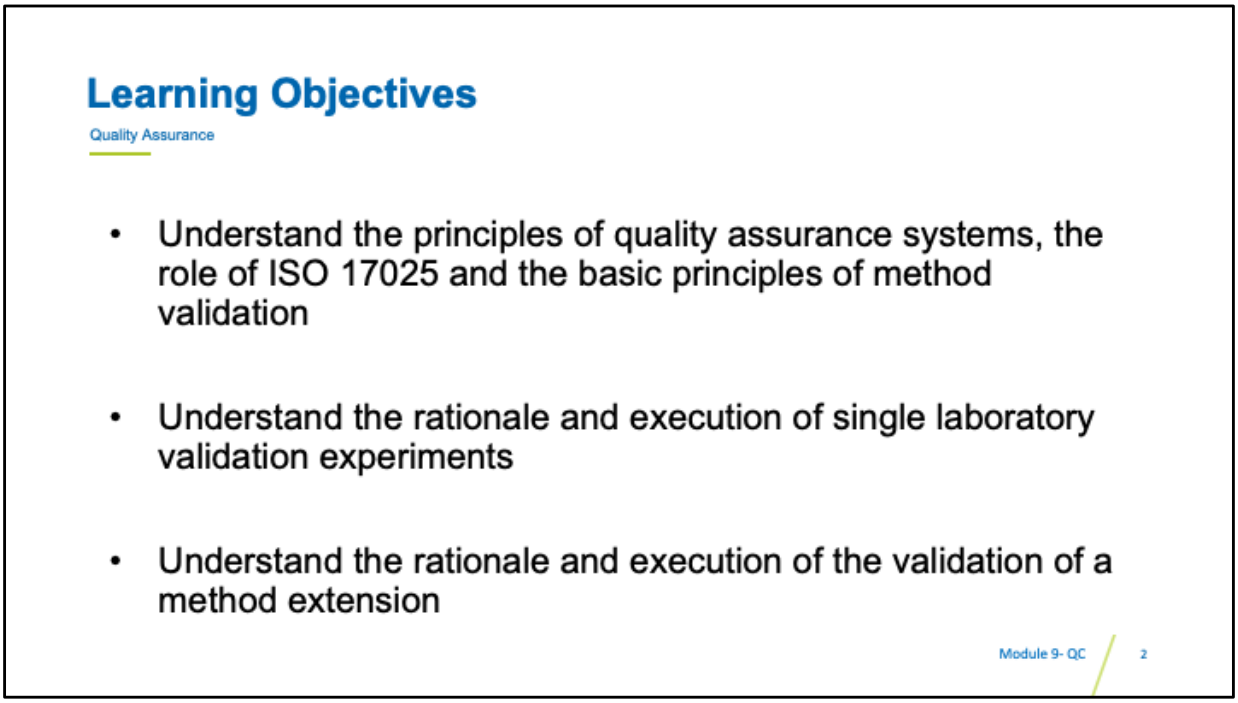
Welcome to Module 9 dedicated to understanding quality assurance



Analytical laboratories, especially those laboratories performing regulatory testing, must maintain a strong quality management system in order to ensure that they provide the correct test results for every sample.

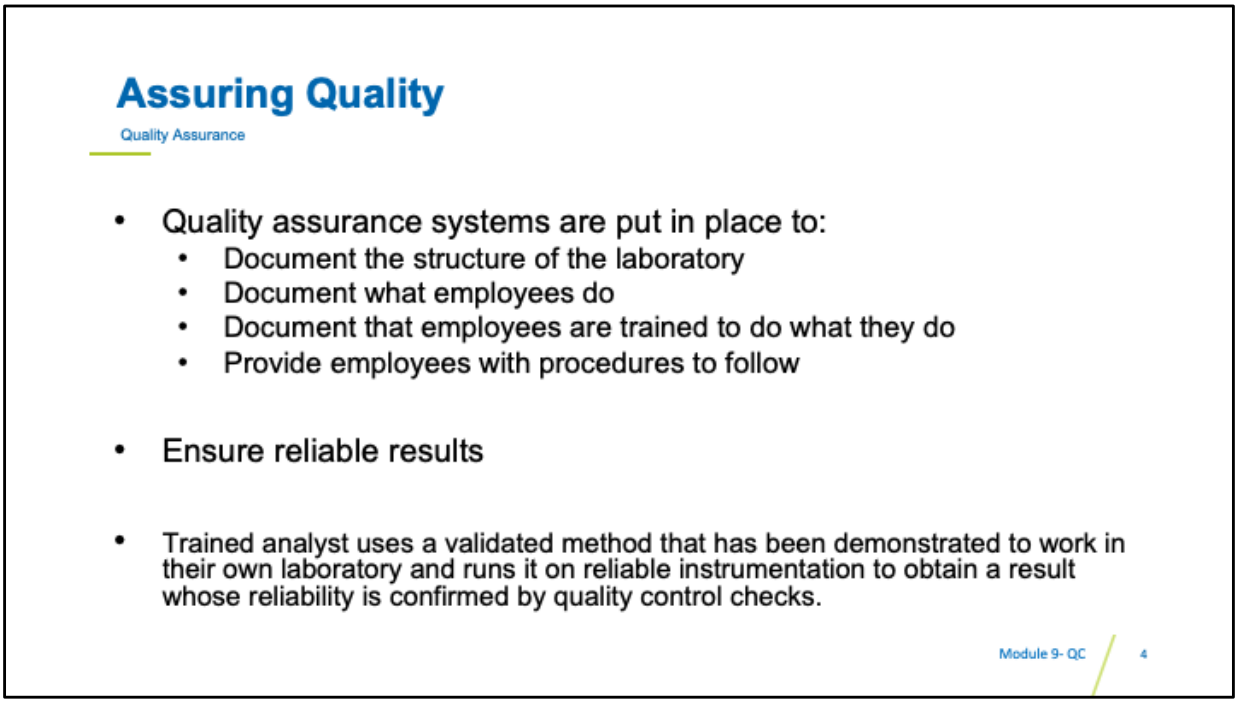
There are three learning objectives in this module.

First, we need to understand the principles of quality assurance systems, the role of ISO 17025 and the basic principles of method validation. We review these topic in section 1

In section 2, we focus on understanding the rationale and execution of single laboratory validation.

Finally, section 3 aims to further our understanding of the rationale and execution of the validation of method extensions.

Section 1: Introduction to quality assurance



Quality assurance system are put in place in laboratories to ensure that there is a clear structure in the organization which promotes confidence in the results. While documentation is often perceived as the objective of the quality assurance system, the focus should be placed on the actions that ensure quality.

For example, the organizational structure of a laboratory should illustrate how the skills and experience, as well as the checks and balances that it represents, will ensure reliable results.

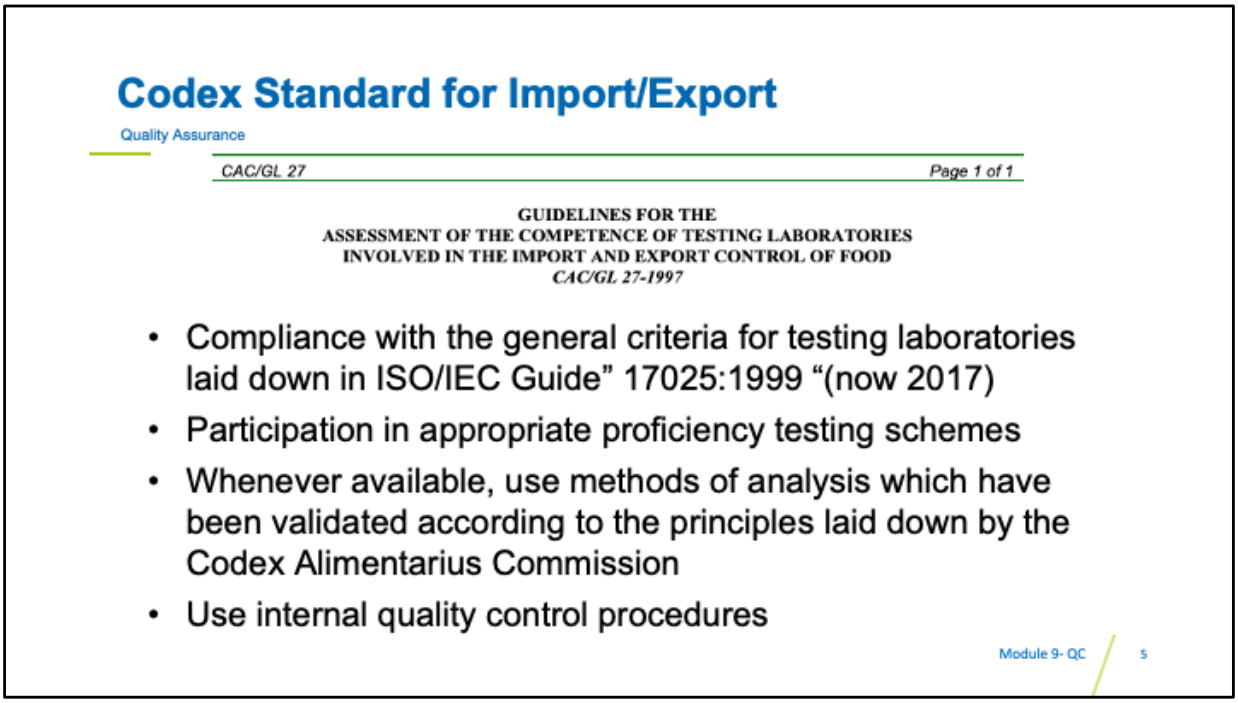
Similarly, the quality assurance system requires the documentation of employees' skills and experience; it is not the CV itself that matters, it is the fact that it shows that the employee is qualified for their job.

Training is also documented because it supports this demonstration that employees are qualified, but also the commitment of management to sustain high-quality results even as technology or best practices evolve.

Standard operating procedures (SOP) are an obvious part of the quality assurance system because they provide employees with procedures to follow. These SOPs are tested, and they produce reliable results.

In summary a trained analyst uses a validated method that has been demonstrated to work in their own laboratory and runs it on reliable

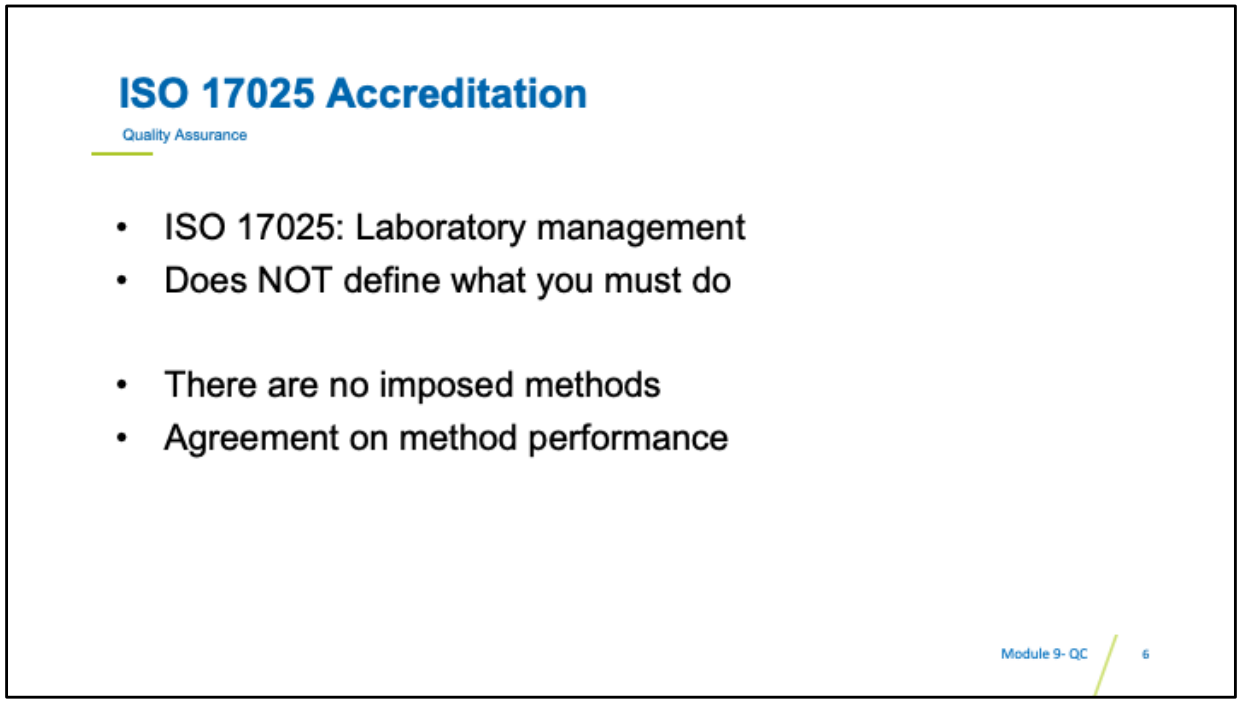
instrumentation to obtain a result whose reliability is confirmed by quality control checks.



The quality assurance systems for food testing laboratories involved in trade increasingly follow the guidelines from ISO/IEC guide 17025. ISO stands for International Standards Organization.

The ISO accreditation system for laboratories started in 1999 but was not broadly implemented in food safety laboratories until the late 2000s. Until then, rigorous internal quality assurance systems were considered sufficient to assure the reliability of the results. However, with the dramatic increase in international trade, deferring to international standards provides the advantage of building trust in each other's laboratories without a need to inspect them ourselves.

Compliance with ISO accreditation requirements includes participation in appropriate proficiency testing schemes. It also emphasizes the use of validated methods and internal quality control that ensure that the application of the validated method produces a reliable result. The latest revision of the standard was published in 2017.



It is important to understand that ISO 17025 accreditation is about laboratory management. It does not in itself define what a laboratory must do. What it does is require documentation for what a laboratory decides to do. For example, different methods could be used to measure contaminants, and the laboratory is free to select the method that is best suited for its instrumentation, environmental controls, staff training, experience, and its sample throughput.

Once this decision is made, the laboratory must document that analysts obtain the correct result reliably when they use this method. This is what we call a single laboratory validation, or verification, for the use of an official method. In the same spirit regulatory agencies typically do not require laboratories providing them with results to use predefined methods, but rather select methods that meet certain performance criteria.

ISO 17025 Accreditation

Quality Assurance

- Requires audits
- Is obtained for each method/commodity or for broad scopes
  - It is not an "accredited laboratory", it is a laboratory accredited for a list of methods

Complying with ISO 17025

A practical guidebook

UNIDO UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION

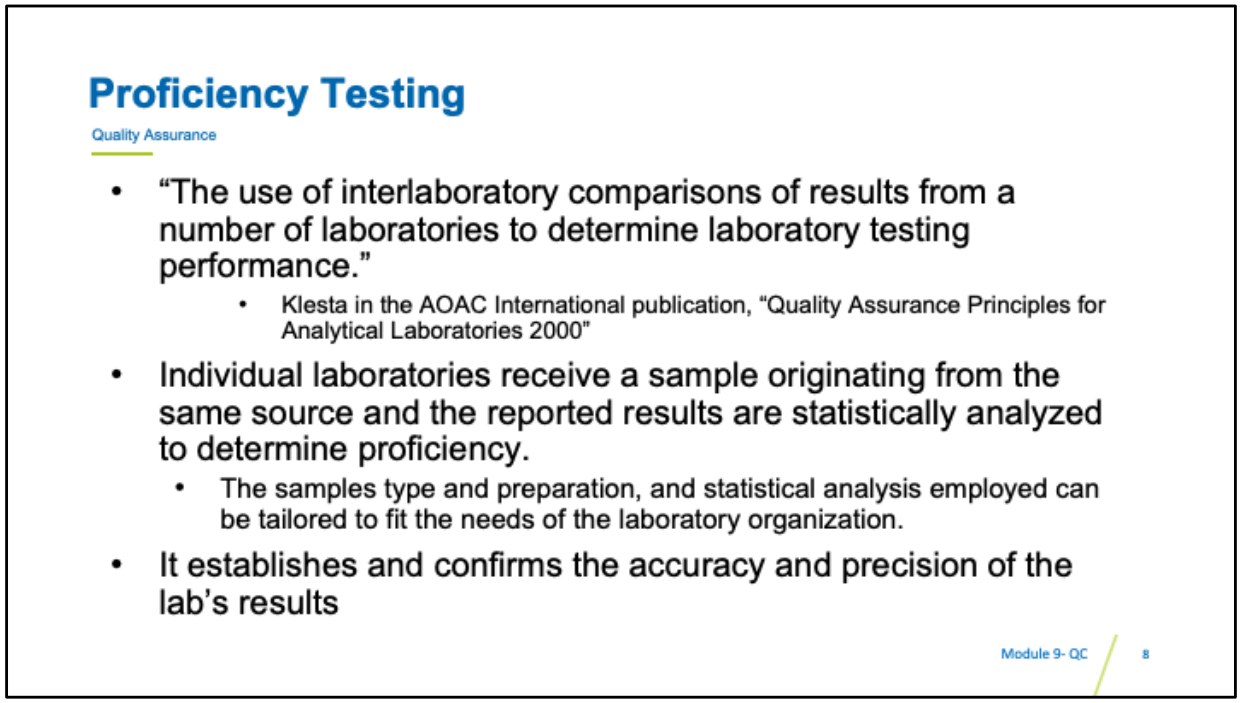
https://www.unido.org/sites/default/files/2010-08/Complying_with_ISO_17025_A_practical_guidebook_0.pdf

Module 9- QC 7

Put simply, ISO accreditation is a verification by a third-party that the laboratory has documented its procedures and quality assurance system. This third party performs audits to determine the suitability of the documentation and procedures to ensure that the laboratory obtains reliable results.

Contrary to popular belief, laboratory accreditation is not a one-time time deal covering everything that happens in the lab. In reality, each method is accredited, including its scope of application. The accreditation must also be maintained.

We recommend reading the UNIDO Practical Guidebook for Complying with ISO 17025 available at this address if you are not familiar with laboratory accreditation.



PT, as it is usually called, is a comparison of results obtained by a large number of laboratories from the same sample to determine laboratory testing performance. In a formal PT program, a provider who is accredited to deliver such services prepares a large quantity of a single sample and obtains a consensus value, or the assigned value. Then, they send portions of this sample to many laboratories, like dozens , and collect the results.

The PT establishes and confirms the accuracy and precision of a laboratory's results.

Results of PT

Quality Assurance

- Z Score: Comparison of all participating labs
  - Need many laboratories to be valid

$$z = \frac{(x - x_a)}{\sigma_p}$$

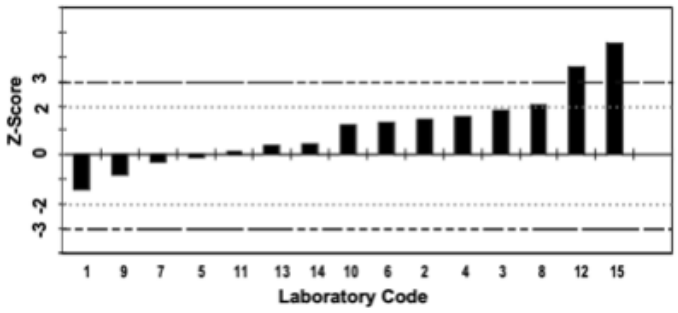
where x = the result reported by the participant
x_a = the assigned value
and σ_p = the standard deviation for proficiency

Using statistical analysis of the results, the PT provider publishes the z-scores for all the laboratories who participated. The names of labs are always kept confidential.

The z-score combines an estimate of the error of a result and the standard deviation. In other words, both the trueness and precision are assessed. Because this is a comparison of participants, there needs to be a fairly large number of laboratories in the round for the z-score to be useful.

Individual Z-Scores in PT

Quality Assurance



Module 9- QC 10

The z-scores are reported for each laboratory identified by a code to keep its identity confidential. A z-score of 0 indicates that the laboratory has obtained the value closest to the assigned value. While increasingly high z-scores both in the positive and negative directions indicate greater deviations from the assigned value, a z-score of up to 2 or minus 2 indicates a satisfactory result whereas a z-score greater than +2 or lower than -2 is considered unsatisfactory. A result greater than 3 indicates that action should be taken to correct the causes of the inaccurate result.

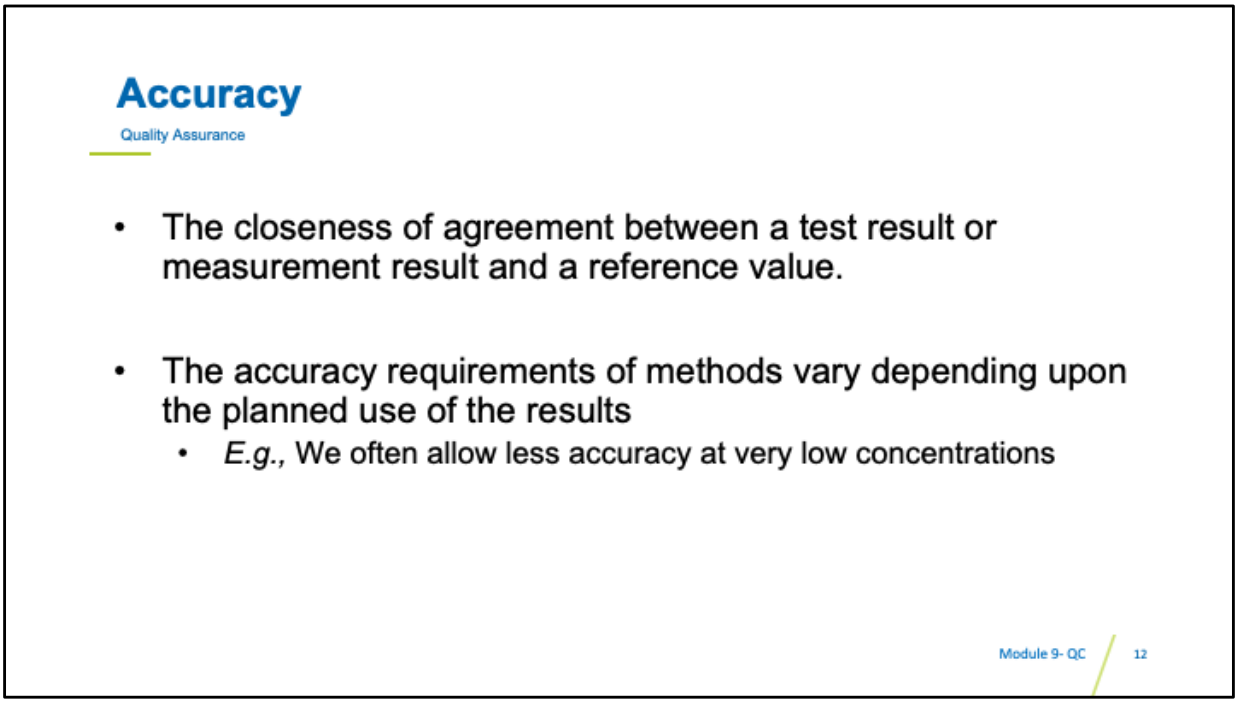
For anyone who is new to PT, it is important to understand that even if all the labs were excellent, some of them would get a score outside the satisfactory range because this is just the characterization of a distribution... We all hope to score 0 or as close to 0 as possible, but it really makes no difference as long as the laboratory obtains a z-score between -2 and and +2.

Method Validation (Codex)

Quality Assurance

- Accuracy
- Precision
- Selectivity
- Limit of quantitation (LOQ)
- Sensitivity
- Applicability
- Practicability

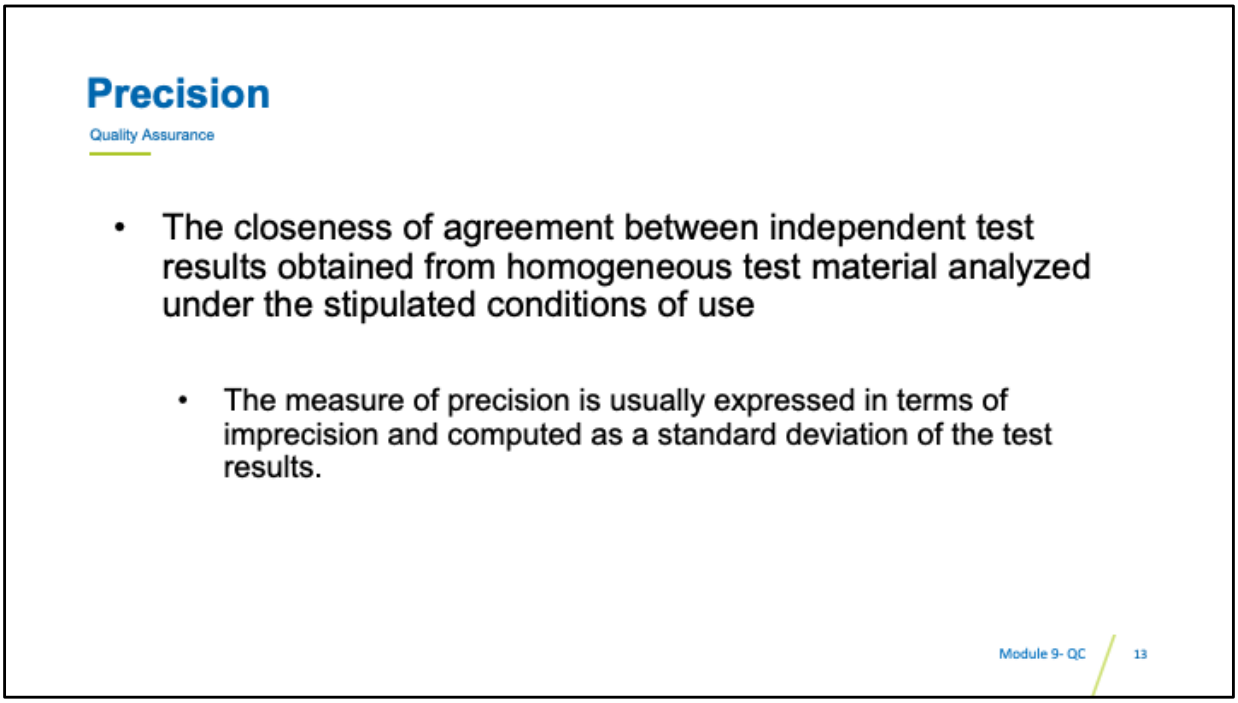
Each method is characterized by number of parameters that essentially define how good it is at measuring the compounds of interest from a matrix. There are more characteristics than shown in this list, such as bias, linearity and limit of detection for example, but we will limit our discussion to this list for the online section of this training.



Accuracy is the closeness of agreement between a test result or measurement result and a reference value. In other words, does this method produce the right result?

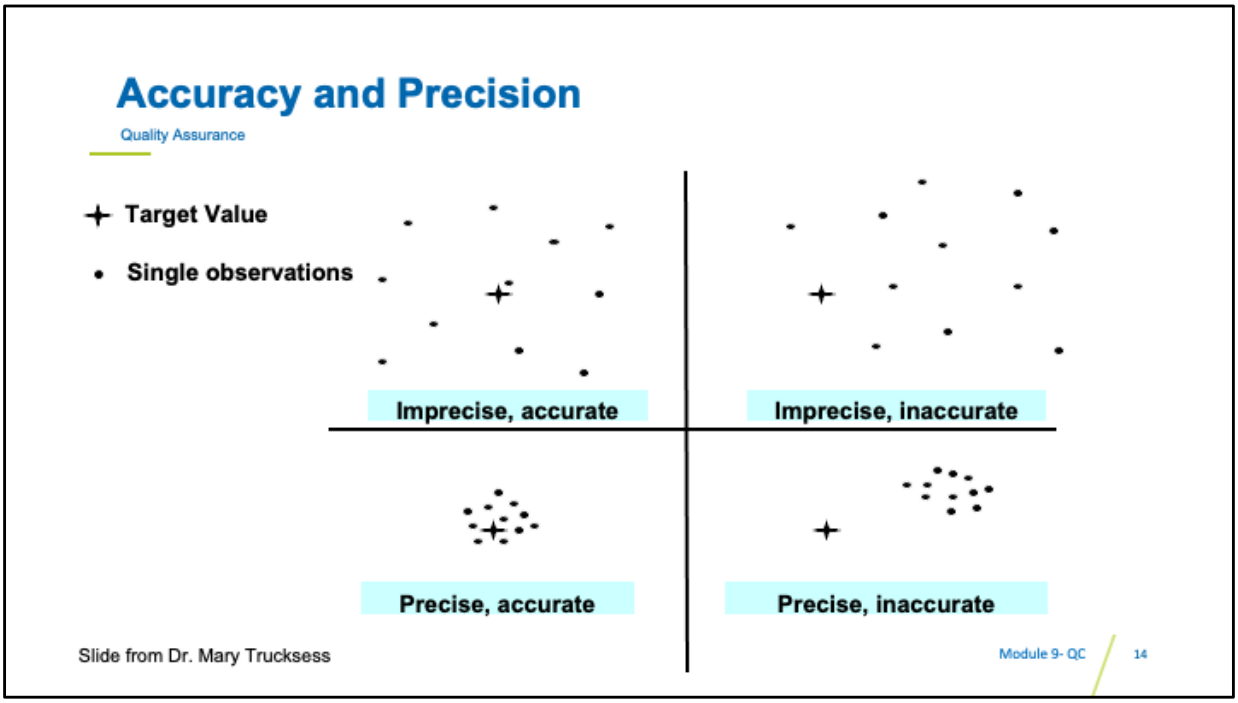
The accuracy requirements of methods vary depending upon the planned use of the results

E.g., We often allow less accuracy at very low concentrations. We will see examples of this later.



The precision of a method is the closeness of agreement between independent test results obtained from homogeneous test material analyzed under the stipulated conditions of use

The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results.



This graphical representation used by my friend Dr. Mary Trucksess is an excellent way to see what these two terms mean and how they differ.

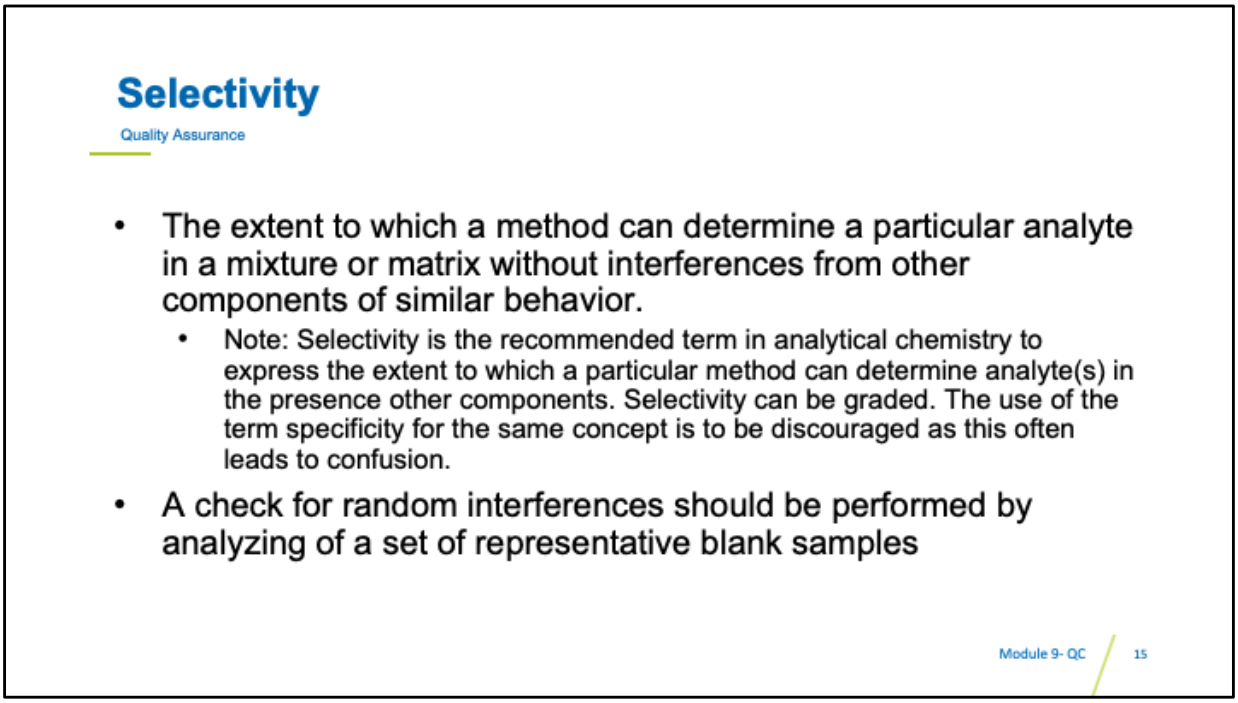
The target value is represented by a star and a number of single measurements are represented by dots. If we start from the top left corner, the method is accurate, because the mean is close the the target value, but it is imprecise because the individual measurements are a bit all over the place. It is correct on average, but if were to take only 3 measurements for example, the result could be way off... So the method is considered accurate, but imprecise.

The bottom right corner represents the opposite, where the method is precise because the points are all together, the individual measurements are close together, but it is inaccurate because the mean value is far away from the target value.

On the top right, the method is both imprecise and inaccurate, which means that the individual measurements are all over the place and don't even come close to the target value on average.

And of course, in the last corner, we have the precise and accurate measurement, where all the individual measurements are close together, and the average

close to the target value.

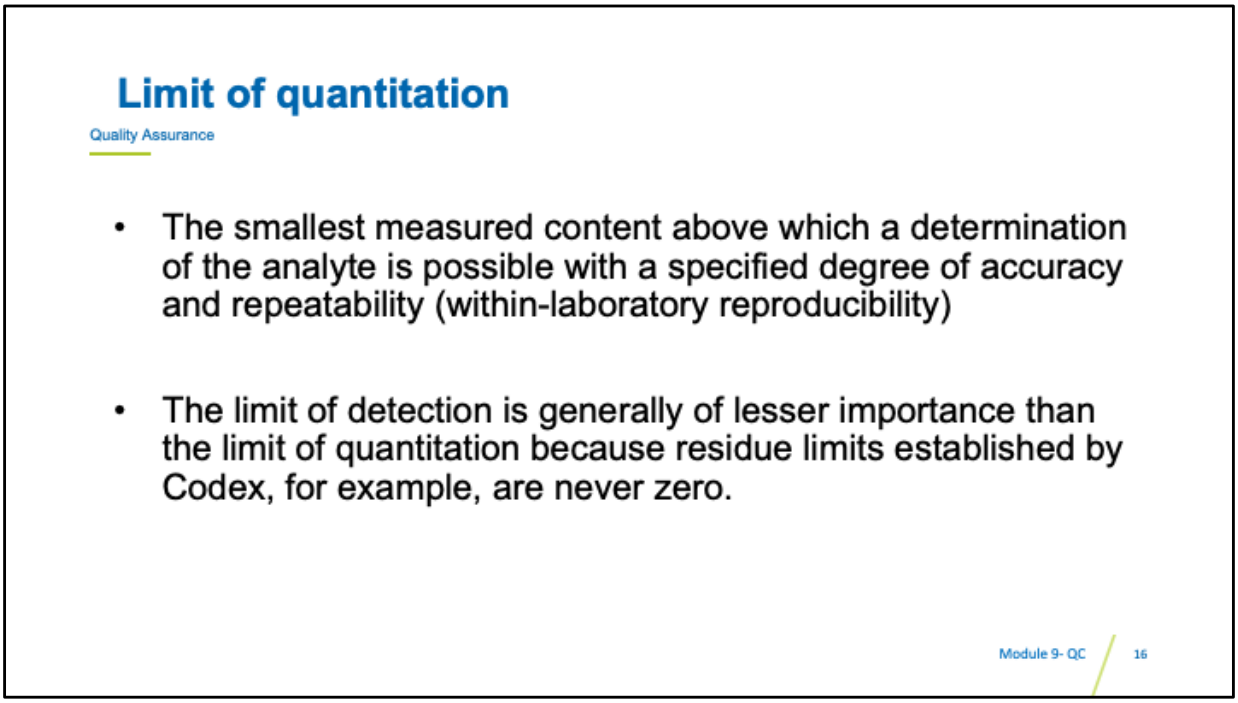


Selectivity: Selectivity is the extent to which a method can determine particular analyte(s) in a mixture(s) or a matrix without interferences from other components of similar behavior.

Note: The selectivity is the recommended term in analytical chemistry to express the extent to which a particular method can determine analyte(s) in the presence other components. Selectivity can be graded. The use of the term specificity for the same concept is common but should be discouraged as this often leads to confusion.

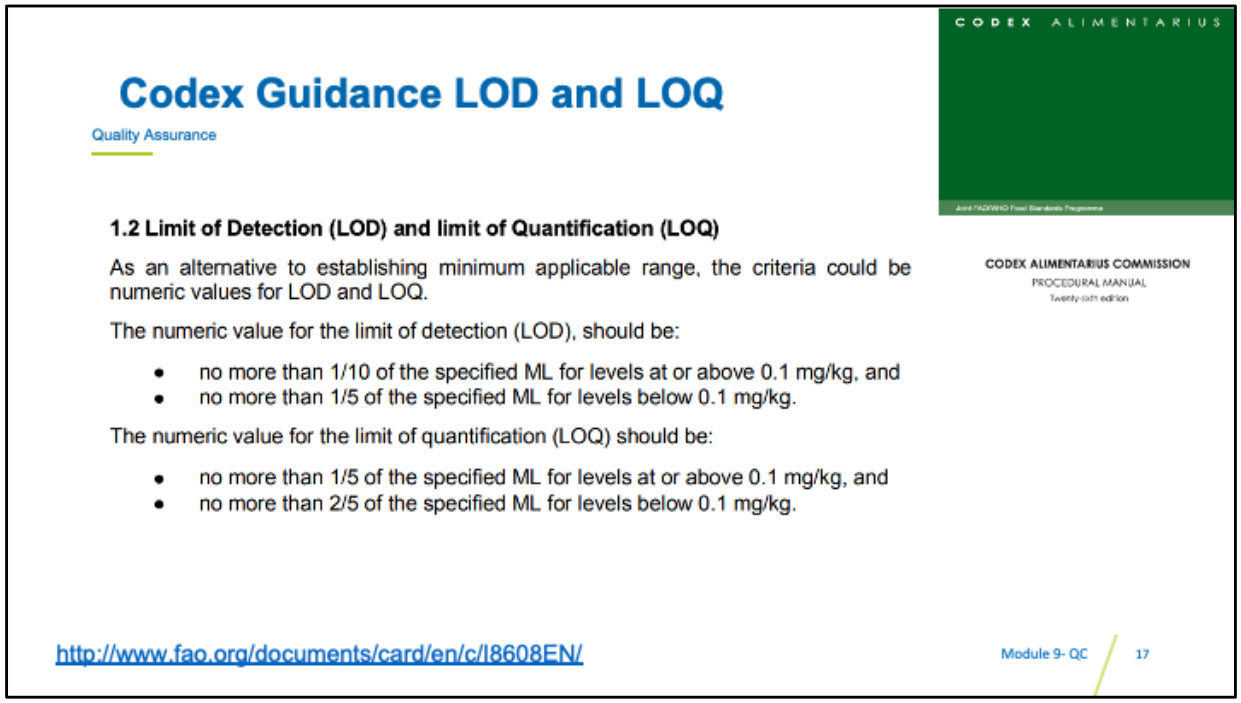
A check for random interferences should be

performed by analyzing a set of representative blank samples



The limit of quantitation is always very low in contaminants analysis. It is the smallest measured content above which a determination of the analyte is possible with a specified degree of accuracy and repeatability (or within-laboratory reproducibility). In other words, how small a concentration can we measure with this method?

The limit of detection is generally of lesser importance than the limit of quantitation because residue limits established by Codex, for example, are never zero.



Codex offers some guidance about desired levels for the LOQ and LOD. What we are trying to achieve is go low enough for our purpose, but not make the method excessively difficult just to go lower than we really need to. So lower is not always better.

Generally speaking, we are looking for an LOQ that is about 1/5th of the MRL when it's above 100 ppb, and 2/5th of the MRL below 100 ppb. So if you are trying to measure a contaminant with an MRL at 1 ppm, the LOQ should be 200 ppb, which us one fifth. If the MRL is 10 ppb, then the LOQ should be 4 ppb, which is two fifths.

The only methods that need extremely low LOQs are those used in risk assessment projects, where we want to be able to measure very low concentration to calculate exposure. These special methods are generally not appropriate for regulatory testing as they would impose complications that are only there to achieve the very low LOQs and serve no purpose in the MRL range. They are fit for the purpose of risk assessment.

Codex Method Criteria

Quality Assurance

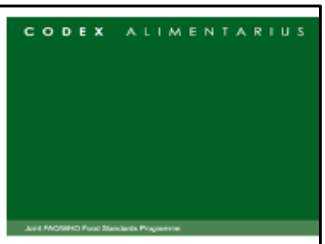
2. Method criteria at different MLs (maximum level, minimum level, normative level or concentration range)

In table 4 examples on method criteria are given for certain MLs.

Table 4: Method criteria for MLs at increasing orders of magnitude.

ML unit	0.001 mg/kg	0.01 mg/kg	0.1 mg/kg	1 mg/kg	10 mg/kg	100 mg/kg	1 g/kg	10 g/kg
Concentration ratio of ML (C_{ML})	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
Minimum applicable Range	From 0.0006 to 0.0014 (mg/kg)	From 0.006 to 0.014 (mg/kg)	From 0.03 to 0.17 (mg/kg)	From 0.52 to 1.48 (mg/kg)	From 6.6 to 13.3 (mg/kg)	From 76 to 124 (mg/kg)	From 0.83 to 1.2 (g/kg)	From 8.8 to 11 (g/kg)
LOD (≤ mg/kg)	0.0002	0.002	0.01	0.1	1	10	100	1000
LOQ (≤ mg/kg)	0.0004	0.004	0.02	0.2	2	20	200	2000
RSD_R (≤ %)	44	44	44	32	22	16	12	8
Recovery (%) *	40 - 120	60 - 115	80 - 110	80-110	80 - 110	90 - 107	95 - 105	97 - 103

* Other guidelines are available for expected recovery ranges in specific areas of analysis. In cases where recoveries have been shown to be a function of the matrix other specified requirements may be applied.



CODEX ALIMENTARIUS COMMISSION

PROCEDURAL MANUAL

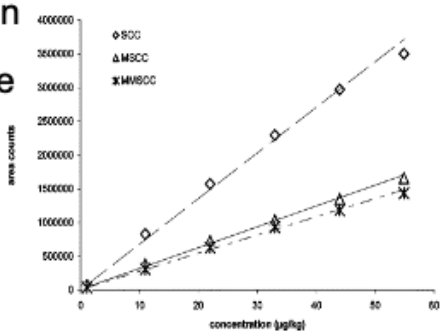
Twenty-sixth edition

I briefly mentioned before that method requirements change depending on the concentration range of interest. We just saw that LOQ is adjusted at very low concentrations. It is the same for the standard deviation and the recovery requirements. As we try to measure smaller and smaller concentrations , extraction is more difficult, which means recovery may not be proportionally as good, and we usually see more relative variability at these low levels and therefore allow more in the requirements.

Sensitivity

Quality Assurance

- The change in the analytical response divided by the corresponding change in the concentration of a standard (calibration) curve, *i.e.*, the slope of the analytical calibration curve.
- In the ideal situation, the calibration curve is a straight line, expressing a direct linear relationship between analytical response and standard concentration



Module 9- QC 19

Sensitivity is the change in the analytical response corresponding change in the concentration of the standard for the calibration curve. It therefore corresponds to the slope of the analytical calibration curve

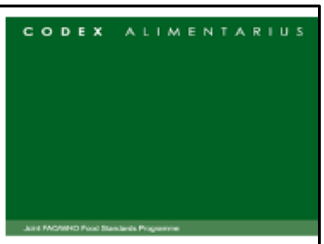
We look for linearity in method, a direct relation between the change in response and the change in actual concentration.

Codex Reference Criteria

Quality Assurance

Table 1: Guidelines for establishing numeric values for the criteria:

Applicability:	The method has to be applicable for the specified provision, specified commodity and the specified level(s) (maximum and/or minimum) (ML). The minimum applicable range of the method depends on the specified level (ML) to be assessed, and can either be expressed in terms of the reproducibility standard deviation (s_R) or in terms of LOD and LOQ.
Minimum applicable range:	For ML ≥ 0.1 mg/kg, [ML - 3 s_R, ML + 3 s_R] For ML < 0.1 mg/kg, [ML - 2 s_R, ML + 2 s_R] s_R [13] = standard deviation of reproducibility
Limit of Detection (LOD):	For ML ≥ 0.1 mg/kg, LOD ≤ ML · 1/10 For ML < 0.1 mg/kg, LOD ≤ ML · 1/5
Limit of Quantification (LOQ):	For ML ≥ 0.1 mg/kg, LOQ ≤ ML · 1/5 For ML < 0.1 mg/kg, LOQ ≤ ML · 2/5



CODEX ALIMENTARIUS COMMISSION

PROCEDURAL MANUAL

Twenty-sixth edition

Codex also recommends an applicable range for a method that fits the MRLs. In general, the method has to be applicable to the particular analyte(s)/provision(s) in the specified matrix/ commodity or food category. We aim for being able to measure at the MRL +- 3 times the standard deviation of reproducibility at minimum. Of course, the range is smaller at lower concentrations.,

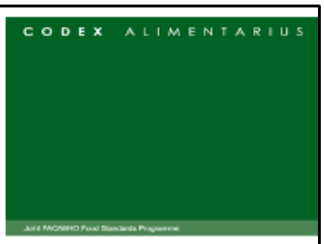
In other words, the method is applicable for concentrations levels around the specified Maximum Limit, or the MRL should be within the validated range.

Codex Reference Criteria

Quality Assurance

Precision:	For ML ≥ 0.1 mg/kg, HorRat value ≤ 2 For ML < 0.1 mg/kg, the RSD_{TR} < 22%. RSD_R [14] = relative standard deviation of reproducibility. $RSD_R \leq 2.\ PRSD_R$			
Recovery (R):	**Concentration**	**Ratio**	**Unit**	**Recovery (%)**
	100	1	100% (100g/100g)	98 – 102
	≥10	10^{-1}	≥ 10% (10g/100g)	98 – 102
	≥1	10^{-2}	≥ 1% (1g/100g)	97 – 103
	≥0.1	10^{-3}	≥ 0.1% (1mg/g)	95 – 105
	0.01	10^{-4}	100 mg/kg	90 – 107
	0.001	10^{-5}	10 mg/kg	80 – 110
	0.0001	10^{-6}	1 mg/kg	80 – 110
	0.00001	10^{-7}	100 µg/kg	80 – 110
	0.000001	10^{-8}	10 µg/kg	60 – 115
	0.0000001	10^{-9}	1 µg/kg	40 – 120
Trueness:	Other guidelines are available for expected recovery ranges in specific areas of analysis. In cases where recoveries have been shown to be a function of the matrix other specified requirements may be applied. For the evaluation of trueness preferably certified reference material should be used.			

The criteria in Table 1 must be approved for the determination in question.



CODEX ALIMENTARIUS COMMISSION

PROCEDURAL MANUAL

Twenty-sixth edition

Module 9- QC 21

The recovery is defined by Codex as the proportion of the amount of analyte, present naturally or added in the analytical portion of the test material. This means the proportion of the contaminant that we are able to extract from the matrix for measurement.

Acceptable recovery changes depending on the concentration range.

Example of Requirement (Mycotoxins FDA)

Quality Assurance

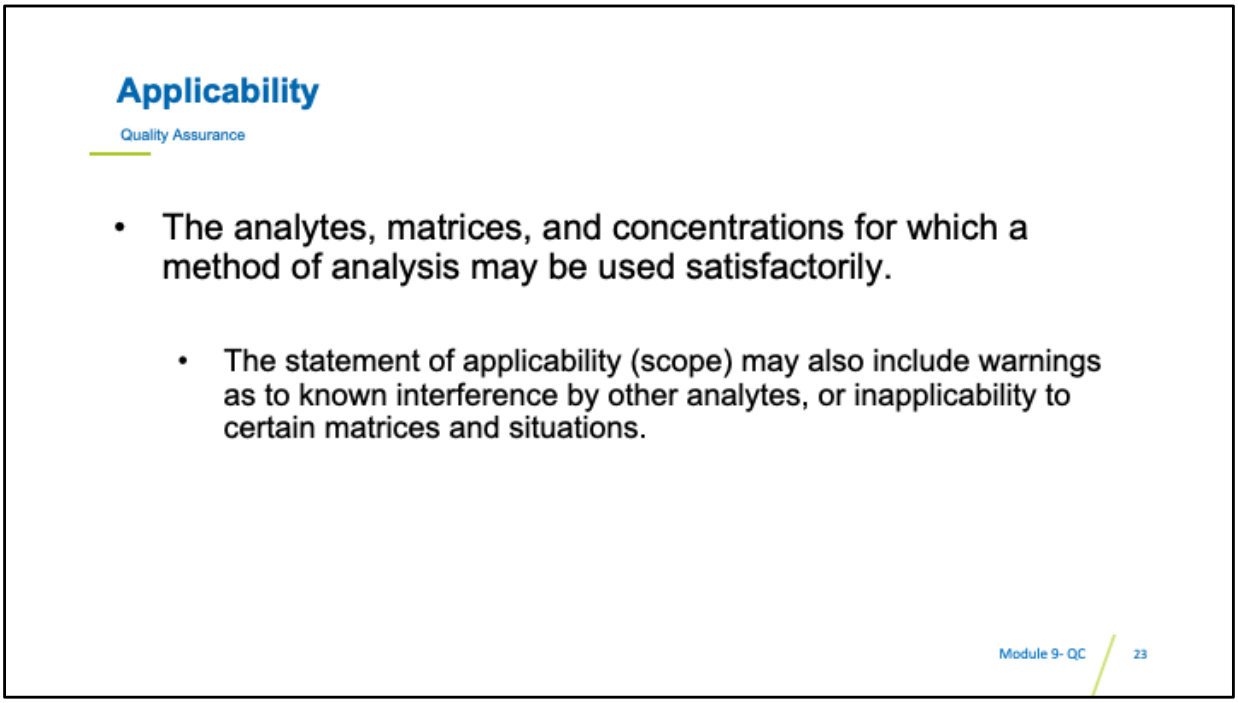
Concentration	Recovery limits %	Calculated PRSDr %*	Calculated PRSDR**
10 µg/g (ppm)	80 - 115	6	11
1 µg/g (ppm)	75 - 120	8	16
10 ng/g (ppb)	70 - 125	15	32
1 ng/g (ppb)	60 - 120	22	45
≤1 ng/g (ppb)	40 - 120		

* $PRSDr = C^{-0.15}$ (C = mass faction)
**$PRSDR = 2C^{-0.15}$
Acceptable values for repeatability (RSDr) are between ½ and 2 times the calculated value
Acceptable values for reproducability (RSDR) are between ½ and 2 times the calculated values
$HORRAT_R$ = RSDR (found, %)/$PRSD_R$ (calculated, %)

Slide from Dr. Mary Trucksess

Each country has the right to set its own criteria for the methods they use or want used for results reported to them. In fact, different agencies can have different requirements because their measurements are for different purposes.

This is a photo of my dear friend Dr. Mary Trucksess at a dinner we had in an Indian restaurant in Malaysia. This was dessert... Mary was the top chemist at the US FDA for mycotoxins until her retirement. She provided this table as an example of the US FDA requirements for methods to measure mycotoxins. These align with the Codex recommendations.



Applicability is the analytes, matrices, and concentrations for which a method of analysis may be used. In addition to a statement of the range of performance for each factor, the statement of applicability, also known as the scope, may also include warnings as to known interferences by other analytes, or matrices for which the method does not work.

Practicability and applicability under normal conditions

Quality Assurance

- Ease with which a method may be applied by those skilled in analysis
- Preference for methods applicable to a broad range of matrices and analytes
  - It also may include application to multi-residue methods.
- The method should be assessed over the relevant range of concentration, taking as a minimum half the value of the specified limit and twice the specified limit.

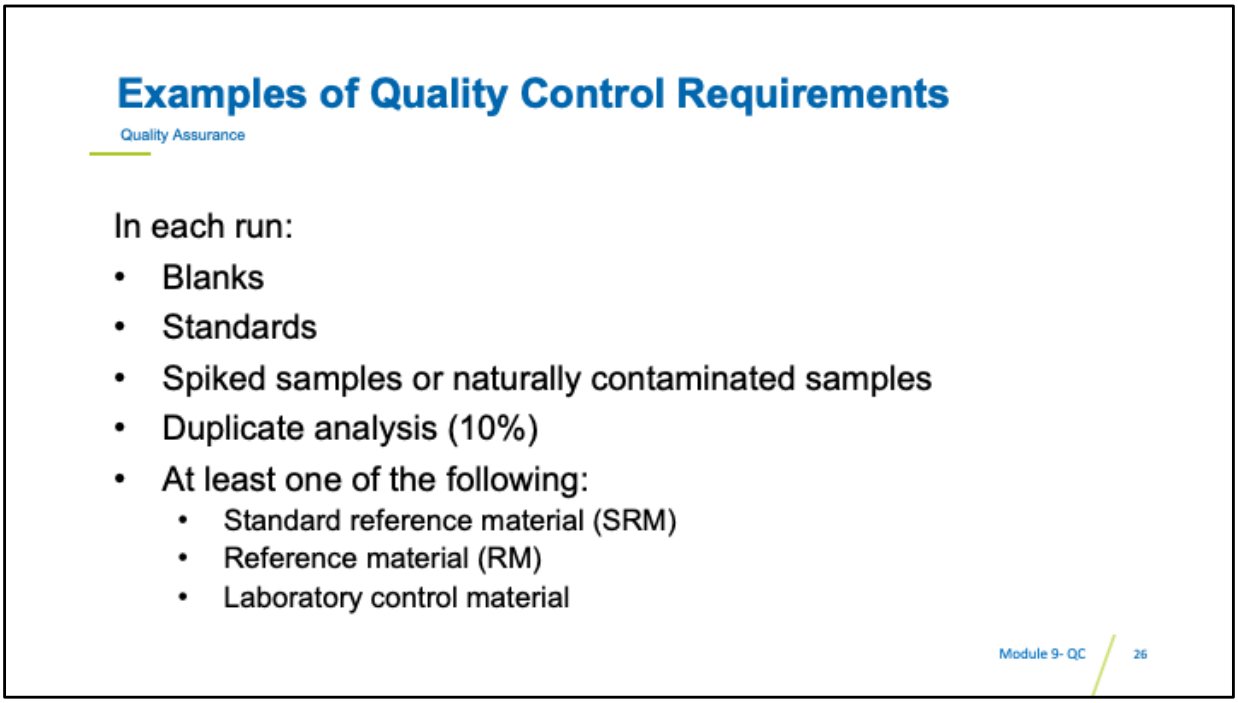
Practicability refers to the ease with which a method may be applied by those skilled in analysis.

Preference should be given to methods of analysis which are applicable to a broad range of matrices and analytes.

It also may include application to multi-residue methods.

The method should be assessed over a relevant range of concentration, taking as a minimum half of the value of the specified limit and twice the specified limit.

Obviously, there is a lot more to quality assurance than using appropriate methods and obtaining the correct value reliably... We cannot go through a thorough review of quality assurance programs, so I thought I would at least provide a reference to look at if you are interested. This is the repository of information about the quality control used in the regulatory laboratories of the US FDA.



Going back to very practical aspects of QA, we look at the steps we take to ensure that we are confident about our results.

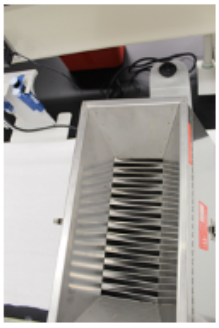
We will discuss each of these steps during our in-person training, as we apply them in our hands-on activities. Briefly, we will run blanks, so samples without any contaminant but with matrix, spiked standards in solvent and their counterparts spiked in blank matrix extracts to evaluate matrix effects, and we will run some of our analyses in duplicate to see that we get the same answer both times.

Then we will also use one or more of the three types of standards, either

Standard reference material (SRM), Reference material (RM) and laboratory control material.

QA Sample Preparation

Quality Assurance



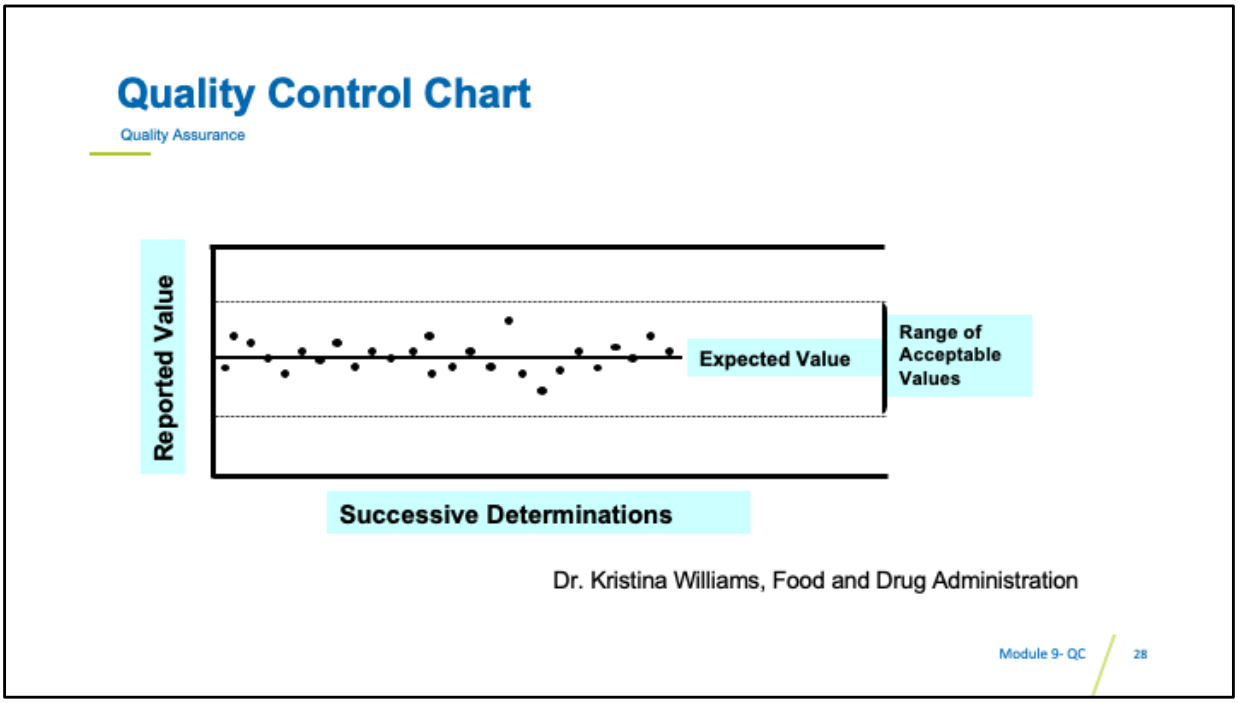
- Avoid errors!
  - Non-homogeneous
  - Large particles
  - Dirty blender, grinder or mixer
  - Poor sub sampling
  - "The usual human mistakes": Inaccurate test sample weight, pipetting, cross-contamination...

Module 9- QC 27

There are some well-known sources of error in the sample preparation. The biggest source of error is a non-homogeneous material. If the contaminant is not evenly distributed in the matrix, it is essential to homogenize it so that each small portion we take to perform the analysis, called the test portion, actually contains about the same amount.

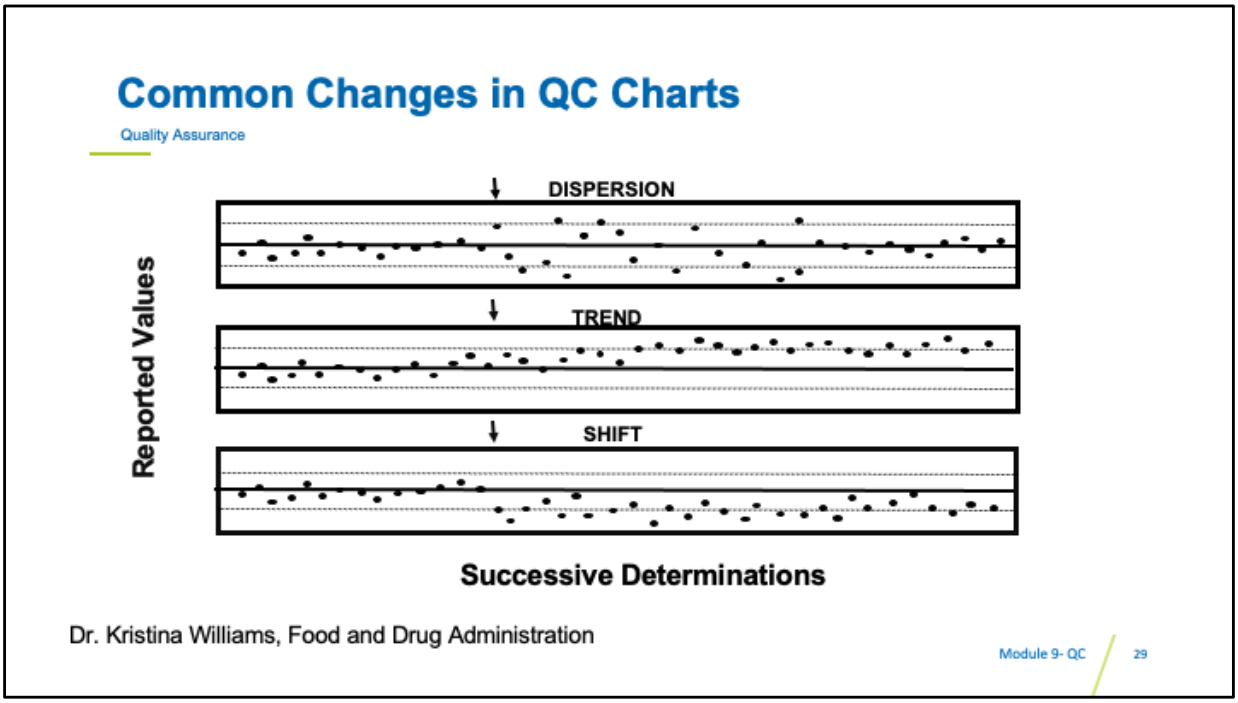
Large particles contribute to poor homogeneity, so grinding, crushing and blending are pretty common in our labs to reduce particle size. Of course, we can cause cross-contamination between samples if we don't clean our tools properly. Sub-sampling is typically a source of error when the sample is not homogeneous.

Finally, we have all the usual human errors that happen every once in a while. We need to have procedures in place to catch ourselves when we make these errors.



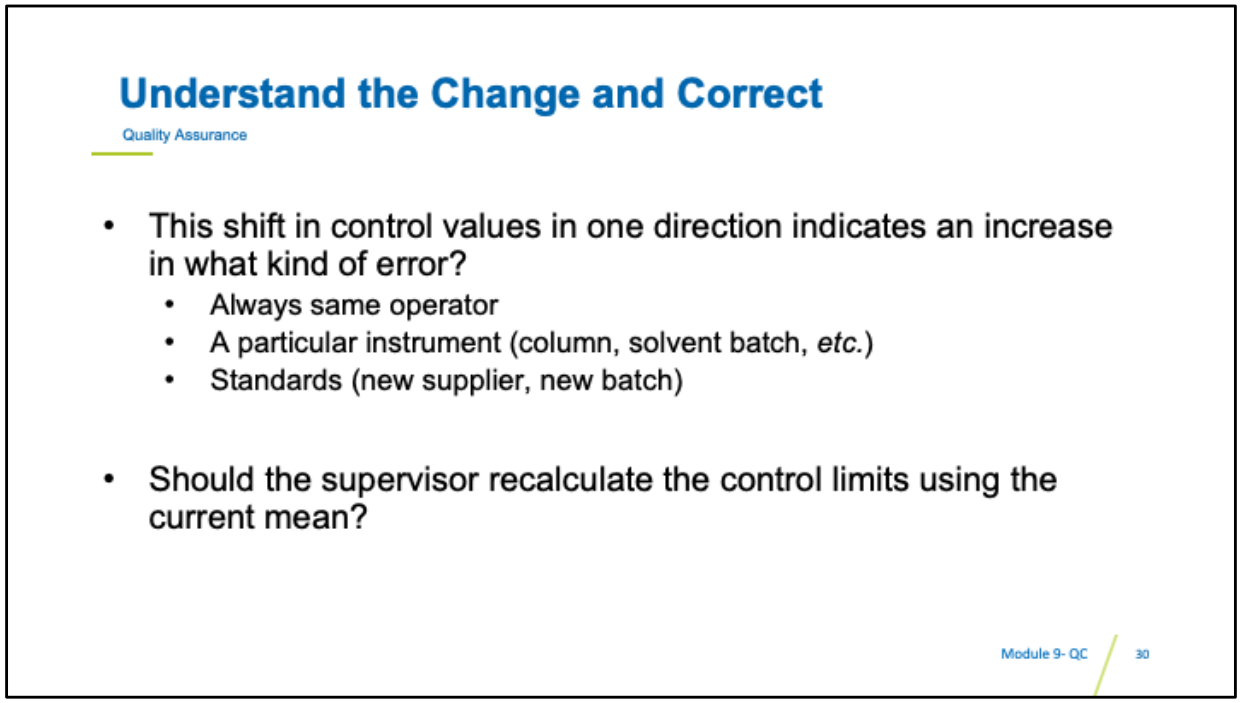
The quality control chart is a very important tool to let us know how things are going in our laboratory. These charts are pretty standard, but I am going to explain them the way I heard Dr. Kristina Williams from FDA explain. I thought she gave the clearest explanation I had ever heard, so I hope I can do it justice!

A QC chart is just a report of the values obtained from your QC sample (one of the three at the bottom of the list in the previous slide). As long as your values hover around the middle, or the expected value, and don't go beyond the two lines delimiting your range of acceptability, you are doing OK. However, there is a lot more to this control chart than staying between the lines.



You can see that the points are starting to get further from the middle where the arrow is pointing. At this point, measurements are still within the acceptability limits, but we can see that something is changing, and we call this change a dispersion. The second chart shows a trend upward that keeps growing. We call this one a trend, while the bottom one shows a trend downward, but then the measurements stay around a new mean... We call this a shift.

All of these are indicative of something changing in our lab and alerts that we need to do something about it. The beauty of the control chart is that you can start seeing the dispersion, the trend or the shift happen while the results are still within the acceptance limits, which means that you have an opportunity to correct or whatever caused it before your result go out of bounds. Once they are out of bounds, or outside of the acceptability range, you know for sure that they are not reliable and therefore you cannot use these results.



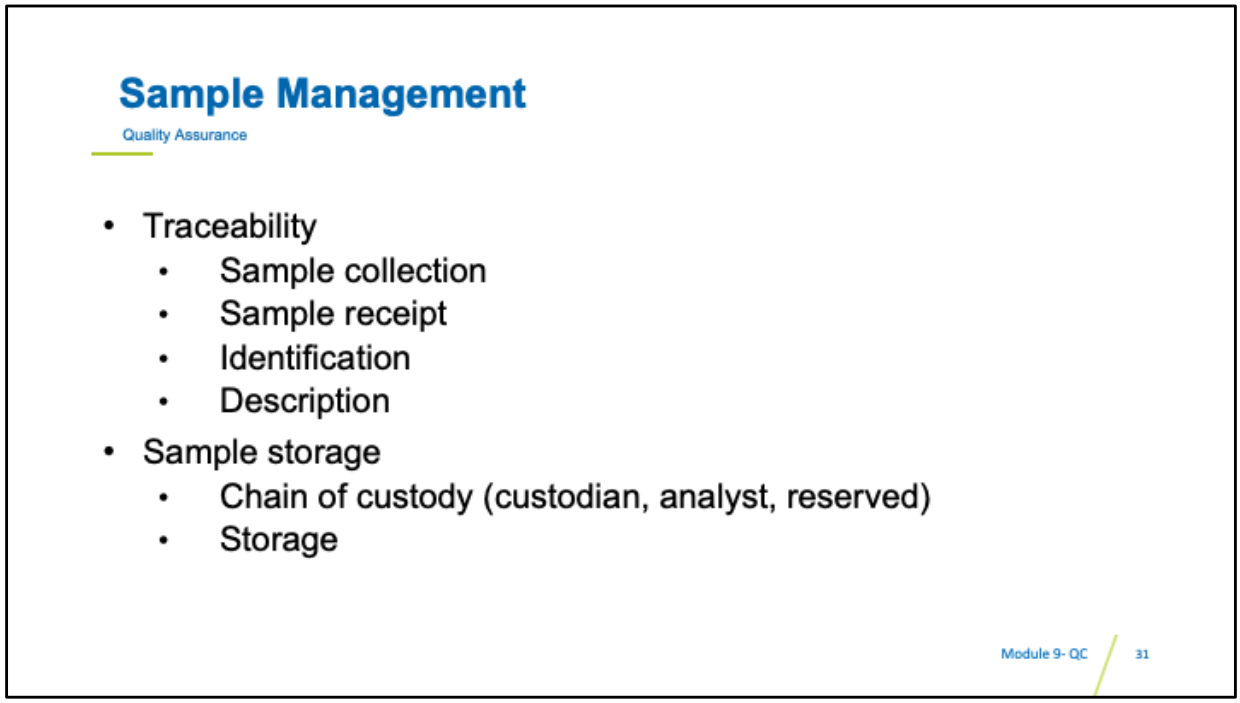
The purpose of the quality control chart is to spot any changes and implement the corrective actions necessary to continue to produce reliable results, ideally before the measured values have exceeded the acceptance limits.

The interpretation of the control chart should be done with care because many sources of error can cause in the same direction. For example, a shift in the reported values associated with specific work periods or shifts maybe due to the operator on duty, or to a change in the environmental conditions of the laboratory.

For example, a laboratory running mass spectrometry during the night must ensure that the air conditioning stays on because if the temperaturein the building changes significantly in the building, the results will be different.

For laboratories that have multiple instruments running the same methods, a trend from one instrument would indicate that it comes from that instrument itself; its column and any one of its components can be the source of error. A batch of solvent or reagents could cause a sudden trend, but this would be seen across all instruments. The same is true for a new batch of standards, whether they come from a new supplier or the usual supplier of the laboratory.

The control limits and the mean might change over time and the laboratory supervisor needs to decide if this is the new reality or if they need to continue to search for the source of the change. For example, as an instrument ages, it may cause a change in the results. We adapt to this very slowly, as the instrument ages. Then, the day we replace it, the results could be quite different! They could be like we used to see some years ago when our instrument was new. In this case, we could adjust the acceptance limits to fit our new reality.



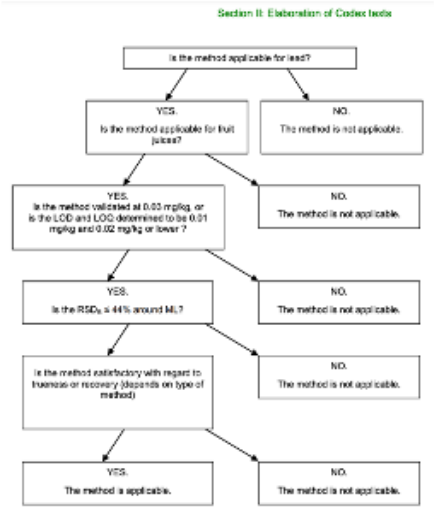
Sample management is also an important component of the quality assurance system. Traceability of the sample, all the way back to sample collection, is insured through proper documentation of the chain of custody. The sample label must always contain sufficient information for traceability before the sample was received in the laboratory, and typically also contains a description of the sample detailed enough to ensure that the label undoubtedly belongs to the sample

. Samples must be stored in appropriate conditions throughout the chain of custody, and this includes the preservation of a portion of the sample for re analysis in the case of disputed results. The sample identity must be tracked through the apportioning process.

Factors to consider when choosing a method

Quality Assurance

- Purpose (fit for purpose)
- Sample matrix
- Sample preparation
- Detection limit
- Dynamic range
- Potential interferences
- Ease of use
- Cost



As we saw in Module 8, we go through a series of questions when trying to determine if a method should be used. We try to decide if it is fit for our purpose and that includes whether is is applicable, and validated, for our matrix. We also consider if we can perform the sample preparation described, what the LOQ and LOD are, how high a concentration we can measure, what affects the instrument response –the interferences, etc. On a very practical note, we also need to decide if the method is easy to use and if we can afford all the reagents, standards and instruments for the volume of samples we need to analyze.

Criteria for Acceptance of LC, GC –MS Methods

Quality Assurance

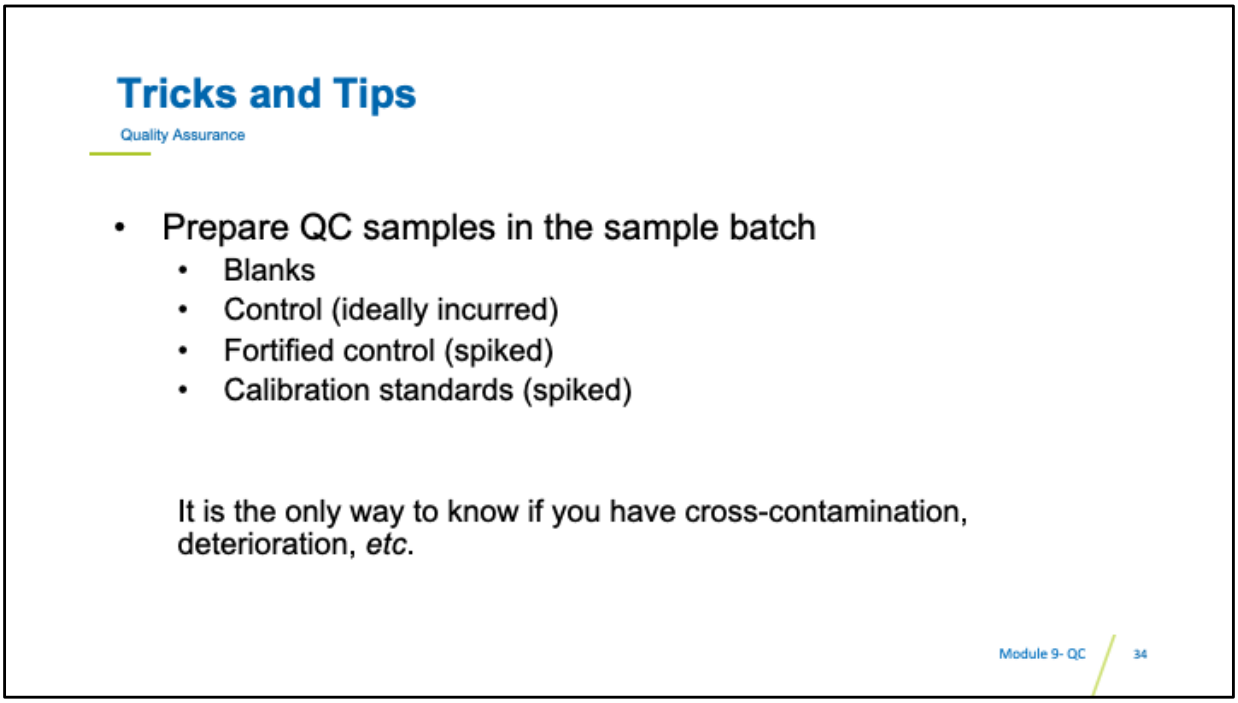
- Retention time (check applicable regulations)

  As compared to standard

  •GC/MS: ± 0.05 minutes

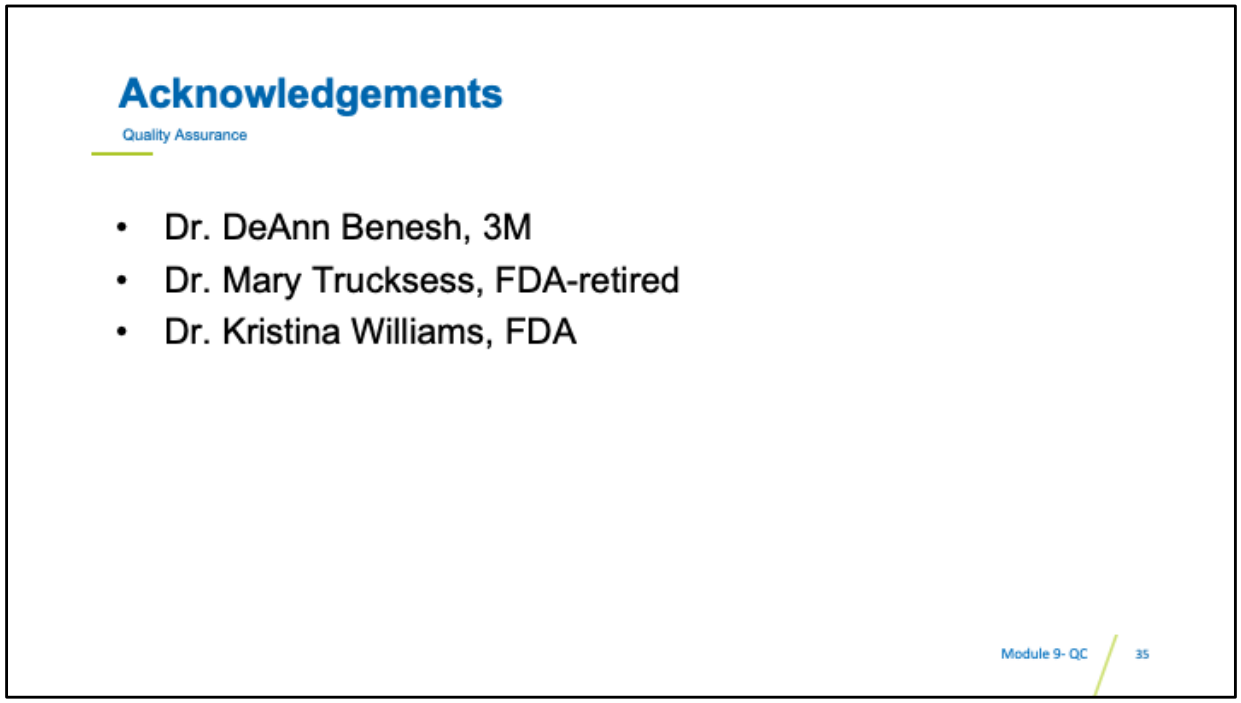
  •LC/MS: ± 5%

The question of instrumentation, or whether our instrument is good enough for a new method, sometimes requires us to verify that we can achieve the acceptability criteria. For example, the retention time of an analyte in the sample needs to be close to the retention time of the standard run on the same instrument and on the same day. If the instrument is not stable, the retention time may vary enough to change between samples... It could be the column that is a problem and if we can't afford a new one, we may be better off using another method...



One final tip to keep in mind: Always prepare your QC samples as part of the sample batch. A fairly typical order of preparation is the blanks, then incurred controls, followed by the spiked control and the calibration curve. Preparing QC samples in isolation would not reflect all the sources of errors that apply to the samples, so they would end up not serving their purpose.

-- Remember, the purpose of the QC steps is to prevent problems, spots small problems before they become big enough to cause unreliable results, and in the worst-case scenario, they help us find what caused the failure and fix it.



As was the case for the other modules, I received a lot of help when preparing the training slides for this section. Special thanks go to Dr. DeAnn Benesh of 3M, Dr. Mary Trucksess who is now retired from the FDA and Dr. Kristina Williams of the FDA

You have reached the end of section 1. In the next section, we discuss method validation.