually inspected, **Note** it is not trivial to remove outliers from a calibration curve. At this stage it is best to identify them, by residual plots and Dixons test, but not to remove them.

Variance should also be tested across the working range.

Testing for linearity

When assessing the linearity of the calibration curve, the first test is often visual inspection. Regression coefficients have limited value for assessing linearity, however significance tests can be applied to set acceptance criteria for **regression (r)**.

Residual plots are very useful for assessing linearity and they can also help to identify potential outliers. **Refer to Process unit 7 in Chapter 5 of this teaching and learning programme.**

The calibration curve comprises three parts:

(i) At low analyte concentrations, where the analyte cannot be detected. If the blank does not give response then the curve intercepts the Y axis at X=0;
(ii) The linear range, starting at the LoD, and finishing when it 'deviates from linearity';
(iii) At very high concentration of analyte the signal no longer varies with increasing concentration and is parallel to the concentration (X) axis.

Using the calibration curve

Once the linear and working ranges are established the calibration curve can be used to determine the amount or concentration of analyte in a sample.

How often calibration curves need to be repeated should be defined in the validation documents. If the response is linear and through the origin, then it maybe acceptable to use a response ratio or a calibration factor, instead of running the full calibration curve.

The calibration factor, if used, should be verified at least daily, using calibration standards. **Refer to Chapter 5 (Process unit 7) of this teaching and learning programme.**

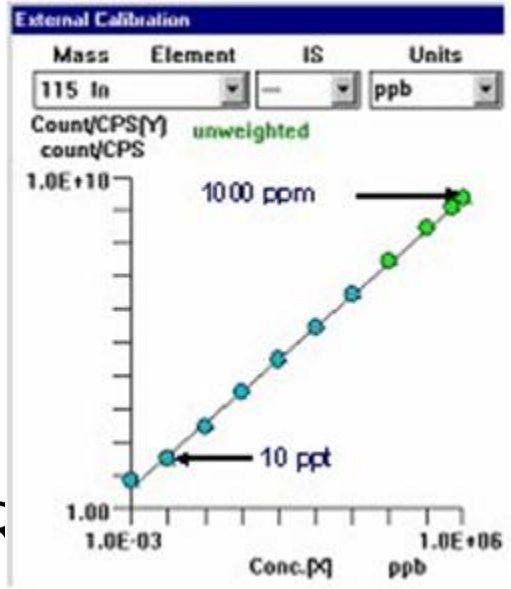


Figure 15.10 - a calibration curve for an new ICP-MS showing a wide dynamic range, from 10 ppt to 10,000 ppm.

Note: This covers the combined range of conventional ICP **and** GFAAS.

Examples of working ranges

Some techniques have very wide working ranges such as Ion Selective Electrodes. A modern combined Fluoride ISE, has a working range from 1 to $1x10^{-6}$ mol/dm^3

Sometimes the useful working range is quite narrow. GF-AAS has a useful range of 0.1 to100 μg/l, but its advantage is its low limit of detection.

ICP instruments generally have linear ranges from μg/l to mg/l, and ICP-MS from pg/l to μg/l

Application	Specification range (%)	Validation range (%)
Release Assay	95-105	80-120
Assay of API in a stability study	90-105	80-120
Content uniformity test	75-125	60-140
Assay of preservative in a stability study	50-110	40-120
Assay of degradation pdts in a stability study	0-10	0-20

Table 15.2 - verification of linearity for pharmaceutical applications:

Sensitivity

The IUPAC definition refers to **sensitivity** (S) as the 'slope of the calibration curve', a function of analyte concentration. Put simply, sensitivity is the ability of an analytical method to respond to small variations in analyte concentration. This can be expressed as the 'change in response of a measuring instrument divided by the corresponding change in the stimulus'. This describes the variation in analytical signal **y** with the analyte concentration, **C**:

$S = y/C$ Equation (15.3)

$S = \Delta y/\Delta C$ Equation (15.4)

The greater the change in the signal produced by a low analyte concentration, the greater the sensitivity.

Where '**y**' is plotted against '**C**', then:

$\mathbf{y} = S \cdot \mathbf{C} + \text{intercept}$ Equation (15.5)

I.e. the slope of the calibration curve.

Measuring sensitivity

The measurement for sensitivity is the slope of the calibration curve, as described on the previous slide. Inevitably, there is a **measurement uncertainty** associated with measuring the response of the analytical system, u(r), and this can be used to measure the quantitative resolution of a method or technique.

Please refer to Process unit 7 of Chapter 5 of this teaching and learning programme

Quantitative resolution (QR) of a measurement system for concentration can be determined using the sensitivity (S) of the technique and a measurement uncertainty:

$QR = u(r) / S$ Equation (15.6)

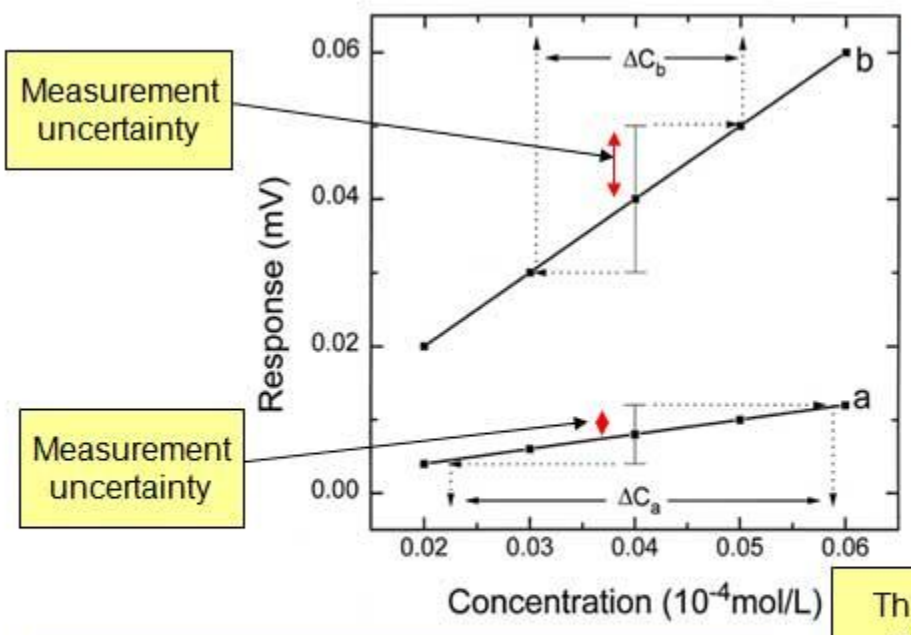
$\Delta C = u(r)/S$ Equation (15.7) Note [**ΔC is shown on the next slide**]

Sensitivity and measurement 'error' are not independent, see illustration overleaf.

Note: The sensitivity of a technique refers only to the measurement process.

Continued on the next slide

Two methods A and B are described here by their calibration curves a and b. It is clear from the plots that method B is more sensitive as the slope illustrates a more significant change in response to a smaller change in concentration. The **measurement uncertainty** for method A is 0.004 mV, but larger at 0.01 mV for method B



$\Delta C_b = 0.01/(1 \times 10^4) = 1 \times 10^{-6}$ mol/L

Notice that ΔC_b is less than ΔC_a, so when the 'error' bars are projected onto the concentration axes, the uncertainty of method B is actually less:

$\Delta C_a = 0.004/2 \times 10^3 = 2 \times 10^{-6}$ **mol/L**

S is 2×10^3 mV L/mmol for curve a, and 1×10^4 mV L/mmol for curve b.

This illustration shows that the apparent measure of measurement uncertainty can be misleading in that it is related to the sensitivity of the method. Thus method 'b' has a lower measure of uncertainty than method 'a'.

Figure 15.11 – relationship between method sensitivity and measurement uncertainty

Limit of Detection

The **limit of detection** is the lowest amount of analyte that can be detected by an analytical method. It should be expressed as an amount or concentration, $C_{LoD.}$
The signal y_{LoD}, corresponding to C_{LoD}, should be statistically distinguishable from the blank or background signal. The blank signal is measured, y_B and a standard deviation, sd_B, is calculated.

$y_{LoD} = y_B + 3sd_B$ Equation (15.8)

C_{LoD} is then calculated from the calibration curve.

The factor '3' is statistically relevant, as there is a 99.865 % probability that the blank signal does not exceed the LoD, assuming normal distribution.

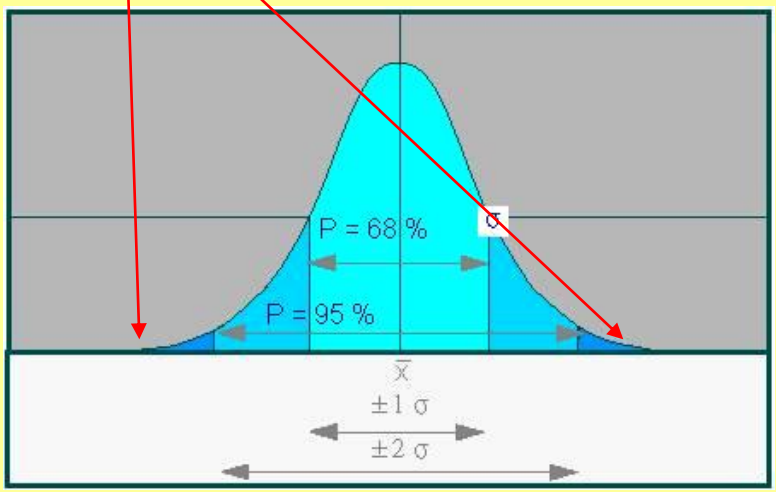


Figure 15.12 - Gaussian plot showing results within 3 standard deviations of the mean

Other approaches to calculating LoD are to be found on the next 5 slides.

Regression approach to calculating limit of detection

There is a significant amount of literature describing different approaches to calculating LoD from real sample data. Another one of which may be useful, is the **regression** approach.

Instead of independent replicates of the blank, the intercept on the Y axis for X=0, is used (q_0). A statistic $s_{y/x}$ is calculated, this represents the variation in signal values (y_i) around the fitted values, or residuals. The LoD, measured as a concentration, is obtained from:

$$LoD = q_0 + 3\, s_{y/x} \quad \text{Equation (15.9)}$$

$$s_{\frac{y}{x}} = \left\{ \frac{\sum_i (y_i - \hat{y}_i)^2}{n-2} \right\}^{1/2}$$

See slide 42 in Chapter 5 of this teaching and learning programme, for an explanation of this equation

This assumes a linear relationship between response and concentration, LOD may also be determined from non-linear calibration plots, but is beyond the scope of this unit.

Continued on the next slide

In some analysis it is necessary to define a **critical value**. This is an instrument response which triggers an action. When referring to a critical value in this context, it means that below the critical value, the sample does not contain analyte.

The LoD is used to define the lower operating concentration of a method and to facilitate the selection of decision criteria such as critical values.

When applying critical values to methods such as limit tests, it is important to be able to state if a sample exceeds the limit or not. There is a requirement for reliability even in qualitative analysis. The frequency of false positives, wrongly declaring a substance to be present and false negatives, wrongly declaring a substance to be absent, are measured to give a certain level of reliability.

LoD is often defined as the critical value where the false negative rate is 50%.

Measuring the limit of detection

To measure the **limit of detection** for a whole method, as opposed to solely for an instrument, the sample blanks must be independent, i.e. separate measurements each requiring the sample blank to be taken through the complete method.

Ten measurements of the blank (all separate blank samples) are required; in some applications low level spiked samples or standards are measured and diluted until approximately 50% of the results show that no analyte is present.

The standard deviation is calculated, and equation (15.9) is used to determine the LoD.

Limit of detection, and critical values, can also be determined for a particular step of the measurement process such as for an instrument. If using a calibration curve, an average value for the blank can be determined from the intersection of the response (Y) axis at concentration (X) = 0 as above.

Evaluating limit of detection from instrument outputs

It is also possible to determine LoD from the instrument output itself. For example a chromatographic peak against a noisy background. The standard deviation of the noise S_n is measured and LoD determined from:

$LoD = baseline + 3S_n$ Equation (15.10)

S_n can be estimated by measuring the height of the peak to peak noise, see figure (15.13) below.

This is a valid method of determining Limit of Detection, and can also be used to set critical values which are useful when designing experiments, see example on the next slide.

Figure (15.13) shows a peak on a noisy baseline. The peak to peak noise is represented by the blue lines.

If the peak to peak noise level was 6 mm, this will give a S_n value of 3 mm so that the LoD will be any signal that is 9 mm above the baseline

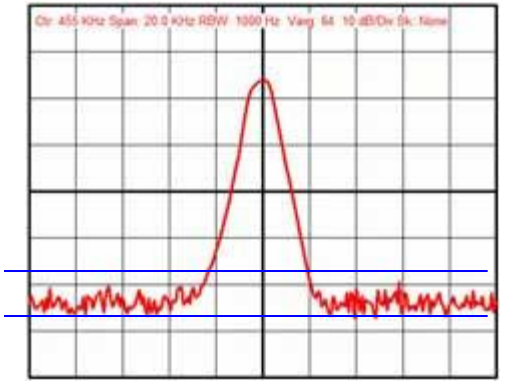


Figure 15.13 – noisy instrumental output

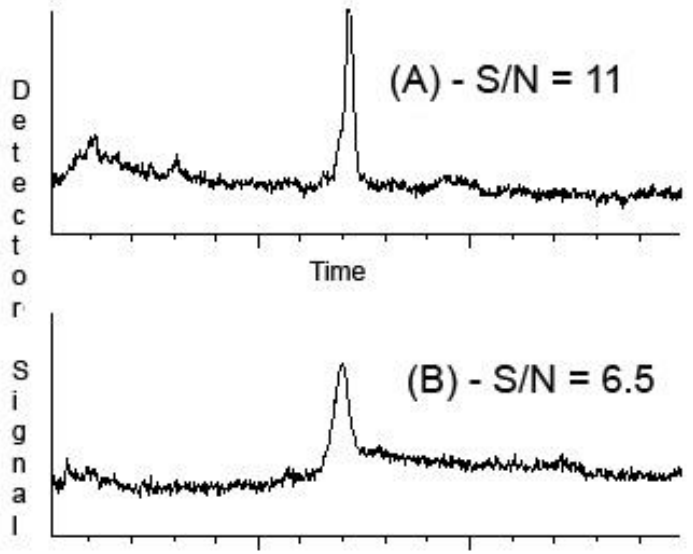
The standard deviation of the noise can be estimated in more than one way depending on the available data. Either by 'eye', in other words the value of the 'peak to peak' noise, see figure (4b.12) above, as the **population SD** for more than 30 measurements of the baseline, or by:

Baseline +(3 X 0.71) (*peak to peak noise*)/2

The assumption in this latter case is that the noise is approximately sinusoidal, (0.71 comes from a sine function ,sine 45 degrees).

The Signal to Noise Ratio (SNR) is another term you might meet when studying LOD.

To determine the signal to noise ratio, SNR, measure the height of the 'peak to peak noise' of the baseline, and the peak height of the analyte peak.
SNR = peak ht/peak to peak noise



Analysis of sample of Tamoxifen by hplc using C18 stationary phases from 2 different manufacturers

Figure 15.14 – measurement of signal to noise ratio

Limit of Quantitation

The **limit of quantitation** is the level at which the uncertainty becomes acceptable. Normally it is calculated at:

$X_{LoQ} = X_B + 10sd_B$ Equation (15.11)

For trace analysis a precision to 10% CV, is useful, this is why $10sd_B$ is chosen.

The factor 10 has no statistical significance, and was chosen to estimate the LoQ. In fact it is a slight overestimate as statistical studies have shown that a factor of 9 would be sufficient.

The most common practical measurement of LoQ is to measure 10 replicates and determine an average response for the blank, Y_B and s_B and determine the LoQ from the calibration graph.

During validation, other significant validation **CV%** parameters such as precision, should be evaluated for concentrations close to the LoQ and LoD.

If the required precision at the LoQ has been identified, then the LoQ could be determined by measuring the blank value for at least 6 samples over the required range and plotting CV% *vs*. concentration. Then reading the concentration equivalent to LoQ at the required precision.

Example (15.iv) - Total Organic Carbon from a cleaning validation in pharmaceutical manufacturing

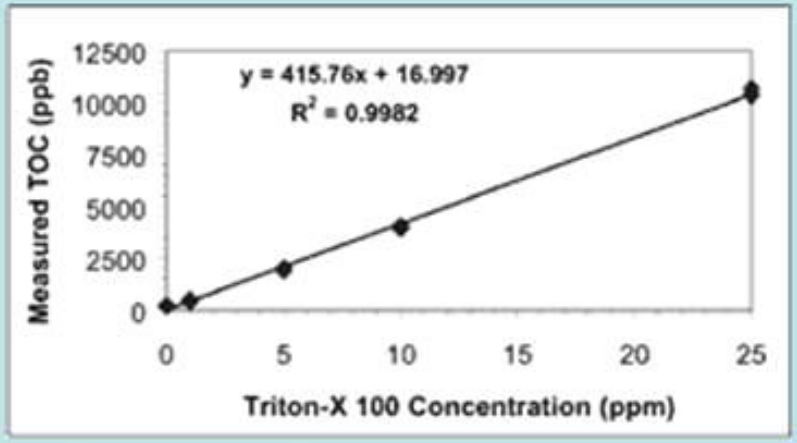
It is not only analytical methods that should be validated, Good Manufacturing Practise, **GMP**, requires validation of the cleaning of pharmaceutical manufacturing equipment. This must **demonstrate** that the equipment is free from residual drug substances and cleaning agents. Total Organic Carbon analysis, TOC, is a non-specific method capable of detecting all carbon containing compounds, including active ingredients (API), excipients and most cleaning agents. The method involves the oxidation of carbon containing molecules to CO_2. In this example the rinse water sample is acidified to release 'inorganic' carbon as CO_2 to be determined by a non-dispersive Infra-red detector (NDIR). Remaining carbon, assumed to be TOC is then oxidised to CO_2, and this is also determined by NDIR.

For validation two types of sampling, swabs and rinse water were tested and the parameters: **Linearity, LOD & LOQ, Precision and Accuracy,** and swab **recovery** were measured.

High pressure washing chamber

Continued on the next slide

Despite foaming, wetting agents such as triton-X gave acceptable linearity, illustrated here with a correlation coefficient of 0.9982



Vial Number	Average TOC (ppb)
1	58
2	72
3	75
4	93
5	79
6	102
7	60
8	83
9	67
10	54
Average	74.3
Standard Deviation	15.5
MDL (Student t, n=10)	50 ppb
LOQ	151 ppb

Calculated TOC averages from 10 blank vials.

In this table, 10 blanks comprising, vial, swab and recovery solution were tested. Four replicates for each vial. The 'Method Detection Limit' was calculated:
SD * 3.25 (Students-t for 10 samples 99% level).

Recovery of an API was measured by 'spiking' stainless steel plates with the substance. After drying the plates were swabbed, as if sampling a manufacturing container, and TOC levels were determined:

Substance	ppm C of Spiked Standard Solution	ppm C of Spiked Plate	% Recovery	% RSD
CIP-100	1810	1710	94.5	1.8
Sucrose	2663	2112	79.3	4.9
Vancomycin	661	634	95.9	3.0
Endotoxin	902	736	80.0	2.8

Representative examples of swab recoveries from cleaning agents and active substances.

Robustness (ruggedness)

Ruggedness (or **robustness**) is a measure of how well an effective analytical technique stands up to less than perfect implementation.

The ICH definition of Robustness

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

These changes include variables in the experimental conditions such as pH, temperature, small changes in mobile phase etc., in addition to environmental variances.

Once the variables that could have a significant impact on the method's performance have been identified, then experiments can be set up to quantify them, with the ultimate aim of excluding them.

Continued on the next slide

Measuring robustness

Robustness is evaluated by making deliberate variations to the method and investigating the subsequent effect on performance. **Robustness testing** could be carried out assessing each effect separately. However, this is time consuming. As most procedures facing validation are well defined anyway, the effects are small, it is possible to vary several parameters at the same time and assess the impact.

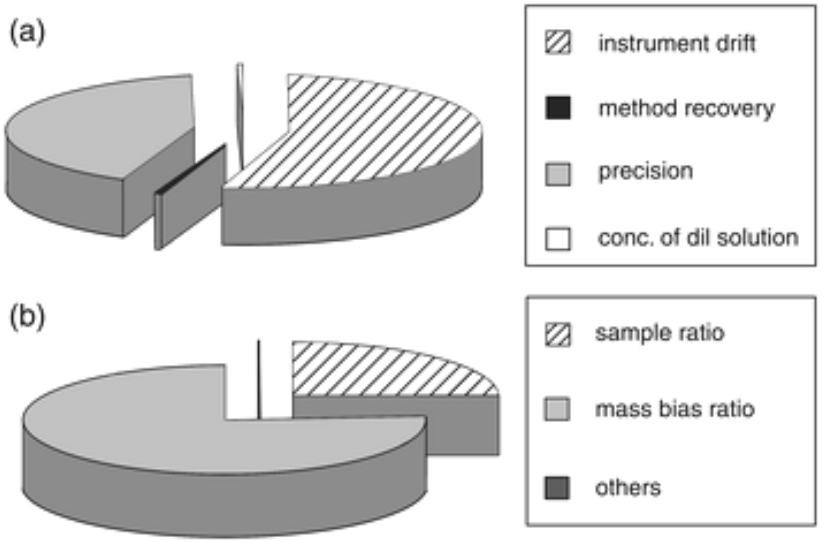
Plackett-Burman design allows seven independent factors be evaluated in eight experiments. It is an experimental design technique used to test the impact a change in a variable will have on the results of an experiment. Used in experimental design as well as robustness testing.

Once the variables that have the most significant effect are identified, they are ranked in order of greatest effect, then the **SOP** can ensure they are tightly controlled.

A description of Plackett Burman experimental design is beyond the scope of this programme.

Measurement uncertainty and validation studies

Measurement uncertainty is not a formal stage of method validation, but it should be considered throughout. Information from the validation can be used when evaluating the **measurement uncertainty budget.** Measurement uncertainty has been described in Chapter 5 (Process unit 7) of this teaching and learning programme. The estimate of uncertainty should include all parts of the analysis which contribute to the error in the final measurement. An example of how this can be displayed in practice is illustrated in figure (15.15).



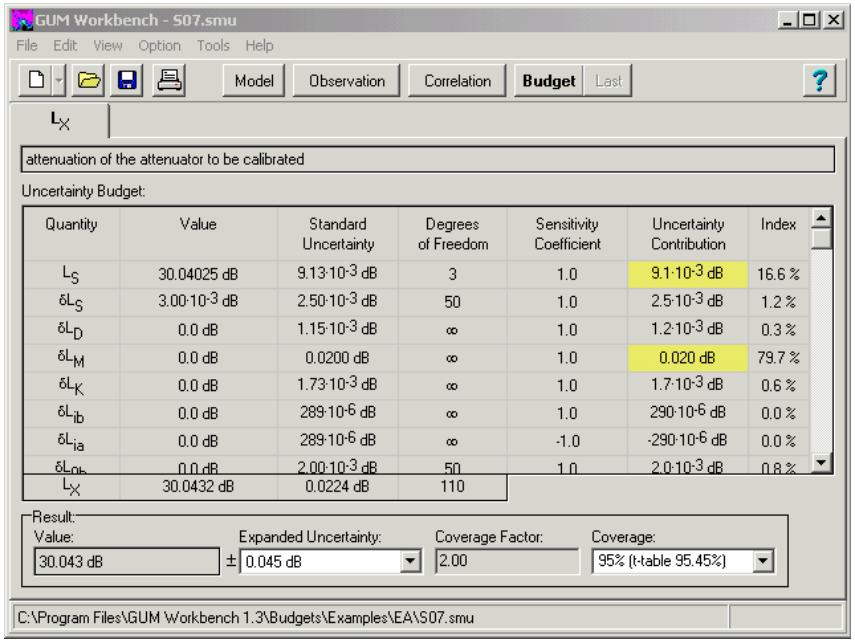


Figure 15.15 - uncertainty measurement

Question 15.1 Write out a standard operating procedure (SOP) for the preparation of 100 cm3 a 0.1M solution of sodium chloride.

Question 15.2 Though not an analysis, the preparation of an analytical reagent can comply with the VAM principles. Go through the SOP you prepared in Q1 above, consider where each of the six VAM principles applies.

Question 15.3 Reflect on the ISO definition of Method Validation, write an explanation of and give examples for each of the following: (a) confirmation, (b) objective evidence and (c) intended use.

Question 15.4 A method for the determination of riboflavin in cereal using fluorescence is described on page 103 of the course textbook. Given that the statistic $S_{y/x}$ is 0.4986 estimate the limit of detection for this analysis

Riboflavin (μg cm^{-3})	Fluorescence (Fl)
0.000	0.0
0.100	5.8
0.200	12.2
0.400	22.3
0.800	43.3

Outline answer to question 15.1

SOP's are dealt with on slides 13 - 16

Before writing the SOP, decide what grade of NaCl is required.
Calculate how much sodium chloride is required to prepare 100 cm^3 of a 0.1M NaCl solution.
The molecular weight of NaCl is X, therefore Y g is required.

Your SOP should include the following detail:

- Weigh accurately Y g of NaCl into a 100 cm^3 beaker. Record the data in your lab notebook.
- Dissolve the solid in 20 cm^3 of distilled water by stirring.
- Quantitatively transfer the solution into a 100 cm^3 class 'A' volumetric flask. Rinsing the beaker, and stirring rod carefully with distilled water and transferring the rinse water to the volumetric flask.
- Make up to the mark with distilled water and seal the flask. Mix the contents thoroughly by inversion.
- The solution may now be transferred into a storage bottle and labelled 0.1 M NaCl include the date, and any other information required, such as operators initials etc.

Outline answer to question 15.2

The VAM principles can be found on slide 17

1. What grade of NaCl is required? How accurate do you need it to be?
2. Is your balance calibrated? Have you selected appropriate grade glassware?
3. Are you competent? Have you been trained in the manipulation of glassware? Do you know how to execute a quantitative transfer? Have you been trained how to use the balance? Is your training up to date?
4. Are you working in an accredited lab, and have you followed procedures. If not have you written out your method clearly in your lab notebook (e.g. an SOP), have you recorded the weights etc, so you can refer to them later if you are queried or encounter a problem with the solution?
5. Where did you get your NaCl? Is it from recognised supplier, with a certificate of analysis?
6. See **4** above.

Outline answer to question 15.3

ISO definition of Method Validation:

The **confirmation** by examination and provision of **objective evidence**, that the particular requirements for a specific **intended use** are fulfilled'.

(a) Have you fulfilled the requirements? Comparison with standards, or reference materials. E.g. comparing the pH of a solution with the pH of known pH standards using a particular pH meter.
(b) Correctly analyisng data and applying the method performance tests. Is a calibration curve linear? Or is the pH measurement precise i.e. do repeated measurements of the pH give an acceptable standard deviation.
(c) Intended use. Preparing a primary standard for a titration requires more accuracy than preparing a 3% saline solution for an eyewash bottle.

Outline answer to question 15.4

Slides 62 - 67 will help with answering this question

1. Using the linear regression program on your calculator work out the equation for the calibration curve.

   Y = 53.75 X + 0.59

2. Use the formula: LOD = intercept + 3 ($S_{y/x}$) to work out the signal for the minimum concentration that can be detected with some degree of confidence.

   0.59 + (3 x 0.499) = 2.087 Fl

3. Then use the equation of the line above to convert the signal response to concentration, quote the limit of detection in units of concentration.

   0.03 μg cm^{-3}