Validation of liquid chromatography mass spectrometry (LC-MS) methods (analytical chemistry) course

This is a practice-oriented on-line course on validation of analytical methods, specifically using LC-MS as technique.

Table of contents

Course introduction

Course introduction

http://www.uttv.ee/naita?id=23245

https://www.youtube.com/watch?v=jbdA8PnPdLY

Short description of the course

This course – *LC-MS Method Validation* – is a practice-oriented on-line course on validation of analytical methods, specifically using LC-MS as technique. The course introduces the main concepts and mathematical apparatus of validation, covers the most important method performance parameters and ways of estimating them. The course is largely based on the recently published two-part tutorial review:

- Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part I. A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio, I. Leito. *Anal. Chim. Acta* **2015**, *870*, 29-44
- Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part II. A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio, I. Leito. *Anal. Chim. Acta* **2015**, *870*, 8-28

The course contains lectures, practical exercises and numerous tests for self-testing. In spite of being introductory, the course intends to offer sufficient knowledge and skills for carrying out validation for most of the common LC-MS analyses in routine laboratory environment. The real-life analysis situations for which there are either examples or self-tests are for example pesticide analyses in fruits and vegetables, perfluoroalkyl acids in water, antibiotics in blood serum, glyphosate and AMPA in surface water, etc. It is important to stress, that for successful validation experience (both in analytical chemistry as such and also specifically in validation) is crucial and this can be acquired only through practice.

Required preliminary knowledge

Introductory level knowledge of analytical chemistry, as well as liquid chromatography and mass spectrometry is required. More advanced knowledge of analytical chemistry and introductory knowledge of mathematical statistics is an advantage.

Study outcomes

The participant who has successfully passed the course knows:

- the main performance parameters of analytical methods, what they show and which of them are particularly important in different situations;
- the main mathematical concepts and tools in method validation;
- the main approaches for evaluation of the performance parameters in the case of LC-MS analysis.

The participant who has successfully passed the course is able to:

- decide what data are needed for evaluating the different method performance parameters, understand the meaning of the available data and decide whether the available data are suitable and sufficient;
- select the validation approach and design the experiments for obtaining suitable data;
- quantify the relevant performance parameters using the available data and assess whether the obtained values are realistic;
- assess the fitness of the method for the intended purpose based on the values of the evaluated performance parameters.

Organization of the course material

The course is organized in 11 thematic sections, of which most are in turn split into smaller subsections. The following parts are found in the sections:

1. The sections (and also many subsections) start with a **brief introduction** stating the main topic(s) and study outcomes of the section.
2. The main topic of the respective section is explained in one or several short **video lectures**. The videos are by default streamed in high quality from the UT Video server, which needs quite good Internet connection. If you have slow Internet connection we recommend watching the videos from YouTube, using the links below video screens.
3. The lecture(s) is(are) complemented by a textual part. The textual part is in most cases meant to complement, not substitute the lecture (although in some cases the contents of the lecture are also repeated in some extent). It rather gives additional explanations and addresses some additional topics that were not covered by the lecture.
4. The learners are introduced to the 

software, which helps to carry out validation of chromatographic methods. Learners are encouraged to use VaLChrom throughout the whole course.

5. Most sections end with a self-test, which enables to assess the acquired knowledge and skills. The self-tests contain questions, as well as calculation problems. The self-tests are on one hand meant for the participants to monitor their progress. On the other hand, however, they also promote thinking and provide (by the feedback of the questions) additional knowledge about validation in different practical situations. So, the self-tests are an intrinsic component of the course and participants are strongly recommended to take all of them.

Self-testing

Throughout the course there are numerous self-tests for enabling the participant to test his/her knowledge and skills in specific topics. Each test is graded as a percentage (100% corresponding to correctly answering all questions and correctly solving all problems).

Feedback is given as:

- ✔ Correct answer, correctly recognised and marked by the student.
- ✔ Correct answer, not recognised and not marked by the student.
- ✘ Incorrect answer, however, considered correct by the student.

Explanatory feedback is displayed when wrong answer is selected. All self-tests can be taken as many times as needed and the success of taking these tests will not influence the final grade. We recommend that you take all the tests and work with them until you achieve score 100% and only then move to next topic.

Terminology and definitions

At the end of the thematic section there is **Glossary**, which gives definitions and/or explanations of the terms used in the course. Wherever possible, the used terminology adheres to the one used in the *Tutorial review on validation of liquid chromatography–mass spectrometry methods* (Literature sources 1 and 2).

Literature sources

The main literature sources of the course are here:

1. Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part I. A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio, I. Leito. *Anal. Chim. Acta* **2015**, *870*, 29-44
2. Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part II. A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio, I. Leito. *Anal. Chim. Acta* **2015**, *870*, 8-28
3. B. Magnusson and U. Örnemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods - A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014).
4. SANTE/11945/2015 (Until 01.01.2016: SANCO/12571/2013), Guidance document on analytical quality control and method validation procedures for pesticides residues analysis in food and feed. European Commission, 2015.
5. European Commission Decision 2002/657/EC implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Off. J. Eur. Commun. L221 (2002) 8-36.
6. JCGM 200:2008, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*, 3rd edition. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, 2008.

This list of main literature references of this course is selective, not exhaustive. The references were selected on the basis of the following criteria: (a) widely used and cited; (b) useful under practical lab conditions (i.e. not too deeply scientific or theoretical) and (c) a fairly recent version is available.

In addition, at the end of the course materials there is a more voluminous list of literature sources that are referred to in various parts of the course and are meant for users who are more deeply interested in specific topics. The references are referred to in the course via numbers in square brackets, e.g.: [ref 15].

Course team

Ivo Leito, professor of analytical chemistry at University of Tartu.

Ivo teaches analytical chemistry and metrology in chemistry at all study levels and organizes short training courses for practitioners on different topics of analytical chemistry and metrology in chemistry.
His research work embraces a wide area of topics ranging from studies of superacids and superbases to LC-MS analysis. He is the initiator and coordinator of the *Erasmus Mundus* joint master's programme Excellence in Analytical Chemistry at University of Tartu.

Anneli Kruve obtained her PhD titled Matrix effects in liquid-chromatography electrospray mass-spectrometry from University of Tartu (UT) in 2011. Since 2005 she has been involved in HPLC and LC-MS method development and validation in various fields: bioanalysis, food analysis as well as industrial analysis. Starting from 2011 she works as a research fellow at UT Institute of Chemistry. In 2008 and 2009 she has worked in University of Helsinki in the field on miniaturization of MS ion sources. Her main research fields are method development for pesticide analysis, fundamental studies of ionization efficiency and design of MS ionization sources.

Riin Rebane obtained her PhD in analytical chemistry from University of Tartu in 2012 with a topic on optimization and validation of liquid chromatographic methods with mass spectrometric detection containing derivatization. For the past eight years her main research area has been LC-MS analysis, including method development and validation for various analytes and development of novel derivatization reagents for LC-MS e.g. the DBEMM reagent for amino acids. She is also a head of Research and

Development department in the Estonian Environmental Research Centre and responsible for transferring LC-MS methods from development to routine analysis.

Maarja-Liisa Oldekop obtained her M.Sc. from the University of Tarty (UT) in 2013 and is now a Ph.D. student in the chair of analytical chemistry at UT. Her main field of research is development of LC-MS methods using derivatization. The focus is on matrix influences on this type of analysis, stressing the importance of trueness of the analysis results but also the sensitivity of the method. Since the beginning of 2013 Maarja-Liisa Oldekop works as the quality manager of the UT Testing Centre, which is an ISO/IEC 17025 accredited testing and calibration laboratory.

Hanno Evard obtained his PhD from University of Tartu (UT) in 2016. His PhD work was about the estimation of detection limit for mass spectrometric analytical methods. Since 2009 he has worked on development of several new ionization sources for MS and validation of different LC-MS methods. Since 2016 he is a research fellow in UT focusing on fundamental research and development of novel ionization sources.

Koit Herodes obtained his PhD from University of Tartu (UT) in 2002. Since 2008 he works as the head of the UT Testing centre – a unit providing testing and analysis services and accredited according to ISO 17025 by the Estonian Accreditation Centre. Since 2005 Koit Herodes works as associate professor of analytical chemistry at UT Institute of Chemistry. He has been principal investigator of numerous projects involving LC-MS analyses. Currently he is the PI of the project Development of software for validation of chromatographic methods, which aims at creating web-based software for validation of chromatographic methods.

Karin Kipper obtained her PhD Titled Fluoroalcohols as Components of LC-ESI-MS Eluents: Usage and Applications from University of Tartu (UT) in 2012. Since 2004 she has been involved in the bioanalytical method development and validation for HPLC-UV/VIS and LC-MS analysis, working at UT Institute of Pharmacology and Institute of Chemistry. Starting from 2012 Karin Kipper works as a research fellow at UT Institute of Chemistry. Her main research fields are pharmaceutical bioanalysis (pharmacokinetic/pharmacodynamic studies), pharmaceuticals' pathways in environment and development of novel eluent additives for LC-MS in order to improve separation and peak shape of basic compounds.

Irja Helm obtained her PhD in analytical chemistry from University of Tartu with a topic on validating and optimizing a high accuracy gravimetric Winkler method for determination of dissolved oxygen in 2012. Since 2011 she works as a research fellow in analytical chemistry at University of Tartu and is involved in different metrology-related activities, such as high-accuracy dissolved oxygen measurements and interlaboratory comparisons. Irja teaches practical classes of analytical chemistry. She takes care that metrological concepts and approaches are introduced to students at as early stage of analytical chemistry studies as possible. Her main responsibilities during development of the present course are design and setting up the materials and tests to the web.

Technical design: Educational Technology Centre, University of Tartu.

Contact and feedback

Ivo Leito
University of Tartu
Institute of Chemistry
Ravila 14a, Tartu 50411
ESTONIA
Tel: +372-5 184 176
e-mail: ivo.leito@ut.ee
Skype: leitoivo

Copyright

1. Validation: General

The procedure for demonstrating the reliability of analytical methods is called validation. Validation is required for all analytical methods, but the amount of work and complexity of the issues encountered increases with increasing sophistication of the technique used.

LC-MS is an extremely powerful and at the same time highly sophisticated technique. It offers various operation modes (scan, MRM, etc), numerous parameters and several possibilities of ionization (protonation, adduct formation, etc). However, these numerous possibilities come at a cost: LC-MS systems are complex and not always robust. Therefore it is extremely important that the reliability of LC-MS based analytical methods is assured and this is done by validating the method.

This section introduces the basics of validation and is divided into three subsections:

1.1. The purpose of validation

Method validation is a key activity in chemical analysis, indispensable for obtaining reliable results. The higher the complexity of the method, the more important and voluminous, as a rule, is validation. Methods based on LC-MS are notorious for their complexity, on the one hand because of the instrument itself and on the other hand because LC-MS is often applied to the most complex samples. Therefore, it is important to demonstrate that methods are working as expected (validation) and the obtained results are reliable. This information is relevant both to the laboratory (to be confident in your results or to make adequate changes in method if the performance is not as expected) as well as for the customer. Besides the intrinsic necessity of validation, there are increasingly more regulations affecting laboratories that stipulate method validation as a requirement.

The importance of validation has led to the emergence of a large number of validation guidance materials for laboratories, both of universal nature and sector-specific. Although there is a general agreement on the various validation parameters to be evaluated, diversity prevails about the details and about the methodology employed for validation and acceptance criteria: different recommendations and different sets of terminology are found in different guidelines. As a conclusion, the analytical chemistry community is still far from a consensus about exactly how validation should be carried out, both in general terms and as applied to LC-MS.

Purpose of validation

http://www.uttv.ee/naita?id=23347

https://www.youtube.com/watch?v=Gv8CoI_3gLc

1.2 Carrying out validation

It is always of importance how to perform validation in the most effective way. Based on the literature data and our own experience we suggest a possible general sequence of validation in Figure 1.

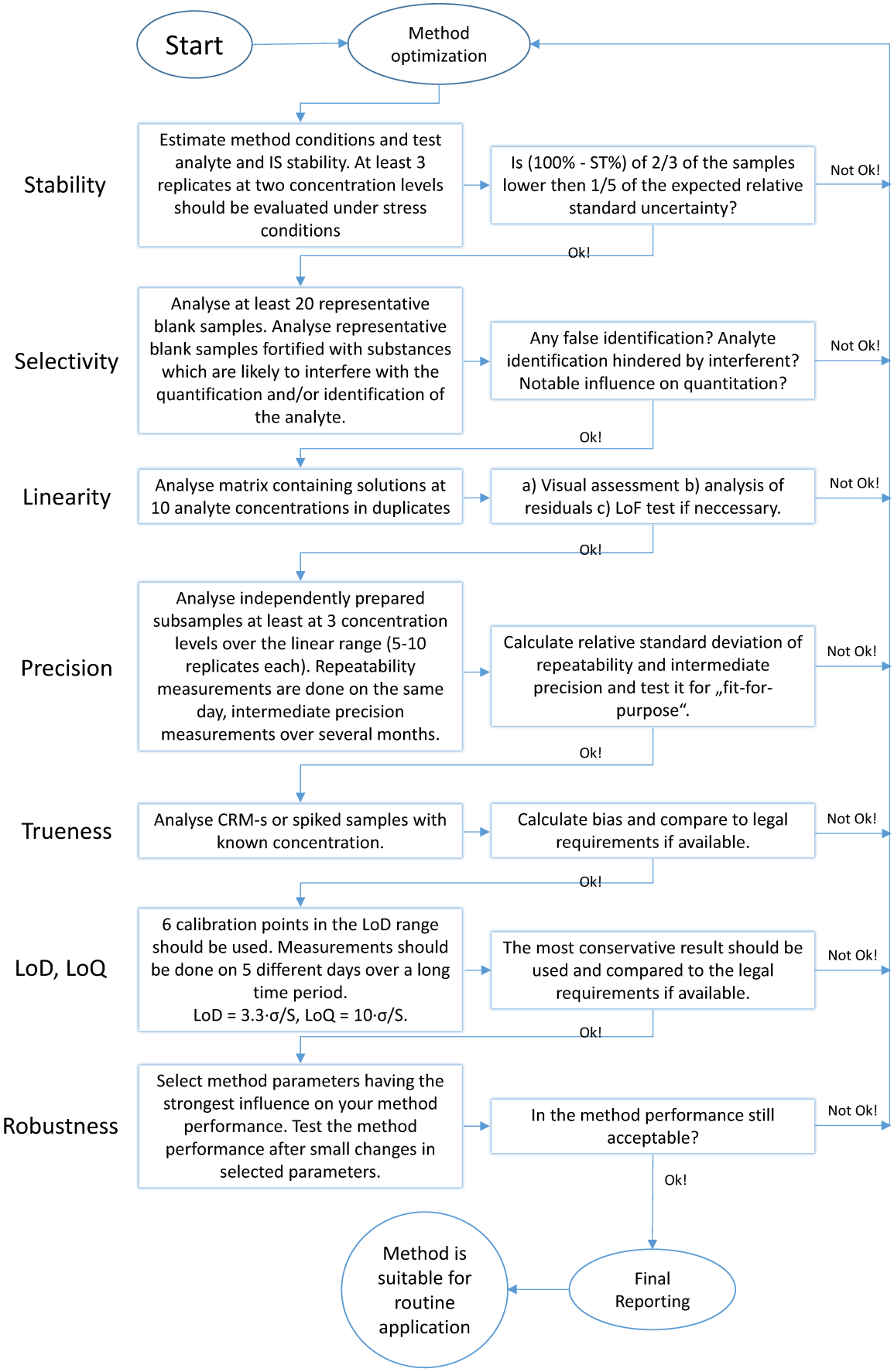


Figure 1. A possible sequence of operations and decisions in LC–MS method validation. All steps are explained in detail in coming chapters. *ST*% refers to stability in per cents, explained in section 8.3. LoF stands for Lack-of-Fit test, explained in 3.3. CRM stands for certified reference material.

Before starting a validation, a clear plan is compiled, which consists of the reason for validation, planned experiments as well as expected outcomes – requirements that need to be met by the method. The requirements often result from guidelines or from other regulatory documents. Making that plan depends on the each different method under development and takes into account all the specific aspects related to that method. After carrying out the necessary experiments, a decision must be made if the results are satisfying and consequently if the method is fit for purpose. Validation is documented in a validation report.

Guidelines generally give recommendations for evaluating separately each performance parameter. At the same time, the guidance on deciding about the whole method's validation is usually very general: validation has to demonstrate that the values of all the evaluated parameters are satisfactory. Few different cases arise.

- (a) When methods are applied in the scope of standards, laws or directives, then the requirements from those documents, if present, must be followed and the decision on validation should be based on these. When the decision on validation suitability is based on the guidelines, then for each parameter a separate decision must be given according to the requirements.
- (b) Sometimes the client can specify the requirements, then the client's requirements are superior to those in the guidelines.
- (c) If there are no external requirements, then the analyst can set up the requirements himself/herself based on his/her knowledge of the subject.

The validation should start with evaluating analyte stability and method selectivity as all the other parameters strongly depend on these. For example, if the analyte extensively decomposes in the autosampler, no linear relation can be achieved. In that case, non-linear calibration models can be considered. Consequently we propose estimation of linearity as the next step, because for an evaluation of trueness and precision we need to know the linear/working range of the method. We propose robustness studies as the last step of validation. It is sometimes suggested to test robustness as one of the first things in method validation or in the end of the method development phase. We find it important to have some insight, as to which are the most important performance characteristics (e.g. closest to the legal limits or the requirements of the client) before deciding which of the method performance characteristics are varied during the robustness studies.

Not only validation but also appropriate documentation of the validation is required for an adequate interpretation as well as on the validity of the obtained results. As the last stage of validation, an assessment of validity (fitness for the intended purpose) of the method should be given, based on the validation results.

Overview of validation

http://www.uttv.ee/naita?id=23287

https://www.youtube.com/watch?v=K312lRGTJFk

1.2_overview_of_validation.pdf 267 KB

1.3. Scope of validation

It is often unnecessary to carry out a determination of all the possible method performance parameters. Validation involving only some of the performance parameters is called partial validation. If a fully validated (e.g. using a collaborative trial) method is set up in the laboratory then it is necessary to carry out so-called verification: verify that it is able to achieve the published values of performance characteristics. Partial validation is justified when a standard method (e.g. ISO, ASTM) is used, small changes are made to previously validated methods or for methods with narrow application range. The small changes can include transfer of the method to another laboratory, adding a new matrix, new reagent in sample preparation, etc. FDA states that in the case of bioanalytical methods it is sometimes sufficient to determine only trueness and repeatability. ICH distinguishes between the methods for identification, impurity analysis and assay analysis. Different performance characteristics need to be determined for different types of methods.

Degree of validation

http://www.uttv.ee/naita?id=23305

https://www.youtube.com/watch?v=sHnYshEUtOs

1.3_degree_of_validation.pdf 198 KB

2. Selectivity and identity confirmation

In this chapter we will rationalize concepts and methods related to selectivity and confirmation of identity of analytes. No matter what kind of analysis is carried out, it is always assumed that the results express the content of analyte and not anything else.

Real-life samples are usually mixtures of many compounds and the analytical method must therefore be selective towards the analyte. IUPAC defines **selectivity** as the extent to which other substances interfere with the determination of a substance according to a given procedure [ref 52]. The larger the interference the less selective is the procedure (1) (method). As the definition implies, methods can be selective to different extent. If a given method is 100% selective, it is said to be **specific**. Analytical techniques are almost never generally specific or it is nearly impossible to prove that. However, analytical methods can be specific within their scope of application, i.e. for a given analyte in a given matrix in a given concentration range.

Note, that terminology regarding selectivity and specificity is not used unanimously in validation guidelines. FDA, AOAC, EMA and Eurachem guidelines use the terms suggested by IUPAC, while ICH and NordVal use the term specificity to denote selectivity.

Analytical method is regarded selective if its results are not affected by other sample components to any significant extent. Compounds, other than analyte, which contribute to the analytical signal are called interfering compounds or interferents. Interfering compounds may:

1. Suppress/enhance the signal by altering chromatographic separation or detector response (the so-called "matrix effect", see sections 5.3 and 5.4).
2. Behave like the analyte and yield the signal indistinguishable from analyte's signal. For example, in case of LC-UV analysis, an interferent would have retention time very similar to that of the analyte and would absorb UV-radiation at the same wavelength as the analyte.

In the case of chromatographic methods it is very common to assess selectivity regarding chromatographic separation of compounds. Ability of detector to further enhance selectivity is sometimes overlooked. But let's start from selectivity provided by chromatography.

***

(1) Instead of the term (analytical) *method* used in this course IUPAC uses the term (analytical) *procedure*.

2.1. Selectivity: LC selectivity

There are a couple of ways to characterize how well chromatographic peaks are separated/resolved. Probably the best known parameter is the **separation factor** (aka selectivity factor) – *a*. For the two peaks of the chromatograms in Figure 1 the separation factor is calculated as follows:

$$\alpha = \frac{t_{RB} - t_M}{t_{RA} - t_M} = \frac{8-1}{6-1} = 1.4 \qquad \textbf{(Eq 1)}$$

The larger is the separation factor, the farther apart are the peak maxima. Separation factors for chromatograms 1 and 2 in Figure 1 are equal, but visually it is clear that peaks in chromatogram 2 are not completely resolved. This is the main drawback of separation factor as numerical characteristic of chromatographic separation – it characterizes only peak separation in terms of retention times, but not peak resolution in terms of lack of overlap.

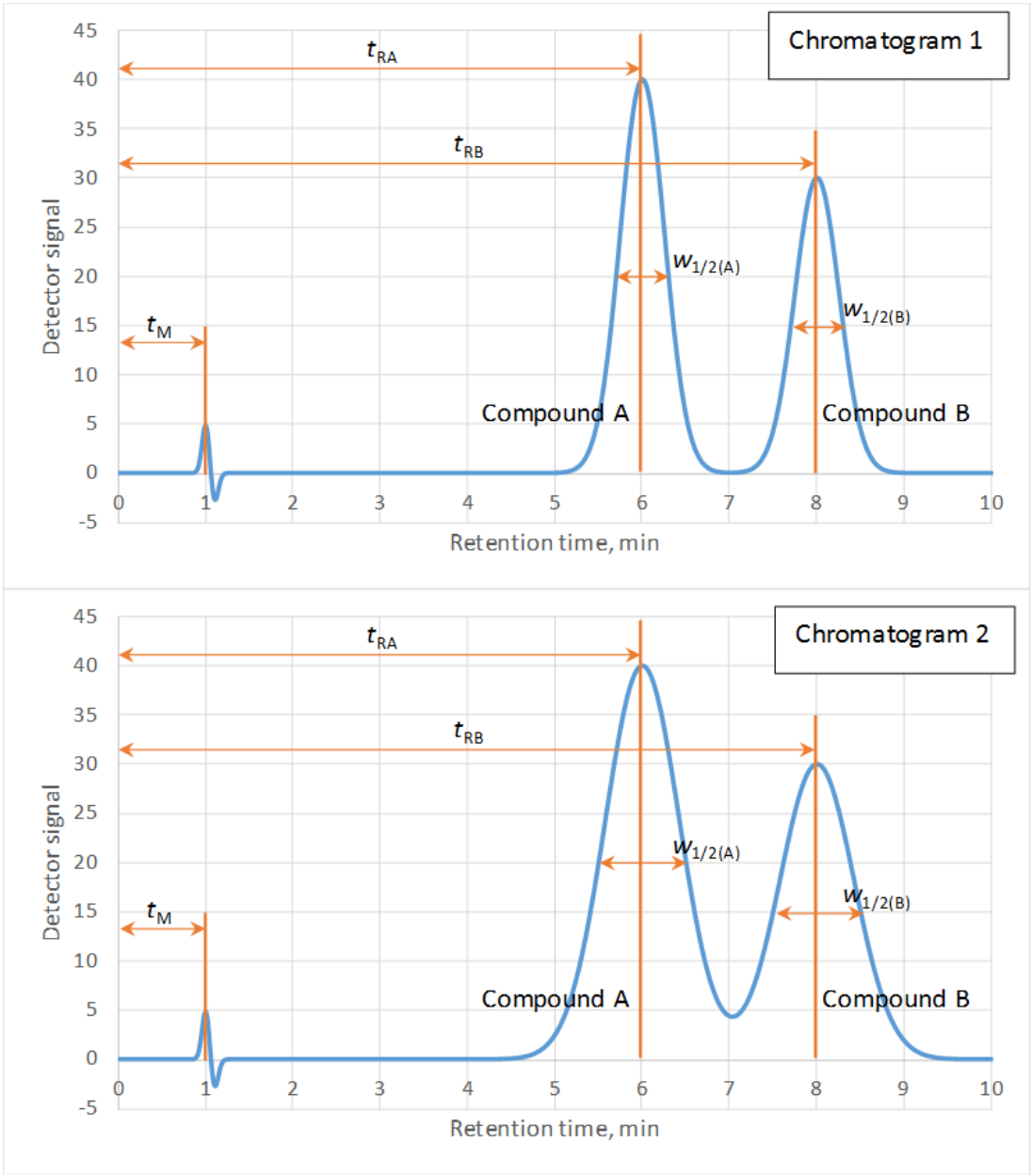


Figure 1. Chromatograms with similar separation factor but different resolution.

The second possibility of quantitatively expressing chromatographic selectivity is using the peak resolution, R_S. R_S takes in addition to retention times into account peak widths at half height, $w_{1/2}$:

$$R_S = \frac{t_{RB} - t_{RA}}{0.85(w_{1/2B} + w_{1/2A})} \qquad \textbf{(Eq 2)}$$

Peak resolution in chromatograms 1 and 2 (Figure 1) are calculated as follows:

Chromatogram 1: $$R_S = \frac{8-6}{0.85(0.6+0.6)} = 1.96$$

Chromatogram 2: $$R_S = \frac{8-6}{0.85(1+1)} = 1.18$$

Different R_S values reflect the actual situation – peaks in chromatogram 1 are better resolved than in chromatogram 2.

In order to numerically express selectivity, some validation guidelines require demonstrating that peak resolution exceeds certain threshold. For example, FDA requires $R_S > 2$ and AOAC requires $R_S > 1.5$. In addition to this, AOAC (also ICH) requires that no other compound should be detectable at analyte retention time, when methods like IR, NMR or MS are used. To check for presence or absence of coeluting compounds Eurachem suggests demonstration of separation on a column of different chemistry.

Introduction to selectivity and identity confirmation. Chromatographic (LC) selectivity.

http://www.uttv.ee/naita?id=23251

https://www.youtube.com/watch?v=uE-qjB0w4Y4&t=332s

2_1_intro_selectivity_identity.pdf 495 KB

2.2. Selectivity: detector-side (MS) selectivity

What is so special about using MS as the detector in LC? MS adds new dimension – *m/z* – to the analytical signal. Without this additional dimension, we would just know whether some compound is eluting from the column or not. With MS detector the information will be more specific: whether compound with specific *m/z* is eluting from the column or not.

As an example total ion chromatogram (TIC) is presented in Figure 1. In TIC intensities of all the ion abundances in mass spectrum are summed, i.e. discriminating power of MS is not used. As a result, several bigger and smaller peaks, corresponding to different compounds, are observed in the chromatogram.

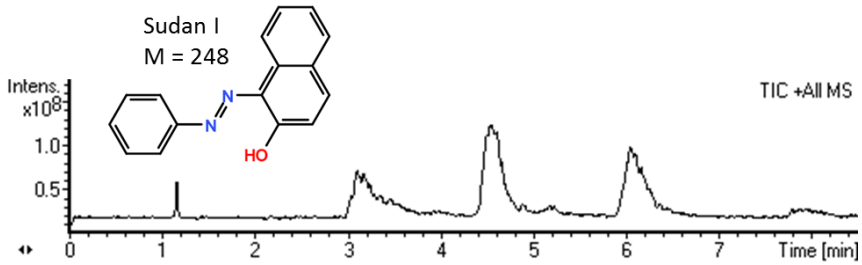


Figure 1. Total ion chromatogram (TIC) of Sudan I dye analysis.

When the additional selectivity provided by MS detection was employed, the chromatogram in Figure 2 was obtained. In the case of Figure 2, only the part of MS signal corresponding to *m/z* 249 was monitored, resulting in an EIC – extracted ion chromatogram. As a result, chromatogram with just one peak, corresponding to Sudan I dye was obtained. This example demonstrates that MS can considerably enhance selectivity.

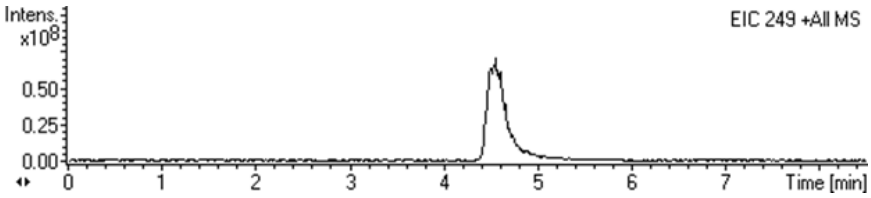


Figure 2. Extracted ion chromatogram (EIC) of Sudan I dye analysis. The *m/z* value 249 correspond to protonated Sudan I molecule $[M+H]^+$.

What if compounds of identical or very close *m/z* are analyzed? In this case there are three approaches. The best (cheapest) would be separation of the compounds chromatographically. If this is not possible, then tandem MS or high resolution MS can be used. In practical LC-MS analysis the first two approaches are routinely used: chromatographic separation is always attempted and as a rule all LC-MS analysis, at least in complex sample matrices, is carried out in tandem MS (also known as MS/MS or MS^2) mode. High-resolution MS is seldom used in routine analysis.

Selectivity: detector-side (MS) selectivity

http://www.uttv.ee/naita?id=23292

https://www.youtube.com/watch?v=cpSy2O-cpSA

2_2_detectorside_selectivity.pdf 485 KB

2.3. Selectivity examples

Example 1

Not all validation guidelines explicitly specify a required limit for chromatographic peak resolution (R_S). In LC-MS incomplete chromatographic separation of analyte peak from neighboring peaks may or may not cause problems. Often *m/z* of analyte differs from that of neighboring peaks or unique MRM transition is used. This way the presence of any potentially overlapping peak is simply neglected. However, even though the analytical signals of these interfering compounds are not registered, they can cause matrix effects. Therefore, it is customary to seek for peak resolution of at least $R_S \geq 1.5$.

Chromatogram in Figure 1 was recorded while developing method for compounds A, B, C and D analysis. One MRM trace is presented for each compound and shown in different color. If one hadn't recorded the transition shown in green, this peak would have missed form the chromatogram and one wouldn't have known that compound C (RT 5.33 min) – is present. But this compound could interfere with detection of other compounds, regardless of whether or not it's signal is recorded. Therefore, one should make sure that the peak is adequately separated from others.

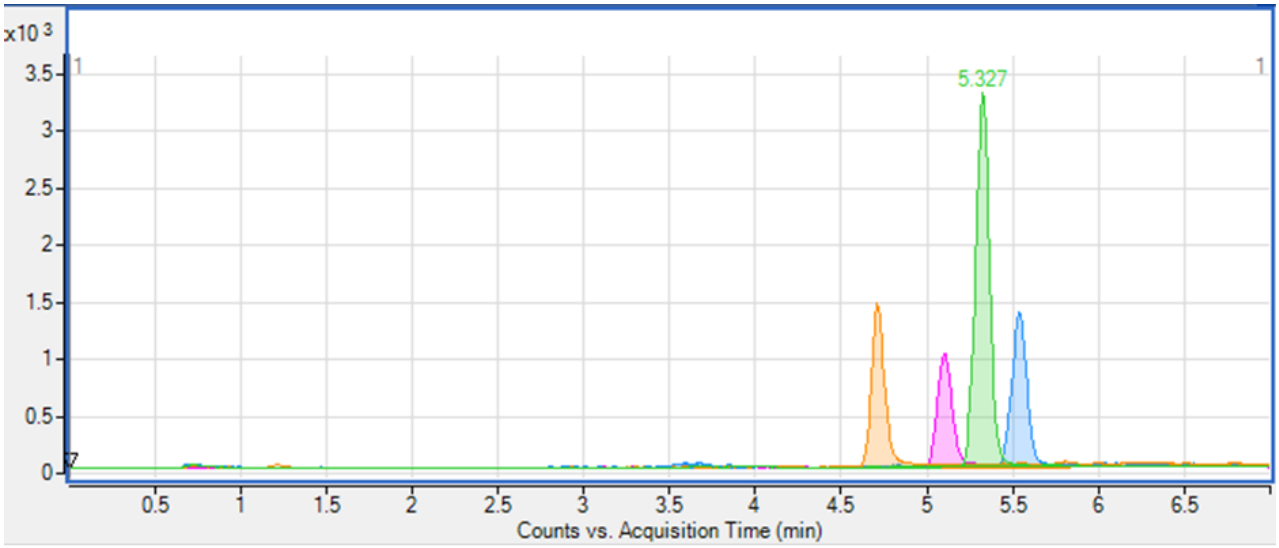


Figure 1. Initial chromatographic separation of compounds A, B, C and D.

Table 1 presents the chromatographic data for above separation. Peak resolution (R_S) from the preceding peak was calculated for the last 3 peaks according to equation 1. It appears that while separation of compounds B and C is at the limit (R_S= 1.5), then separation of C and D is not satisfactory (R_S = 1.4).

$$R_s = \frac{t_{RB} - t_{RA}}{0.85 \cdot (w_{1/2B} + w_{1/2A})} \quad \textbf{(Eq 1)}$$

Table 1. Peak data of the chromatogram in Figure 1.

Analyte	t_R	$w_{1/2}$	R_S
A	4.710	0.079	
B	5.102	0.094	2.7
C	5.327	0.088	1.5
D	5.538	0.093	1.4

Chromatographic separation was further optimized (1) and chromatogram presented in Figure 2 was obtained. As a result, sufficient peak resolution (R_S > 1.5) was achieved for all analytes (Table 2).

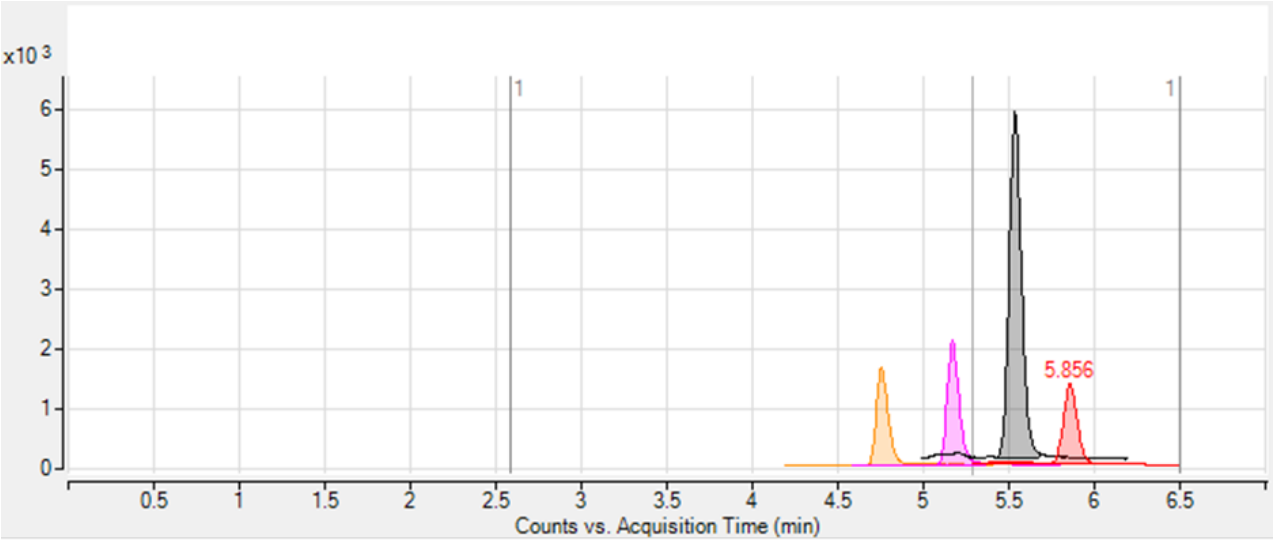


Figure 2. Chromatographic separation of compounds A, B, C and D.

Table 2. Peak data of the chromatogram in Figure 2.

Analyte	t_R	$w_{1/2}$	R_S
A	4.752	0.075	
B	5.169	0.074	3.3
C	5.533	0.077	2.8
D	5.856	0.088	2.3

Example 2

As seen in the previous example, while respecting the selectivity-enhancing power of MS, one must not forget the importance of chromatography. Chromatography becomes crucial if discrimination of isomers – having the same molecular mass – is desired.

For example, two major metabolites of morphine are morphine-3-glucuronide (M3G) and morphine-6-glucuronide (M6G) (Figure 3).

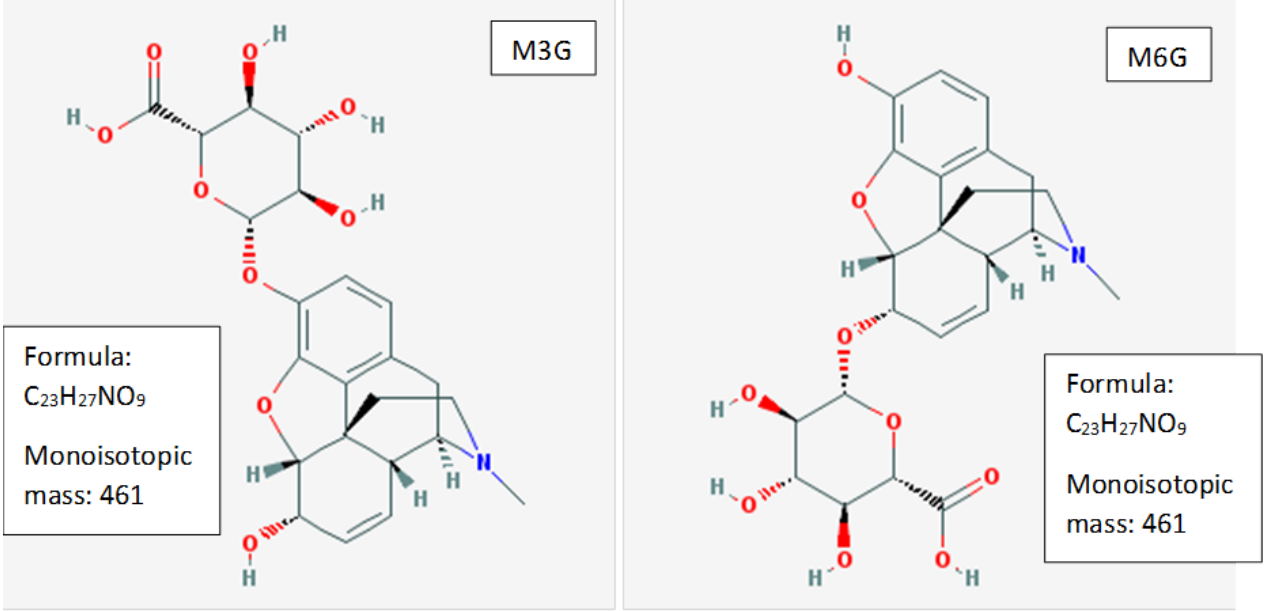


Figure 3. Morphine-3-glucuronide (M3G) and morphine-6-glucuronide (M6G).

M3G and M6G are structural isomers with identical molecular formulae and, consequently, identical molecular masses. In such situation discrimination of the isomers relies completely on chromatographic separation. Mass spectrometer, even if high resolution (HRMS) type, can't distinguish molecules of identical masses and fragmentation patterns (which is the case of these two compounds). M3G and M6G can be conveniently separated chromatographically (Figure 4).

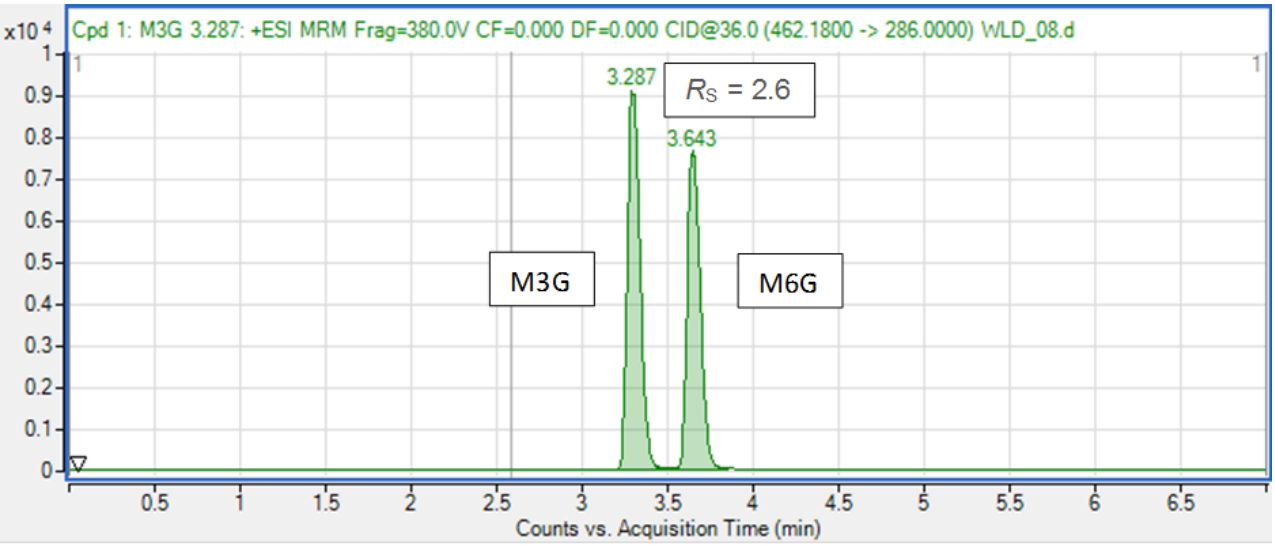


Figure 4. Chromatographic separation of morphine metabolites Morphine-3-glucuronide (M3G) and morphine-6-glucuronide (M6G).

***

(1) There are several options for selectivity optimization in reversed phase LC:

- Percent of organic solvent in mobile phase
- Organic solvent itself (methanol, acetonitrile, THF or their mixture)
- pH of eluent (note the pH limits of your column; in case of LC-MS only volatile pH-buffer components can be used)
- Different column
- Temperature (the effect is usually small, but may help)

2.4. Introduction to identity confirmation

Confirmation of analyte's identity is closely related to selectivity assessment of the method. Validation guidelines usually discuss these topics together and it might be confusing. In case of selectivity assessment, we know where the analyte peak is, or if it is missing. In case of identity confirmation, the problem is reversed – there is a peak on the chromatogram, and the analyst must provide evidence that the peak belongs to the analyte. The information required to collect such evidence is obtained during validation.

The following evidence is used in chromatographic methods for confirming analyte's identity:

1. Using chromatographic retention time. The basic requirement in chromatography is that retention time of a compound must be the same in standard solution and in sample. But similar retention time alone does not confirm the identity of the compound as there are potentially many compounds (sample constituents) that may elute at the same (or very similar) retention time.
2. Spectrometric methods used for detection can provide evidence that the detected compound is the same as the analyte. UV-Vis and fluorescence detectors are classical spectrometric LC detectors, which provide some information about structural features of compounds, but UV-Vis and fluorescence spectra are usually not sufficiently characteristic to the analytes for their positive identification. Mass spectrometry, especially high resolution MS or tandem MS (MS/MS), is the most powerful method for identity confirmation.

Most of the validation guidelines consider LC and some recognize the power of MS for identity confirmation. In contrast, ICH is more general summarizing the general requirements instead of specific techniques (LC, spectrophotometry, etc). In order to demonstrate that the developed method is able to discriminate between compounds of closely related structures, positive results must be obtained for samples containing the analyte and negative results for samples which do not. Also, negative results must be demonstrated for compounds, which are structurally similar or closely related to the analyte. Proper selection of potentially interfering compounds requires scientific judgement, experience and knowledge.

The following two sections look in more detail at identity confirmation via retention time and MS, respectively.

Introduction to identity confirmation

http://www.uttv.ee/naita?id=23630

https://www.youtube.com/watch?v=wKW_1SEsy5M

2_4_intro_identity.pdf 368 KB

2.5. Identity confirmation: retention time

Retention time and peak shape of analyte in standard solution and in sample must match (within some tolerance limit, see below). If they do, then analyte may be present in the sample, but additional confirmation is still required. On the other hand, if retention time or peak shape differ then the peak under question cannot be due to analyte (assuming properly functioning LC system).

For proper chromatographic separation from fast-eluting sample components (which are often abundant), the analyte retention time should be at least twice the dead time of the system. This means that retention factor k (aka capacity factor, k') should be ≥ 1. This recommendation is general in chromatography and is specifically important in LC-MS, because the matrix effects (ionization suppression/enhancement) are most probable close to dead time of the system (see sections 5.3 and 5.4).

According to the most recent SANTE/SANCO validation guidelines retention times of analyte in sample and in standard solution must not differ more than by 0.1 min (1) . This requirement is valid in the case of classical HPLC as well as UHPLC. The European Commission Decision 2002/657/EC specifies a tolerance criterion for analyte retention time relative to the standard substance retention time. The relative retention time of analyte in standard solution must correspond to that in the sample within tolerance of 2.5%.

The shape of chromatographic peak may also be used to evaluate identity of analyte. If peaks in standard solution and sample differ in shape, it may indicate that the compound detected in sample is not the analyte or some compound coelutes with the analyte. It is possible that the analyte peak is distorted in sample injection because of sample matrix. In order to check for this, matrix matched standard solutions or isotopically labelled internal standard (ILIS) should be used. ILIS, although expensive, are rather common in LC-MS analysis. If ILIS are used, then peak shape of analyte and respective ILIS shall match.

Identity confirmation: retention time

http://www.uttv.ee/naita?id=23631

https://www.youtube.com/watch?v=LWAeC2Xr2f0

***

(1) The previous SANCO guidelines established the retention time tolerance as ± 0.2 min [ref 4].

2_5_identity_lc.pdf 369 KB

2.6. Identity confirmation by MS

For identity confirmation MS provides *m/z* values of ions formed from analyte molecules and relative abundancies of the ions. Validation guidelines by SANTE/SANCO [ref 4] and European Commission Decision 2002/657/EC [ref 5] are the most specific with respect to using MS data for identity confirmation and are recommended by the authors of this course. In the rest of this chapter we rely on these two validation guidelines.
For comparison, reference mass spectra should be recorded on the same instrument and using the same operating mode as used for sample analysis. Preferably, reference spectra should be recorded within the same analysis batch with samples.
Recommendations for choosing good diagnostic (characteristic) ions:

- The quasimolecular ion should be involved in identification.
- High *m/z* (*m/z* > 100) ions are more specific than low *m/z* ions.
- Product ions by common losses (eg H_2O, NH_3) are of little diagnostic value.
- The choice of diagnostic ions depends on matrix interferences.

Identity confirmation by MS

http://www.uttv.ee/naita?id=23688

https://www.youtube.com/watch?v=G2IISK_yM0s

The usage of principles outlined in the above video are illustrated in section 2.7.

2_6_identity_ms.pdf 467 KB

2.7. Identity confirmation examples

Example 1

It is the basic assumption of liquid chromatography, that retention time of analyte is the same in chromatograms of calibration solutions and in sample solutions. Two validation guidelines have set clear criteria for retention time tolerance: 2002/657/EC [ref 5] and SANTE/SANCO [ref 4]. 2002/657/EC limits tolerance of relative retention time to 2.5% and the new SANTE requires 0.1 min absolute retention time tolerance (the previous SANCO limit was 0.2 min).

Assessment of retention time deviation according to SANTE.

Analysis of pharmaceutical residues in sewage sludge compost was carried out. Retention times of three analytes in calibration solutions are presented in Table 1 along with average retention times for each analyte.

Table 1. Experimental and average retention times (in minutes) of metformin, diclofenac and carbamazepine in calibration solutions.

	Metformin	Diclofenac	Carbamazepine
Cal 1	7.675	19.172	21.263
Cal 2	7.675	19.162	21.242
Cal 3	7.664	19.172	21.263
Cal 4	7.675	19.172	21.253
Cal 5	7.664	19.161	21.252
Cal 6	7.664	19.161	21.252
Cal 7	7.664	19.172	21.263
Average (*RT*):	7.669	19.168	21.255

Average retention times in calibration solutions are used as reference points to assess deviations of retention times of analytes in sample solutions. Table 2 presents experimental retention times of analytes in sewage sludge compost extracts. Difference of retention time from respective reference value is calculated. Absolute value of this difference must not exceed the tolerance limit set by SANTE (0.1 min). Calculated differences and assessment are also included in Table 2.

Table 2. Retention times of analytes (in minutes) in sewage sludge compost extract sample chromatograms. Differences of retention times from respective reference values are presented along with assessment of the result with respect to the SANTE 0.1 min tolerance limit.

	Metformin			Diclofenac			Carbamazepine		
	RT	Difference	Complies?	*RT*	Difference	Complies?	*RT*	Difference	Complies?
Sample 1	7.598	-0.071	Yes	19.063	-0.105	No	21.23	-0.025	Yes
Sample 10	7.566	-0.103	No	19.183	0.015	Yes	21.285	0.03	Yes
Sample 20	7.577	-0.092	Yes	19.139	-0.029	Yes	21.263	0.008	Yes
Sample 30	7.533	-0.136	No	19.194	0.026	Yes	21.296	0.041	Yes
Sample 40	7.544	-0.125	No	19.183	0.015	Yes	21.285	0.03	Yes
Sample 50	7.522	-0.147	No	19.194	0.026	Yes	21.285	0.03	Yes
Sample 55	7.435	-0.234	No	19.172	0.004	Yes	21.274	0.019	Yes

From Table 2 it is evident that all carbamazepine results and all but one diclofenac results comply with SANTE requirement. But situation is different with metformin – only two results comply. Study of the metformin retention time differences reveals that all the differences are negative. This indicates that some systematic effect is present. In some cases, sample matrix affects the retention. Matrix matched calibration could then solve the problem. It is also possible, that contamination gradually builds up in the column and alters the stationary phase properties and consequently the retention – more extensive cleanup of sample extracts could help.

In conclusion, SANTE retention time tolerance check can be rather restrictive, but it is also helpful for diagnosing possible problems in analytical methods.

Assessment of retention time deviation according to 2002/657/EC.

Validation guideline 2002/657/EC by European Commission uses the concept of relative retention time (*RRT*) for establishing the tolerance limit of retention time. *RRT* is defined as ratio of retention time of analyte (*RT*) to retention time of internal standard (IS). For this example, metformin can be regarded as retention time internal standard. Experimental data obtained from calibration solutions along with *RRT* values are presented in Table 3. Averages of *RRT*-s obtained from calibration solutions serve as reference values for the analytes.

Table 3. Absolute (*RT*) and relative (*RRT*) experimental retention times obtained from calibration solutions. Metformin is used as internal standard (IS). Average relative retention times of diclofenac and carbamazepine.

	Metformin (IS)	Diclofenac		Carbamazepine	
	RT	*RT*	*RRT*	*RT*	*RRT*
Cal 1	7.675	19.172	2.498	21.263	2.770

Cal 2	7.675	19.162	2.497	21.242	2.768
Cal 3	7.664	19.172	2.502	21.263	2.774
Cal 4	7.675	19.172	2.498	21.253	2.769
Cal 5	7.664	19.161	2.500	21.252	2.773
Cal 6	7.664	19.161	2.500	21.252	2.773
Cal 7	7.664	19.172	2.502	21.263	2.774
Average (RRT_{Ref}):			2.499		2.772

Similarly to calibration data, *RRT* values are now calculated for analyte peaks in samples (Table 4). To compare *RRT* in sample to reference value relative difference (as percent) is calculated as follows:

$$Rel.Diff = \frac{RRT - RRT_{Ref}}{RRT_{Ref}} \cdot 100\% \quad \textbf{(Eq 1)}$$

where *RRT* is relative retention time of analyte in sample injection and RRT_{Ref} is average of calibration solution *RRT*-s.

Unsigned value of *Rel.Diff* is compared to the tolerance limit set by 2002/657/EC: 2.5%. Results of this assessment are also presented in Table 4.

Table 4. Absolute (*RT*) and relative (*RRT*) experimental retention times with respect to the internal standard (IS) in sewage sludge compost extract sample chromatograms. Relative differences (as percentage) of relative retention times from respective reference values are presented along with assessment of the result with respect to the 2002/657/EC tolerance limit of 2.5%.

	Metformin (IS)		Diclofenac			Carbamazepine			
	RT	*RT*	*RRT*	Rel. Diff.	Complies?	*RT*	*RRT*	Rel. Diff.	Complies?
Sample 1	7.598	19.063	2.509	0.40%	Yes	21.23	2.794	0.80%	Yes
Sample 10	7.566	19.183	2.535	1.40%	Yes	21.285	2.813	1.50%	Yes
Sample 20	7.577	19.139	2.526	1.10%	Yes	21.263	2.806	1.20%	Yes
Sample 30	7.533	19.194	2.548	1.90%	Yes	21.296	2.827	2.00%	Yes
Sample 40	7.544	19.183	2.543	1.70%	Yes	21.285	2.821	1.80%	Yes
Sample 50	7.522	19.194	2.552	2.10%	Yes	21.285	2.83	2.10%	Yes
Sample 55	7.435	19.172	2.579	3.20%	No	21.274	2.861	3.20%	No

According to Table 4 data, almost all retention times are judged as compliant. Only retention times of analytes in the last sample were found non-compliant. Reason for this is retention time of IS, which is smaller than expected.

Compared to tolerance limits set by SANTE, these of 2002/657/EC appear to be more lenient. Internal standard is required for retention time checking by 2002/657/EC, which makes the approach slightly less comfortable to use.

Example 2

SANTE/SANCO criteria for identification and identification points of 2002/657/EC.

Requirements for number and type (molecular, adduct or product ion) of monitored ions in SANTE/SANCO validation guidelines are presented in a straightforward manner in Table 4 of both guidelines. For example, signal from at least two ions must be recorded with any mass spectrometer operated in MS/MS mode regardless of whether mass spectrometer is of low or high resolution. For further requirements see Table 4 in SANTE/SANCO guidelines [ref 4].

The system used by 2002/657/EC [ref 5] is more elaborate and is based on identification points. The required number of identification points depends on the type of analyte (see [ref 5] for full details):
4 points for substances having anabolic effect and unauthorized substances;
3 points for veterinary drugs and contaminants.

Table 5 presents numbers of identification points earned per ion for different MS techniques.

Table 5. Identification points earned for MS techniques according to 2002/657/EC.

MS technique	Identification points earned per ion
Low resolution (LR) MS	1.0
LR-MS^n precursor ion	1.0
LR-MS^n transition products	1.5
High resolution MS	2.0
High resolution MS^n precursor ion	2.0
High resolution MS^n transition product	2.5

In addition to the data in Table 5, there are two additional rules – each ion may be counted only once and "transition product" includes all generations of transition products. Few examples:

- Unit resolution MS operated in SIM mode and three ions recorded – 3 points.
- Triple quadrupole MS (unit resolution, MS/MS) recording 2 products ions for the same precursor – 4 points (1*1.0+2*1.5=4).
  - For example transitions 142.1→94.1 and 142.1→112.2 for methamidophos (pesticide) analysis.
- LC-MS^3 experiment with ion trap MS (unit resolution) recording 2 product ions – 5.0 points (1.0+1.0+2*1.5=5.0).
- Combined quadrupole-TOF MS with low resolution precursor ion and 2 high resolution product ions – 6 points (1.0+2*2.5=6.0)

Example 3

Calculation and assessment of ion ratios.

SANTE/SANCO and 2002/657/EC limit ion ratio tolerances according to Table 6.

Table 6. Ion ratio tolerances of SANTE/SANCO and 2002/657/EC.

Ion ratio	Relative tolerance (SANTE/SANCO)	Relative tolerance (2002/657/EC)
> 0.5	± 30%	± 20%
0.20 – 0.50	± 30%	± 25%
0.10 – 0.20	± 30%	± 30%
< 0.10	± 30%	± 50%

Ion ratio is calculated as intensity (or peak area) ratio of less intense ion to that of more intense ion. Reference ion ratio value is calculated as an average of ion ratios of calibration solutions. Table 7 illustrates the process of calculating reference ion ratio on the example of propamocarb (pesticide) calibration.

Table 7. Example of reference ion ratio calculation based on peak areas (S1, S2) of two MRM transitions.

Propamocarb	189→102	189→144	Ion ratio
	S1	S2	S2/S1
Calibrant 1	555	259	0.47
Calibrant 2	1176	499	0.42
Calibrant 3	1707	803	0.47
Calibrant 4	2404	991	0.41
Calibrant 5	3031	1312	0.43
		Average:	0.44

In similar manner, ion ratios are calculated for analysis of samples (Table 8).

Table 8. Calculation of ion ratio from peak areas (S1, S1) of two MRM transitions recorded from sample chromatograms.

Propamocarb	189→102	189→144	Ion ratio
	S1	S2	S2/S1
Sample 1	821	281	0.34
Sample 2	2221	1251	0.56

To assess compliance of ion ratios in samples, the following equation can be used:

$$Rel.Diff = \frac{Ion\_ratio_{Sample} - Ion\_ratio_{Ref}}{Ion\_ratio_{Ref}} \cdot 100\% \quad \textbf{(Eq 2)}$$

Relative difference for Sample 1 is -23% and for Sample 2 it is 27%. Comparison with limits in Table 6 (second row), reveals that identity of propamocarb in both samples is confirmed by SANTE/SANCO rules. Regarding 2002/657/EC limits, presence of propamocarb is not confirmed in Sample 2 as 27% > 25%.

Example 4

Selection of product ions.

SANTE/SANCO and 2002/657/EC give recommendations for proper selection of diagnostic ions. In general, fragmentation due to loss of common fragments (H_2O, NH_3 etc) should be avoided, as more background noise and interfering compounds are expected. For example, Figure 1 shows transitions 230→213 and 230→160 registered for clonidine (blood pressure lowering drug). Transition 230→213 is due to loss of NH_3, which leads to noisier background.

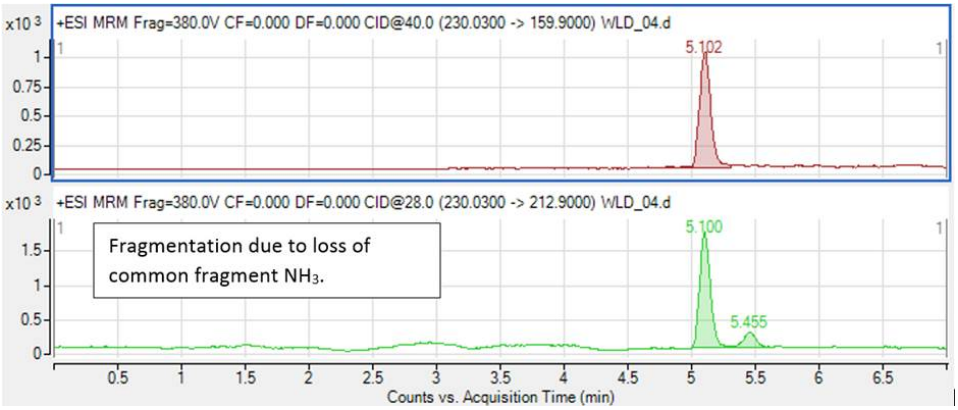


Figure 1. Influence of fragmentation pathways on chromatographic background.

Example 5

How well do all the limitations work?

SANTE/SANCO and 2002/657/EC set rather restrictive rules to LC-MS analysis. Do all these criteria guarantee unambiguous confirmation of analyte? Or maybe the rules are too strict?

A method was developed for tebufenpyrad (pesticide) analysis in ginger extract. In Figure 2 overlaid chromatograms of two transitions of tebufenpyrad are presented for calibration solution and ginger extract.

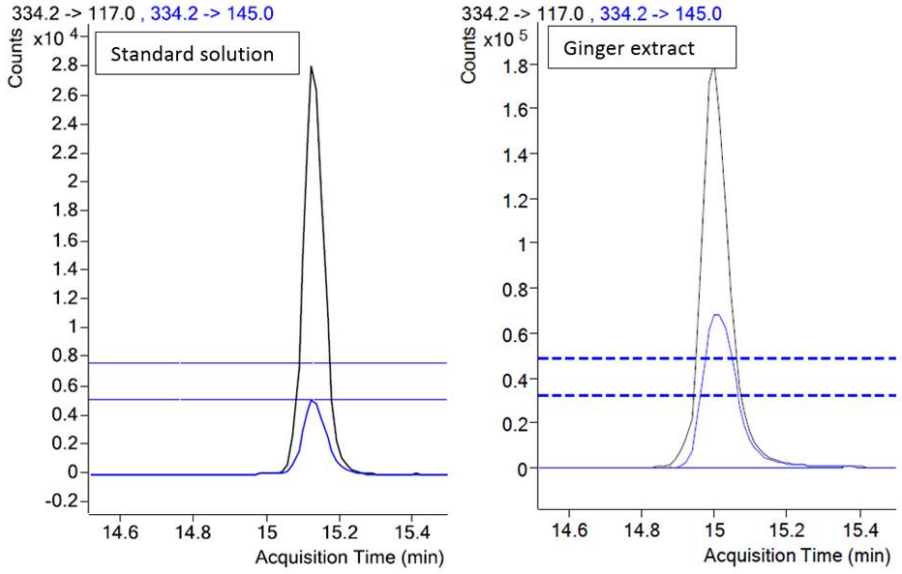


Figure 2. Chromatograms of two transitions of tebufenpyrad in standard solution and sample. (Data from http://www.agilent.com/cs/library/eseminars/public/Triggered%20MRM%20for%20the%20Analysis%20of%20Pesticides.pdf)

Retention times of peaks are similar: 15.1 min in standard and 15.0 min in extract and would pass requirements for retention time match by SANTE (0.1 min). Ion ratio of the two transitions for tebufenpyrad in calibration standard is 0.21 and for peaks in ginger extract 0.469. Relative difference of ion ratios is 123% which exceeds the tolerance limits set by SANTE/SANCO and 2002/657/EC. Therefore, presence of tebufenpyrad in ginger could not be confirmed. Indeed, further analysis revealed that it was an endogenous ginger compound, which yielded this peak.

So, thanks to the strictness of the rules a false positive result was avoided.

3. Linearity of signal, linear range, sensitivity

This chapter covers the topics related to the linearity of the signal: linearity, linear range and sensitivity. These parameters are defined, practical suggestions are given for setting up and carrying out experiments for their determination, as well as for the corresponding data analysis.

Introduction to linearity, linear range and sensitivity

http://www.uttv.ee/naita?id=23249

https://www.youtube.com/watch?v=XXIKFmzgJ8A

All section 3 slides are downloadable in one file:

3_linearity.pdf 6.7 MB

3.1. Linearity

Linearity is the method's ability to obtain test results, which are directly proportional to the concentration of analyte in the sample. (1)

The term linearity of signal can, in the context of LC–MS, have two closely linked meanings: (a) linear relationship between analyte signals and analyte concentrations in calibration samples and (b) linear relationship between analyte signals and analyte concentrations in samples containing matrix components. The latter meaning is becoming increasingly more used and is used also in this course. The reason is that, if the analyte signal in samples is linear, then almost certainly it is linear also in calibration solutions, while the opposite is not necessarily true. The most common cause for this is the phenomenon of matrix effect, which is addressed in chapter 5 of this course. Linearity of the calibration graph is closely related to choosing calibration model and working range.

Factors affecting linearity:

- Firstly, **the ion source** behaves linearly only if the ionization efficiency of the analyte is independent of its concentration in the effluent. As the ionization efficiency of the analyte and its behaviour at different concentrations depends on the used ion source, the linear ranges differ between different ion sources. In the ESI source the linear dependence usually holds at lower concentrations, but at higher concentrations the excess charge on the surface of the droplets becomes limiting and linearity is lost. Also the co-eluting compounds can influence the ionization process (the so-called matrix effect) and lead to decrease or loss of linearity. Therefore it is most important to investigate linearity in the presence of matrix compounds.
- Secondly, during **the ion transport** from the ion source to the mass analyzer the number of successfully transported ions must be proportional to the number of ions formed in the source. (2)
- Thirdly, the linearity of ion signal depends on the mass analyzer design and on the linearity of the detector's signal. The contemporary ion detectors are highly linear, so that mass analyzer design is the deciding factor here. Mass analyzers are characterized by **transmission efficiency** (3) , which is the ratio of number of ions that are finally detected and number of ions that entered the mass analyzer. In order to display linear behavior the transmission must be independent from concentration.

***

(1) The term "linear" can in mathematics refer to the fact that the statistical models considered are linear in terms of all parameters (coefficients), also including polynomial relationships. However in analytical method validation, the term "linear" is used to describe the case where a straight line can be used to describe the relationship between LC–MS signal and analyte concentration.

(2) This ion transport is, in essence, the transport of ions from the atmospheric pressure region in the ion source to the low pressure region of mass spectrometer with minimum loss. Collisions or formation of clusters can cause losses of ions.

(3) Transmission (in mass spectrometry) - the ratio of the number of ions leaving a region of a mass spectrometer to the number entering that region. [http://goldbook.iupac.org/T06478.html]
Transmission efficiency – how many of the ions produced in the source region actually reach the detector. This is an important measure of sensitivity for mass spectrometers.[http://science.widener.edu/svb/massspec/massspec.pdf]

3.2. Experiment setup and evaluation of the data

Experiment planning for evaluation of linearity

http://www.uttv.ee/naita?id=23307

https://www.youtube.com/watch?v=PdLsxDExgV0

1) Type of calibration samples

When choosing the proper quantitation method, we can choose between calibration samples (calibration standards, calibrants) containing matrix and calibration samples that are matrix free. In case of LC-MS analysis, we should prefer samples containing matrix in order to take into account possible matrix interference on the ionization of the analyte. Blank matrix extracts of as similar as possible matrix type as the sample, are suitable for this.

If the sample to be analysed is diluted prior to the analysis, the matrix concentration in the matrix-matched calibration standards should be diluted proportionately, so that the matrix amount in each analyzed sample is constant.

If solvent calibration is used, a comparison of calibration graphs in matrix and solvent should be carried out. (1)

2) Concentrations of calibration samples

The highest and lowest concentrations should be appropriate for the method keeping in mind the predicted variation of analyte levels in the samples. As a minimum, 6 different concentrations are necessary according to most validation guidelines. This is also acceptable for statistical tests carried out later for linearity evaluation, where 4 degrees of freedom is considered minimal. However, as we do not know the span of the linear range at this step of validation, some concentrations might fall out of the linear range. Therefore, using 10 concentration levels encompassing the expected linear range is recommended. Moreover, the concentrations should be approximately evenly spaced over the chosen concentration range, to ensure that different parts of the calibration graph are covered with approximately the same density of data points.

3) Measurement protocol

For LC-MS, the order of measuring the solutions in the series and the number of replicate measurements with every solution is important due to the possible drift or contamination of the instrument. Therefore, the analysis order of calibration samples should be random.
It is useful to analyze calibrants in a manner as similar as possible to the unknown samples, i.e. calibration samples should be randomly ordered and placed between unknown samples in the analytical run. The calibration samples should be analyzed at least twice (and average values are used in linearity calculations).
In the following video the preparation of matrix matched calibration samples on an example of pesticide analysis in tomato is shown.

Carrying out the experiment for linearity evaluation

http://www.uttv.ee/naita?id=23480

https://www.youtube.com/watch?v=x8KaQ7aC_mI

Evaluation of linearity from the experimental data

For quantitative analysis the calibration data is plotted on a calibration graph, where concentrations are on the x-axis and the signals are on the y-axis. (2)

From this graph we can have the first evaluation of linearity using the visual evaluation approach. In addition, calculating the residuals can also be useful.

Evaluation of linearity

http://www.uttv.ee/naita?id=23351

https://www.youtube.com/watch?v=4c3EMRpFDf0&t=29s

Absolute residuals are found as the difference between experimental (y_i) and calculated ($\hat{y}_i$) signal values: $y_i - \hat{y}_i$. In addition, relative residuals can be used:

$$Y_i = \frac{y_i - \hat{y}_i}{\hat{y}_i} \quad \textbf{(Eq 1)}$$

For more complex cases, where the linearity cannot be confirmed by neither the visual evaluation nor the residuals, statistical approaches can be of help.

Statistical approaches for evaluation of linearity

http://www.uttv.ee/naita?id=23683

https://www.youtube.com/watch?v=QFNuo-Jk2Ws

Different expressions of signal values are used for statistical approaches:

y_i is the experimental signal value at concentration level i.

$\acute{y}_i$ is the average value of experimental signal from replicate measurements at concentration level i.

$\hat{y}_i$ is the signal value at concentration level i, calculated using the calibration function.

In addition, n is the number of concentration levels and *p* is the number of replicate measurements at one concentration level.

Tabulated *F*-values can be found here: http://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.htm

The $F_{tabulated}$ can also be found in excel using the following function:
=F.INV.RT(*p*;DoF1;DoF2),
where:
p is the probability of the F distribution (on 95% confidence interval p=0,05),
DoF1 is the number of degrees of freedom of the numerator MSS_{LoF} (*n*-2),
DoF2 is the number of degrees of freedom of the denominator MSS_{error} (*n**(*p*-1)).

Lack-of-fit test

IUPAC validation guideline suggests using the lack-of-fit test. The extent of deviation of the points from the line caused by random scatter of the points is estimated from replicate measurements (mean sum of squares of random error (MSS_{error})).

This is compared to the extent of deviation of the points from the line caused by mismatch of the calibration model (mean sum of squares due to lack of fit MSS_{LOF}).

$$F = \frac{MSS_{\text{LoF}}}{MSS_{\text{error}}} = \frac{\sum(\bar{y}_i - \hat{y}_i)^2 \Big/ (n-2)}{\sum(y_i - \bar{y}_i)^2 \Big/ n(p-1)} \quad \textbf{(Eq 2)}$$

If the $F_{calculated}$ is higher than $F_{tabulated}$, the model cannot be considered fit for the data.

Goodness- of-fit test

The goodness-of-fit test uses the mean sum of squares of the factors (MSS_{factor}) describing the variance described by the model and the mean sum of squares of the residuals ($MSS_{residuals}$).

$$F = \frac{MSS_{\text{factor}}}{MSS_{\text{residuals}}} = \frac{\sum(\hat{y}_i - \bar{y}_i)^2 \Big/ (p-1)}{\sum(y_i - \hat{y}_i)^2 \Big/ (n-p)} \quad \textbf{(Eq 3)}$$

If the $F_{calculated}$ is higher than $F_{tabulated}$, the model differs systematically from the data.

Mandel's fitting test

This test compares the fit of two models: the fit of a linear model with a fit of the nonlinear model. For both of the models, the residual standard deviation is found using this equation:

$$S_{Y_i} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}} \qquad \textbf{(Eq 4)}$$

The $F_{calculated}$ is found:

$$F = \frac{S_{y1}^2 \cdot (n-2) - S_{y2}^2 \cdot (n-3)}{S_{y2}^2} \qquad \textbf{(Eq 5)}$$

If the $F_{calculated}$ is higher than $F_{tabulated}$, the linear model cannot be applied.

Intercept

One important issue concerning the calibration graph is how to handle the intercept. As usually a linear calibration graph model is used, the model has two parameters: slope and intercept. Slope gives us an estimation of the sensitivity of our method (signal per one concentration unit, see section 3.4), while intercept shows the estimate of the signal for a blank sample.

For most of the HPLC detectors, including MS, it is fair to assume that sample without analyte gives no signal. Also for most detectors intercept is not consistent with the physical model behind the calibration principle. For example in case of HPLC-UV/Vis the calibration model should follow the Beer´s law:

$$A = c \cdot l \cdot e \qquad \textbf{(Eq 6)}$$

where A is measured absorbance, c is analyte concentration, l is optical path length and e is the molar absorption coefficient. Therefore, as the physics behind the HPLC-UV/Vis signal does not contain intercept, it is worth checking if the intercept is statistically significant at all.

The statistical way to evaluate the importance of intercept would be via t-test. In order to carry out t-test the linear regression is run with intercept, i.e. in the form LINEST(Y1:Y2; X1:X2; 1; 1), and the obtained intercept value is compared to zero taking into account the standard deviation of the intercept and the number of points on the calibration graph. However, a simpler method can be used that is based on the assumption of normal distribution and also assuming that there is a sufficient number of points on the calibration graph. In this case the t-value is substituted with 2 (referring to the 95% confidence level in normal distribution).

If

Intercept < 2 · Stdev_intercept

then it can be assumed with 95% confidence that intercept is insignificant and can be disregarded in the calibration model. The following form of LINEST spreadsheet function is used in this case: LINEST(Y1:Y2; X1:X2; 0; 1). Setting the third parameter in the function to zero forces intercept to zero.

If intercept, however, is statistically significant, but the physical model behind the detection does not contain intercept, it may be an indication of problems. Most commonly:

- (a) a linear model is fitted to data that are in fact nonlinear (e.g. saturation of signal at higher concentrations);
- (b) blank samples produce signal because of carryover, contamination, etc.

Both of these should be carefully studied and if possible removed.

Evaluation of linearity (visual evaluation, residuals)

http://www.uttv.ee/naita?id=24974

https://www.youtube.com/watch?v=l-AXgA31xRY&feature=youtu.be

Evaluation of linearity (lack-of-fit test)

http://www.uttv.ee/naita?id=24975

https://www.youtube.com/watch?v=dHf7e70uLao NEW, with better quality

An example of linearity evaluation in VaLChrom validation software

http://www.uttv.ee/naita?id=24976

https://www.youtube.com/watch?v=Vu3RKf5Kti4&feature=youtu.be

***

(1) Both standard (in solvent) and matrix—matched calibration curves should be constructed. If the matrix does not interfere with the analysis and the use of standard (in solution) calibration curve is justified, the slopes of these two graphs should not differ statistically. This can be shown using a *t*-test. In order to do so the residual variances of the two graphs should be equal. This can be confirmed using an *F*-test.

An example where the slopes of the two graphs do not differ statistically and the use of standard calibration graph is justified can be found in ref 53.

An example where the matrix interferes with the analysis and matrix matched calibration is used can be found in ref 54.

(2) Calibration curve definition by VIM - the expression of the relation between indication and corresponding measured quantity value. "The term "curve" implies that the line is not straight. However, the best (parts of) calibration lines are linear and, therefore, the general term "graph" is preferred." [http://www.fao.org/docrep/w7295e/w7295e09.htm#TopOfPage]

linearity_example_lack_of_fit.xlsx 67 KB

3.3. Estimating the linear range

Linear range, different ranges

http://www.uttv.ee/naita?id=23250

https://www.youtube.com/watch?v=uc6ZqIKC-Qc

Dynamic range - In the dynamic range the response changes when the analyte concentration is changed but the relationship may be non-linear. If the response is linear it can be specified as a dynamic linear range. (1)

Working range - The working range is the range where the method gives results with acceptable uncertainty. Working range can be wider than linear range. [ref 3]

This means that there is a correlation between the analyte concentration and instrument's signal and the concentrations of analyte can be determined within this range - from LoQ to anomalies of the sensitivity.

Calibration range - Calibration range is the interval between the upper and lower concentration of analyte which can be determined with demonstrated precision, accuracy and response function. [ref 7]

Linear range or linear dynamic range – Range of concentrations where the signals are directly proportional to the concentration of the analyte in the sample.

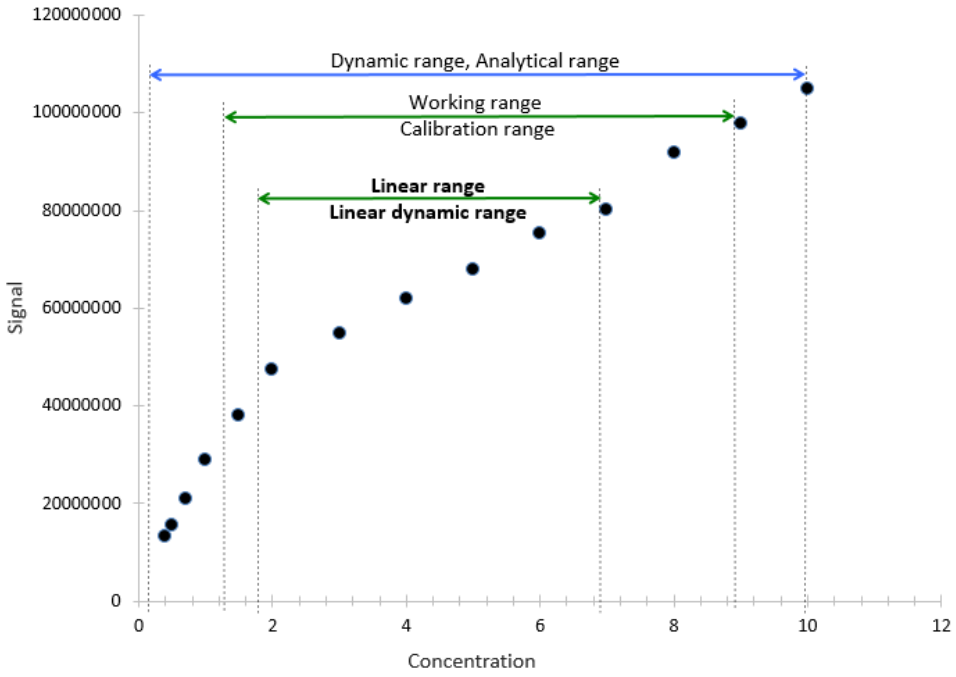


Figure 1. Different ranges

Linearity should be confirmed for the expected working range, including the chosen matrix. Linear range can be found from linearity assessment experiments, however, the criteria for the linear range can be different. The linear range should cover 0–150% or 50–150% of the expected analyte concentration.

While the linear range for LC–MS instruments is usually fairly narrow (and depends on the compound), several possibilities have been used in order to widen the linear range. Using isotopically labeled internal standard (ILIS) is one of them. While the signal-concentration dependence of the compound and the ILIS may not be linear, the ratio of the signals may be linearly dependent on the analyte concentration. But even then, care has to be taken when working outside of the linear range of the signals.

One way of increasing the linear range is by working at lower concentrations (and diluting samples), if the analyte signal is intense enough. Another way to increase the linear range, specifically in the case of LC-ESI-MS is to decrease charge competition by lowering flow rate in ESI source, e.g. by using nano-ESI.

***

(1) http://goldbook.iupac.org/D01874.html

3.4. Sensitivity

Sensitivity

http://www.uttv.ee/naita?id=23289

https://www.youtube.com/watch?v=h59-5voDcQU

Another parameter concerning the relationship between the analyte signal and concentration is sensitivity. **Sensitivity is the change in the response of a measuring instrument divided by corresponding change in the stimulus or simply the gradient of the calibration function** (1) The numerical value (and unit) of sensitivity is arbitrary and depends on the instrument used and its settings. Sensitivity is not an essential parameter during method validation, however it is very important in method optimization and quality assurance procedures for routine monitoring of the instrument's performance. Sensitivity of a method can be easily estimated from the linearity evaluation experiments.

Although sensitivity is not to be confused with limit of quantitation (LoQ), these terms are interrelated: with a given signal to noise ratio, the higher the sensitivity, the lower the LoD and LoQ. Sensitivity is also directly related to ionization suppression – in fact the essence of ionization suppression is decrease of sensitivity due to co-eluting compounds (see chapter 5).

The main use of sensitivity as a parameter is threefold:

1. optimization (for maximizing sensitivity) of the method parameters during method development,
2. daily optimization of the instrument parameters and
3. monitoring of the instrument performance.

***

(1) Sensitivity is often interpreted as related to the detection/determination ability. For example, in the recent FDA's Bioanalytical Method Validation guidance document sensitivity is defined as "the lowest analyte concentration that can be measured with acceptable accuracy and precision (i.e., LLoQ)". In this course the term sensitivity is not used in this meaning.

4. Precision

In the next parts of the course three important validation parameters (method performance parameters) are introduced and explained: precision (this section), trueness (section 5) and accuracy (section 7). In section 6, different aspects of practical determination of precision and trueness are explored.

Precision characterizes the closeness of agreement between the measured values obtained by replicate measurements on the same or similar objects under specified conditions. Precision relates to the *random error* of a measurement system and is a component of measurement uncertainty. Precision can be evaluated with a sample that doesn't necessarily have to have a known analyte content.

For the sake of completeness, let us also briefly address trueness and accuracy here:

Trueness relates to the *systematic error* of a measurement system and if rigorously defined, refers to the agreement between the average of infinite number of replicate measured values and the true value of the measured quantity. In practice, trueness is evaluated from a finite, but reasonably large number of measurements and reference values are used instead of the true value.

Measurement **accuracy** expresses the closeness of a single measurement result to a reference value. Method validation seeks to investigate the accuracy of results by assessing both **systematic** and **random** effects (errors) on single results. These errors are caused by a range of reasons, such as imperfect analyte recovery during sample preparation, possible ionization suppression of the analyte, possible instability of the analyte and others. These errors put together give us the (total) error. Since this error is not experimentally accessible, we operate with estimates of errors – the performance characteristics.

Accuracy embraces both **trueness** and **precision** and can in principle be characterized via precision and trueness. A single-parameter expression of accuracy is **measurement uncertainty**. (Figure 1)

Introduction to precision, trueness and accuracy

http://www.uttv.ee/naita?id=23353

https://www.youtube.com/watch?v=_04F0ViTgPw

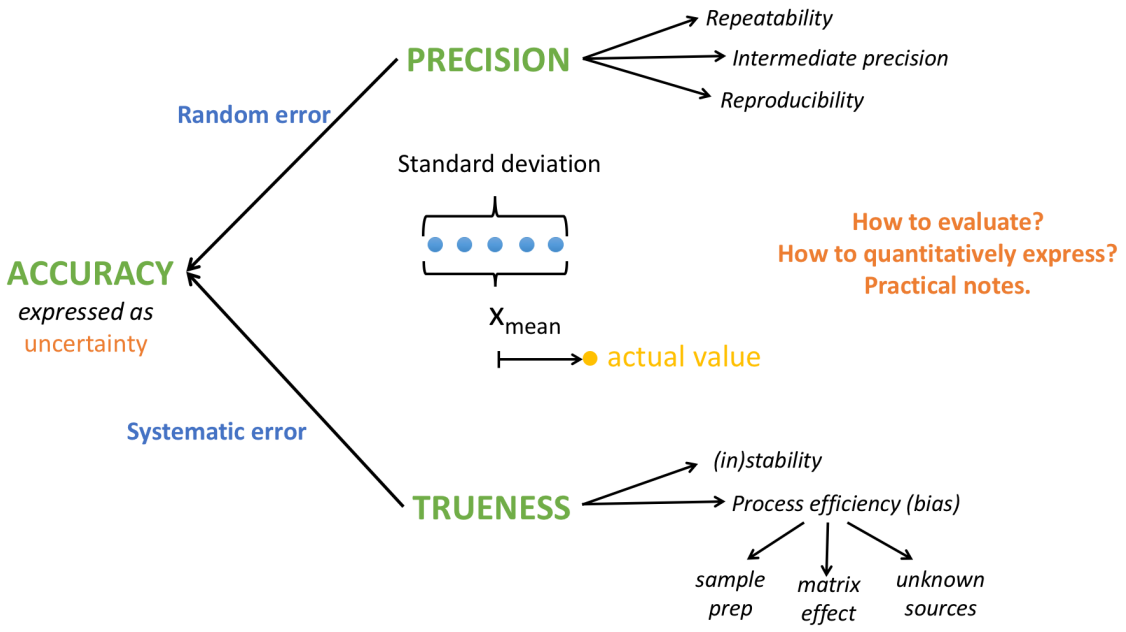


Figure 1. Summarized graphical presentation of the sections 4 to 7.

4.1. Repeatability, intermediate precision and reproducibility

Precision can have different meanings, depending on what level of variability is included.

Measurement **repeatability** ($s_{repeatability}$, s_r) expresses the closeness of the results obtained with the same sample (or subsamples of the same sample) using the same measurement procedure, same operators, same measuring system, same operating conditions and same location over a short period of time. These conditions are called repeatability conditions. The short period of time is typically one day or one analytical run. Repeatability is expected to give the smallest possible variation in results.

Intermediate precision ($s_{intermediate\ precision}$, s_{RW}) (occasionally called within-lab reproducibility) is, differently from repeatability, the precision obtained within a single laboratory over a *longer period of time* (generally at least several months) and takes into account more changes than repeatability. In particular: different analysts, calibrants, batches of reagents, columns, spray needles etc. These factors are constant within a day (i.e. behave systematically within day timescale) but are not constant over a longer time period and thus behave as random in the context of intermediate precision. Because more effects are accounted for by the intermediate precision, its value, expressed as standard deviation (see the next section), is larger than the repeatability standard deviation.

Reproducibility (occasionally called between-lab reproducibility) expresses the precision between measurement results obtained at *different laboratories*. Sometimes a mistake is made and the term reproducibility is used for within-laboratory studies at the level of intermediate precision. Reproducibility is not always needed for single-lab validation. However, it is beneficial when an analytical method is standardized or is going to be used in more than one laboratory (e.g. method developed in R&D departments).

Repeatability, intermediate precision and reproducibility

http://www.uttv.ee/naita?id=23352

https://www.youtube.com/watch?v=ur9gRxpFpPw

4.1_precision.pdf 83 KB

4.2. Calculation of precision

Evaluation of precision requires a sufficient number of replicate measurements to be made on suitable materials. The materials should be representative of test samples in terms of matrix and analyte concentration, homogeneity and stability, but do not need to be Certified Reference Materials (CRMs). The replicates should be independent, i.e. the entire measurement process, including any sample preparation steps, should be repeated. The minimum number of replicates specified varies with different validation guidelines, but is typically between 6 and 15 for each material used in the study.

Precision is usually expressed as "imprecision": an absolute standard deviation (*s*) (Eq 1), relative standard deviation (RSD) (Eq 2), variance (s^2) or coefficient of variation (CV) (same as RSD). Importantly, precision is not related to reference values.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_{mean})^2}{n-1}} \quad \textbf{(Eq 1)}$$

$$RSD = \frac{s}{x_{mean}} \quad \textbf{(Eq 2)}$$

s – standard deviation

n – number of measurements made

x_i – each measurement result

x_{mean} – mean value of the measurement results.

It is not possible to estimate a reliable standard deviation from data sets with few replicate measurements. If acceptable, the values calculated from several small sets of replicate measurements can be combined (pooled) to obtain estimates with sufficient number of degrees of freedom. The pooled standard deviation is calculated with Eq 3 or with Eq 4, if the number of measurements made with each sample is the same. Pooled standard deviation is useful if for some reason it is impossible to make a sufficient number of replicates with one sample – either because of time constraints (e.g. stability) or because the available amount of sample is insufficient. Standard deviations can also be pooled when the same analyte is determined in different samples. However, in this case the samples should be reasonably similar (but not necessarily identical) in terms of analyte content and matrix.

$$s_{pooled} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + \cdots + (n_k-1)s_k^2}{n_1 + n_2 + \cdots + n_k - k}} \quad \textbf{(Eq 3)}$$

$$s_{pooled} = \sqrt{\frac{s_1^2 + s_2^2 + \ldots + s_k^2}{k}} \quad \textbf{(Eq 4)}$$

k – number of samples

s_1, s_2 etc – sample standard deviations

n_1, n_2 etc – number of measurements made for different samples

General formula for the case when experiment is done with different samples, each measured on a single day:

$$RSD_{pooled} = \sqrt{\frac{(n_1-1)RSD_1^2 + (n_2-1)RSD_2^2 + \ldots + (n_k-1)RSD_k^2}{n_1 + n_2 + \ldots + n_k - k}} \quad \textbf{(Eq 5)}$$

k – number of groups (in this case samples)

RSD_1, RSD_2, etc – within group standard deviations

n_1, n_2, etc – numbers of measurements made for different samples

Pooled standard deviation is weighted root mean square (RMS) average of standard deviations, whereby numbers of degrees of freedom of the individual standard deviations serve as weights.

4.3. Acceptance criteria

Specification of acceptance criteria for precision is quite difficult because the criteria should be tailored to each specific application, or as it is called, methods should be "fit-for-purpose". ICH, Eurachem and IUPAC do not specify the acceptance criteria for precision. FDA and EMA say that within-run and between-run CV should be within 15% of the nominal value (20% at LoQ level). SANTE/SANCO and EU directive 2002/657/EC state that CV should be smaller than 20%.

It is important to keep in mind that ideally determination of precision should be carried out for all matrices that are within the scope of application of the method. This general principle becomes even more important for LC–MS methods, where matrix influence can be significant (sample preparation, ionization suppression, etc.).

For LC–MS, it is important to determine precision at different concentrations (at least at two levels: low and high), as most guidelines suggest, since there is strong evidence that repeatability and intermediate precision in LC-MS is concentration-dependent.

4.4. Examples

Example for calculating repeatability

http://www.uttv.ee/naita?id=24824

https://www.youtube.com/watch?v=ckcU4iPvhlg

Example for calculating intermediate precision

http://www.uttv.ee/naita?id=24825

https://www.youtube.com/watch?v=97oc1HpgnEM

precision_blank.xlsx 109 KB

5. Trueness

Part 5 of this course is focused on the different aspects of trueness and its quantitative estimate, bias. In section 5.1 the general information and definitions are given. In particular, this section explains the bias constituents (i.e. causes of bias). Section 5.2 addresses evaluation of (overall) bias in practice. Sections 5.3 and 5.4 look at more specific aspects in relation to trueness (bias), specifically in LC-MS analysis, including experimental determination of some of the bias constituents.

5.1 Bias and its constituents

Different guidance materials use different terms for expressing trueness (Table 1). In this course we use the term trueness with the meaning given in the International Vocabulary of Metrology (VIM) [ref 6] - closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value. The term accuracy has a different meaning and will be discussed more thoroughly in section 7.

Table 1. Terms used for trueness in different guidance materials.

Organization	Term	Meaning according to VIM
Eurachem, AOAC, ISO	Accuracy Trueness	Accuracy Trueness
ICH, FDA, EMA	Accuracy	Trueness
IUPAC, NordVal	Trueness	Trueness

As a rule, trueness of a method is quantitatively expressed as bias or relative bias. Bias is defined as the estimate of the systematic error. In practice bias is usually determined as the difference between the mean obtained from a large number of replicate measurements with a sample having a reference value. It can be expressed as absolute bias (Eq 1 below), i.e. simply the difference, or as relative bias (Eq 2 below), i.e. as difference divided by the reference value. The main causes of bias in LC-MS results can be termed as bias constituents:

1. Bias caused by analyte loss during sample preparation, expressed quantitatively by recovery (R);
2. Bias due to the limited stability of the analyte (see in Stability section) of the analyte in sample solution (B_{stab});
3. Bias due to the ionization suppression/enhancement , i.e. matrix effect ($ME_{ionization}$);
4. Bias due to other possible effects (B_{other}), e.g. purity of the standard substance, calibration bias of volumetric ware.

Table 2 presents the relations between these terms.

Table 2. Bias and related terms together with the equations for calculating them.

Expression	Calculation (1)	Comments
Bias	*Absolute bias:* $bias = X_{lab} - X_{ref}$ **(Eq 1)**	Bias takes into account the effects influencing the result that are systematic over a long term, (5) occurring at any stage of the analytical process. Bias can be expressed as absolute or relative bias.
	Relative bias (2) $bias(\%) = \frac{X_{lab} - X_{ref}}{X_{ref}} \cdot 100\,\%$ **(Eq 2) (3)**	Absolute bias is useful when it is either constant over the used concentration range or it will be evaluated separately at different concentrations.
	$bias = \frac{X_{lab}}{X_{ref}} = ME_{ionization} \cdot R \cdot B_{stab} \cdot B_{other}$ **(Eq 3) (4)**	Relative bias is useful when the absolute bias is proportional to the analyte concentration and it is desired that the same bias estimate could be used at different concentration levels.
Process efficiency (2) , *PE*	$PE = \frac{X_{lab}}{X_{ref}}$ **(Eq 4) (6)**	Process efficiency refers to the joint effect of possible losses during sample preparation and ionization suppression/enhancement in the ion source. *PE* is a useful parameter for characterizing an analysis method when it is either required for characterization of the method or when it is intended to carry out correction with *PE* (more generally: bias correction).
	$PE = \frac{m_{analyte\ detected}}{m_{analyte\ in\ sample}} = ME_{ionization} \cdot R$ **(Eq 5)**	
Recovery (2) , *R*	$R = \frac{m_{analyte\ extracted}}{m_{analyte\ in\ sample}}$ **(Eq 6)**	Recovery (7) expresses the efficiency of the sample preparation step: the proportion of an analyte obtained from the sample during sample preparation (see also (4)).
Ionization suppression/enhancement (matrix effect), $ME_{ionization}$ (2) , (8)	$ME_{ionization} = \frac{m_{analyte\ detected}}{m_{analyte\ extracted}}$ **(Eq 7)**	In LC-MS the term matrix effect refers to the suppression (usually) or enhancement (rarely) (9) of

		analyte ionization in the ion source by co-eluting compounds originating from the sample matrix. $m_{\text{analyte detected}}$ – analyte amount detected in the sample $m_{\text{analyte extracted}}$ – analyte amount actually extracted from the sample.
Stability, B_{stab}	See section 8 for discussion	This bias constituent takes into account losses due to analyte decomposition. Depending at which stage of sample preparation decomposition occurs there are different types of stability (see section 8 below).
B_{other}		This bias constituent takes into account other bias sources (e.g. calibration of glassware) that are not connected to the above mentioned factors.

(1) X_{lab}: average of results obtained by laboratory; X_{ref}: reference value.

(2) Can be expressed as simple ratio or as percentage (in the latter case the ratio is multiplied by 100).

(3) This bias is expressed as percentage. Absence of bias corresponds to 0%. Negative bias values indicate negative and positive bias values positive bias.

(4) This way of expressing relative bias differs from the one in Eq 2. Absence of bias in this case corresponds to 1. Bias values below 1 indicate negative and bias values above 1 indicate positive bias. In many areas the ratio $X_{\text{lab}}/X_{\text{ref}}$ is interpreted as recovery, i.e. the four bias components in this equation would combine into recovery. However, in LC-MS it is useful to make distinction between recovery – relating specifically to sample preparation – and other bias components.

(5) Bias can also be determined as short-term i.e. within-day bias (see below), but long-term bias is more useful, e.g. for measurement uncertainty estimation, as is explained in section 7.

(6) This equation holds, if B_{stab} and B_{other} are insignificantly different from 1.

(7) In the case of most other analytical techniques recovery would include also the possible matrix effects, so that it would effectively be equal to *PE* as defined above.

(8) There are different ways of expressing the matrix effect. We use the way which is similar to expressing recovery and process efficiency.

(9) In the case of alleged ionization enhancement it may under closer examination turn out that the reason for enhanced signal is not analyte signal enhancement but interference from some other compound in the matrix, which accidentally gives precursor and product ions with the same m/z value as the analyte [ref 23] i.e. apparent ionization enhancement may in fact be caused by insufficient selectivity.

Trueness and Precision

http://www.uttv.ee/naita?id=23293

https://www.youtube.com/watch?v=NvmMbrrDjD4

Matrix effect

http://www.uttv.ee/naita?id=23247

https://www.youtube.com/watch?v=nevwRplJNKU

Figure 1 illustrates interrelations between the different bias components. Process efficiency embraces both sample preparation recovery and possible ionization suppression/enhancement in the ion source. An additional important source of bias is possible instability of the analyte. The remaining area on the figure stands for all other (usually smaller) bias components, such as, e.g. calibration of glassware. In LC-MS literature process efficiency (*PE*) is often used as an LC-MS specific term for the overall trueness (if B_{other} and B_{stab} are insignificant).

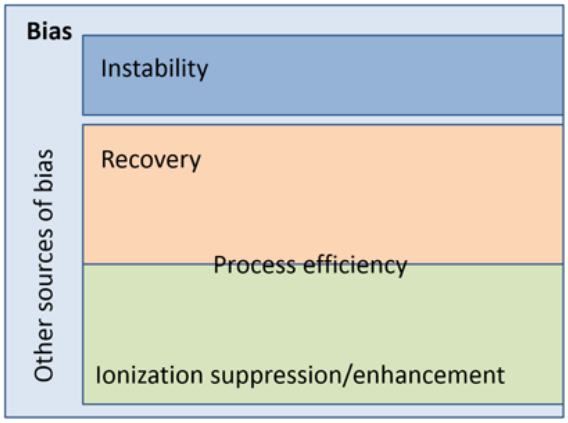


Figure 1. Relations between the bias constituents.

Bias is also dependent on the examined timeframe. In the short term, e.g. within a day, a number of effects cause bias, such as e.g. deviation of the calibration graph of that day from the "true" calibration graph. In the long term, e.g. half a year, if a new calibration graph was made every day, its effect becomes random.

For this reason, within-day bias is larger than long-term bias – the longer time we examine the more effects will switch their status from systematic to random. Consequently, within-day precision (repeatability, s_r) is smaller than long-term precision (within-lab reproducibility, s_{RW}). Figure 2 explains these relations.

Determining precision is significantly easier and the obtained precision estimates are generally more reliable than bias estimates. Therefore, if possible, it is more useful and informative to work in such a way that as few as possible effects are accounted for in bias and as many as possible in precision. This means that when possible, using intermediate precision and long-term bias estimate is more useful and reliable than repeatability and short-term bias estimate.

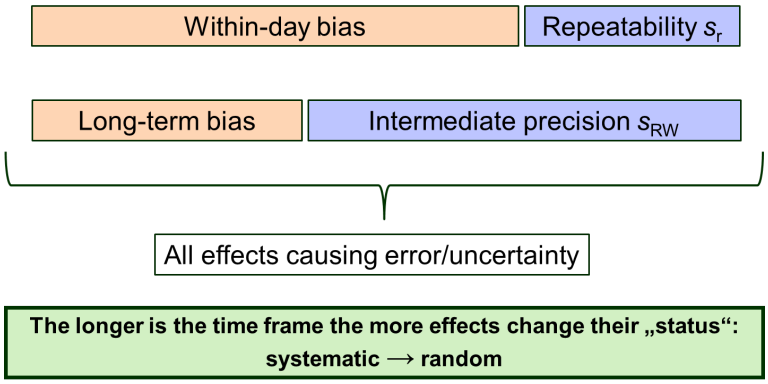


Figure 2. Dependence of bias on the examined timeframe.

5.1_trueness_and_precision.pdf 61 KB
5.1_matrix_effect.pdf 237 KB

5.2 Determining bias in practice

The following practical aspects are important for determining trueness:

1. Obtaining the reference value (see also section 6.2)

Reliable reference value is essential for determining bias. There are essentially 4 approaches of obtaining reference values:

(a) Certified reference materials (CRMs) with sufficiently small uncertainty of the reference value.

(b) Reference materials without a certificate. This includes laboratory reference materials and samples of proficiency testing exercises.

(c) Reference values obtained from spiking studies. Spiking means adding known amount of analyte to a known amount of the matrix, which may or may not contain the analyte. If the amount of the added analyte and the amount of matrix are accurately measured and if the analyte is carefully dispersed into the matrix then the *increase* of analyte content in the matrix will be accurately known and can be used as reference value.

(d) Reference value for a sample can be obtained from analysis with a reference method. Reference method in this context means a method that is more reliable and provides lower uncertainty that the method currently under validation.

The reference values provided by these four approaches have different quality. Reference values from CRMs are generally considered the most reliable, but an important prerequisite is matrix match between the CRM and actually analyzed samples. The reference values of reference materials without certificate are usually less reliable (and can be significantly less reliable) and matrix match is required as well as in the case of CRMs. Spiking can give reliable reference values if done carefully. The main issue is dispersing the spiked analyte in the matrix the same way as the native analyte. The possibility of this varies widely between matrices (easily doable with water and blood plasma, very difficult with soil or minerals). If spiking is doable then ideal matrix match can be achieved, because routine samples can be spiked. Measuring samples with a reference method can in principle give very reliable reference values, but this approach is seldom used because a suitable reference method is usually not available.

Some further aspects of reference values are addressed in section 6.2.

2. Number of replicates: this should be higher than 2 and it should be proportionally larger if the intermediate precision standard deviation is higher. Replicate measurements are needed for separating systematic effects from random effects. The number of replicates is just as important in determining bias as it is in determining precision.

3. Time range of analyzing the replicates: this time range should be at least few weeks (preferably few months). This way the obtained bias will be long-term bias (see section 5.1).

The following video presents a practical example of estimating trueness/bias:

Trueness (CRM)

http://www.uttv.ee/naita?id=24826

https://www.youtube.com/watch?v=htHmkmbTcMI

The trueness/bias estimate obtained in this video is bias according to Equations 1 and 2 in section 5.1. It embraces all four constituents as explained in 5.1 but this way of estimating bias does not enable "dissecting" the obtained bias into its constituents.

trueness_blank.xlsx 585 KB

5.3 Qualitative estimation of matrix effect

This and the following section will look in more detail at some of the bias components. Ionization suppression/enhancement (matrix effect), as one of the most problematic issues in LC-MS analysis receives the most attention, followed by recovery (and the composite of the two – process efficiency). Analyte stability, another important issue, especially in biomedical analysis, is discussed in a separate section 8.

Qualitative estimation of matrix effect

http://www.uttv.ee/naita?id=23477

https://www.youtube.com/watch?v=zZNIkDy9a-Y

The first method of qualitative estimation of matrix effect is based on detecting the presence of ionization suppression (1) by recording the matrix effect profile with post-column infusion [ref 25]. For this the blank sample extract – not containing the analyte – is injected into the LC. At the same time a stream of analyte solution is mixed with the chromatographic effluent exiting the column (post-column mixing) and the mixture is delivered into the ion source. MS monitors the analyte signal.

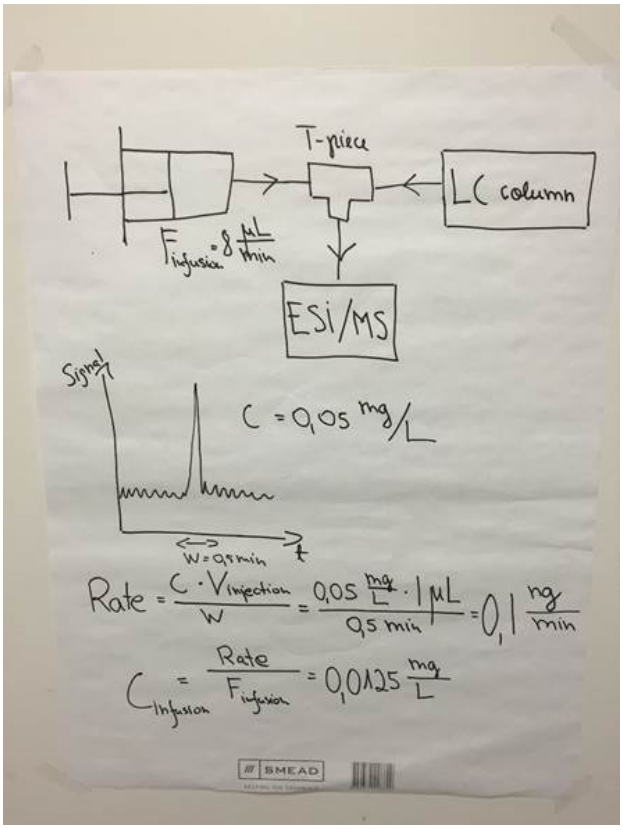


Figure 1. The experimental setup for recording the matrix effect profile.

The same is done while injecting a blank solvent into the LC column. If no suppression is present, then the analyte MS signal intensities are equal for sample and solvent injection within the precision of the MS signal. For the retention time region where ionization suppression occurs, the MS signal in the sample injection decreases. In order to have a method not affected by ionization suppression the analyte peak should elute away from the suppression region. This approach does not enable quantitative evaluation of ionization suppression/enhancement but only confirmation of its presence/absence and is very useful for tuning chromatographic separation of the analyte and possible compounds causing ionization suppression.

The new FDA validation guide proposes evaluation of ionization suppression by assessing the parallelism of dilution plots [ref 8].

We suggest building these plots in axes calculated analyte concentration vs dilution factor (defined as V_{sample}/V_{sum}).

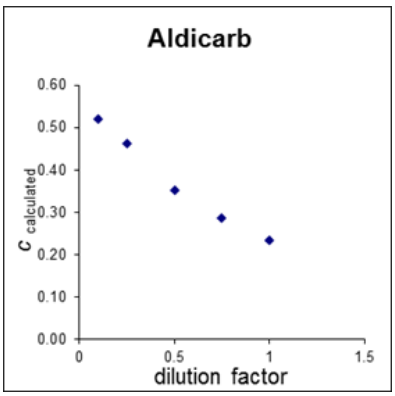


Figure 2. Example dilution plot of Aldicarb determination in garlic.

From these plots matrix effect occurrence can be observed as a decrease (enhancements) or increase (suppression) of calculated concentration with decreasing dilution factor. For example on a plot given above Aldicarb is experiencing strong ionization suppression, that is decreased with dilution.

This approach is also applicable if blank matrices are not available although no quantitative conclusions can be made based on the comparison of dilution curves.

***

(1) The following text is relevant for both ionization suppression and enhancement. However, since ionization suppression is significantly more common, we occasionally use only the term "suppression" in the following text.

5.4 Quantitative estimation of matrix effect, recovery and process efficiency

Quantitative estimation of ionization suppression is possible with post-extraction addition methods as is explained in the following videos. The first video explains the principles of evaluating matrix effect and also touches upon its relations with recovery and process efficiency:

Calculating matrix effect, recovery and process efficiency

http://www.uttv.ee/naita?id=23248

https://www.youtube.com/watch?v=vwRVkhZ8GiY&feature=youtu.be

The second video explains evaluating matrix effect, recovery and process efficiency in practice:

Practical evaluation of matrix effect, recovery and process efficiency

http://www.uttv.ee/naita?id=23476

https://www.youtube.com/watch?v=vwRVkhZ8GiY

For this approach the analyte standard solution with known concentration is prepared in solvent and analyzed with LC-ESI-MS giving the peak area (signal) $S_{standard}$. Also a blank sample extract is prepared and spiked with the analyte at the same concentration level and thereafter analyzed giving peak area (signal) S_{sample}. The ionization suppression/enhancement effect can be calculated:

$$ME_{ionization} = \frac{S_{post\,extraction\,spiked\,matrix}}{S_{Solvent}} \cdot 100\% \qquad \textbf{(Eq 1)}$$

$ME_{ionization}$ value 100% indicates no effect, less than 100% indicates ionization suppression and $ME_{ionization}$ over 100% indicates ionization enhancement due to coeluting sample compounds. From this definition, most often used in LC-MS literature, some possible misunderstandings arise. The expression "reduce matrix effect" does not mean reduced value of %ME, but a $ME_{ionization}$ value becoming closer to 100%.

Calculating matrix effect based on signals

http://www.uttv.ee/naita?id=24822

https://www.youtube.com/watch?v=XZALDnUV9xs&t=32s

Calculating matrix effect based on concentrations

http://www.uttv.ee/naita?id=24818

https://www.youtube.com/watch?v=_BkvRORa4ZM

Calculating matrix effect based on slopes

http://www.uttv.ee/naita?id=24821

https://www.youtube.com/watch?v=WJ9-O4gJxlk

Sometimes also the positive/negative $ME_{ionization}$ scale is used, where 0% denotes no effect, values above 0% ionization enhancement and below 0% suppression. The corresponding equation is:

$$ME_{ionization} = \frac{S_{post\,extraction\,spiked\,matrix} - S_{Solvent}}{S_{Solvent}} \cdot 100\% \qquad \textbf{(Eq 2)}$$

Instead of comparing the peak areas, calibration graph slopes can be compared [ref 26].

A similar approach is described in the FDA upcoming validation guide. In this approach, two calibration graphs are constructed, one in the solvent and the other in the post-extraction spiked samples. This approach is usable also in the case when blank matrix is unavailable.

Several aspects have to be kept in mind:

(a) The intercepts of both calibration graphs have to be negligible so that the ionization suppression/enhancement would not depend on the analyte concentration. Unfortunately, the latter is not always true.

(b) Before using the approach based on slope, linearity needs to be studied. In the literature this approach for ionization suppression/enhancement is often used and sometimes also combined with *F*- and *t*-test or ANOVA to evaluate statistical significance of the obtained matrix effect values.

(c) All of the above described calculations can be done either in signal scale or concentration scale. The obtained results are fairly similar if the samples used for ionization suppression/enhancement study are within the linear range and the intercept of the calibration graph is negligible. If these requirements are not fulfilled, it is more useful, from the method point of view, to use concentration-based calculations.

If signal- or concentration-based calculations are used (not slope-based), the number of samples and replicates used for suppression/enhancement assessment during validation becomes an issue. Often several replicates are run at one or more concentration levels.

(d) It has been often shown that matrix effect depends on the sample source (eg different variety of fruit [ref 27]). It is therefore also recommended to use different matrices for suppression/enhancement evaluation. In the literature the number of matrices used varies a lot.

(e) In the literature [ref 28, ref 29] it has been observed that ionization suppression/enhancement may strongly vary from day to day and it cannot be estimated once during method optimization/validation and be used later for result correction.

Reducing matrix effect

Reducing matrix effect

http://www.uttv.ee/naita?id=23288

https://www.youtube.com/watch?v=Oh3eZKYpa6g

Whenever possible and practical, ionization suppression (matrix effect) should be eliminated or significantly reduced. If it is not possible to reduce ionization suppression to the level of being insignificant, it should be taken into account in calculation of the results. Several approaches have been suggested and tested for reducing the ionization suppression effect, mainly focusing on ESI ionization source. In broad terms the approaches can be categorized as based on (a) sample preparation, (b) instrumental modifications and (c) modifications in LC method:

(a) Less than ideal sample preparation may be viewed as the main reason of occurrence of ionization suppression. In case of a perfect sample preparation combined with the perfect chromatographic separation – leading to the chromatogram where the analyte is completely separated from all matrix components – ionization suppression would not occur and would not have to be considered. Unfortunately, perfect sample preparation methods are not available in most cases. A number of literature sources address choosing the most effective sample preparation method from the matrix effect point of view. In LC-MS solid phase extraction (SPE), liquid-liquid extraction (LLE), precipitation/centrifugation or combinations of these as well as other methods are used for sample preparation.

Different sample preparation techniques have been compared and found that for phenacetin and caffeine determination in endogenous plasma protein precipitation is the least favorable technique for LC-ESI-MS analyses while LLE was the most favorable [ref 30]. Additionally LLE has been found to be more effective sample preparation technique than SPE for methadone determination, because the latter tends to concentrate not only the analyte but also matrix compounds similar to the analyte (i.e. potentially co-eluting from HPLC with the analyte) [ref 30]. The reason probably is that for LLE a larger selection of extracting solvents is available and therefore more freedom in varying selectivity is achievable. On the other hand, in the case of SPE, a solid phase similar to the HPLC stationary phase is often used (often both are low polarity C18 or C8 phases) and therefore little additional/different selectivity is obtained during sample preparation. Additionally, it has been shown that sample pre-concentration may significantly increase ionization suppression.

(b) The main instrumental modification that can be considered is using a non-ESI ion source, such as APCI instead of ESI, since ionization in the APCI source has been demonstrated to be less affected by matrix effects [ref 31, ref 32, ref 33, ref 34]. Switching the ESI source from positive to negative ionization mode or reducing the flow rate of the effluent have also been demonstrated to be efficient in some cases [ref 35]. Unfortunately, there are numerous analytes for which neither APCI nor negative mode ESI are suitable. Furthermore, among the different LC-MS ion sources ESI in general tends to have the lowest limits of detection [ref 36].

(c) The two main LC-method-related matrix effect reduction possibilities are improvement of chromatographic separation, e.g. with ultra-high performance liquid chromatography (UPLC/UHPLC), and sample dilution. Both have been used by numerous authors. Dilution has been shown to significantly reduce the ionization suppression [ref 37]. However, it is often impossible to dilute the sample sufficiently so that ionization suppression will completely disappear, because the analyte concentration may fall below the limit of quantification. In such cases, the so-called extrapolative dilution approach [ref 37] has been found useful, which consists in diluting the sample as far as possible and if the suppression is still present then extrapolating the analyte concentration mathematically to infinite dilution.

The following video presents a practical discussion on these issues:

Accounting for matrix effect

http://www.uttv.ee/naita?id=23308

https://www.youtube.com/watch?v=5lonQCdmcis

Sometimes it is too difficult (and therefore impractical) or impossible to remove all of the matrix effect, therefore, approaches accounting for the matrix effect have also been developed. Most of them fall either into the category of internal standard usage or matrix-matched calibration [ref 38].

matrix_effect_unsolved.xlsx 27 KB
matrix_effect_resolved.xlsx 29 KB
5.4_evaluating_matrix_effect.pdf 361 KB

6. Precision and trueness: some additional aspects

In the preceding sections the concepts of trueness and precision as well as their practical evaluation were explained. This section examines some additional aspects of trueness and precision determination. The experiments for precision and trueness evaluation are often carried out in parallel (simultaneously). If this is not the case, then before trueness experiments, precision of the method must be checked. This is because precision influences evaluation of trueness (or its constituents), but not vice versa.

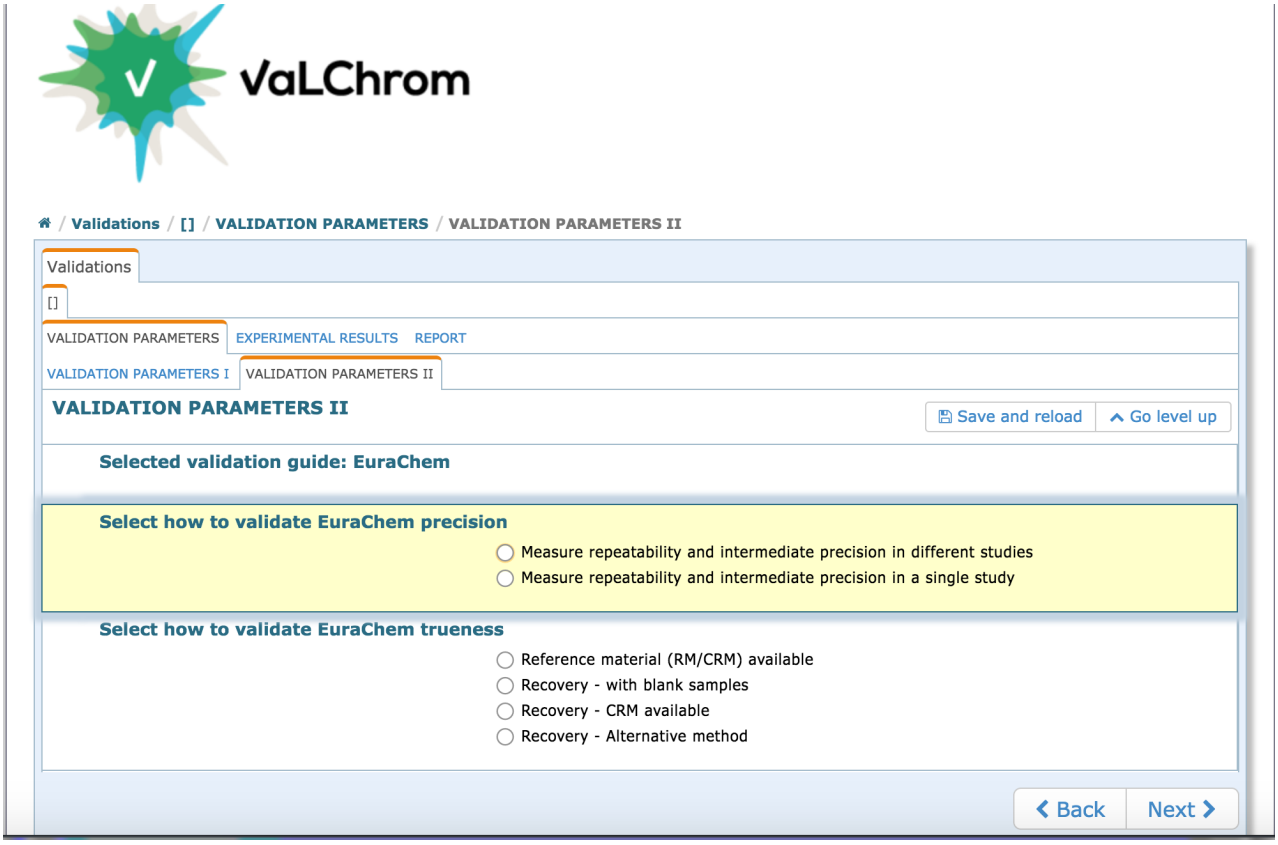
In this section you will find practical information on how to choose the sample types, concentration levels, numbers of replicates and time ranges for precision evaluation (section 6.1), but also how to find reference values for bias calculations (section 6.2). Section 6.3 discusses on how to decide whether or not to correct results for bias.

Practical aspects of precision and trueness: Different guidelines

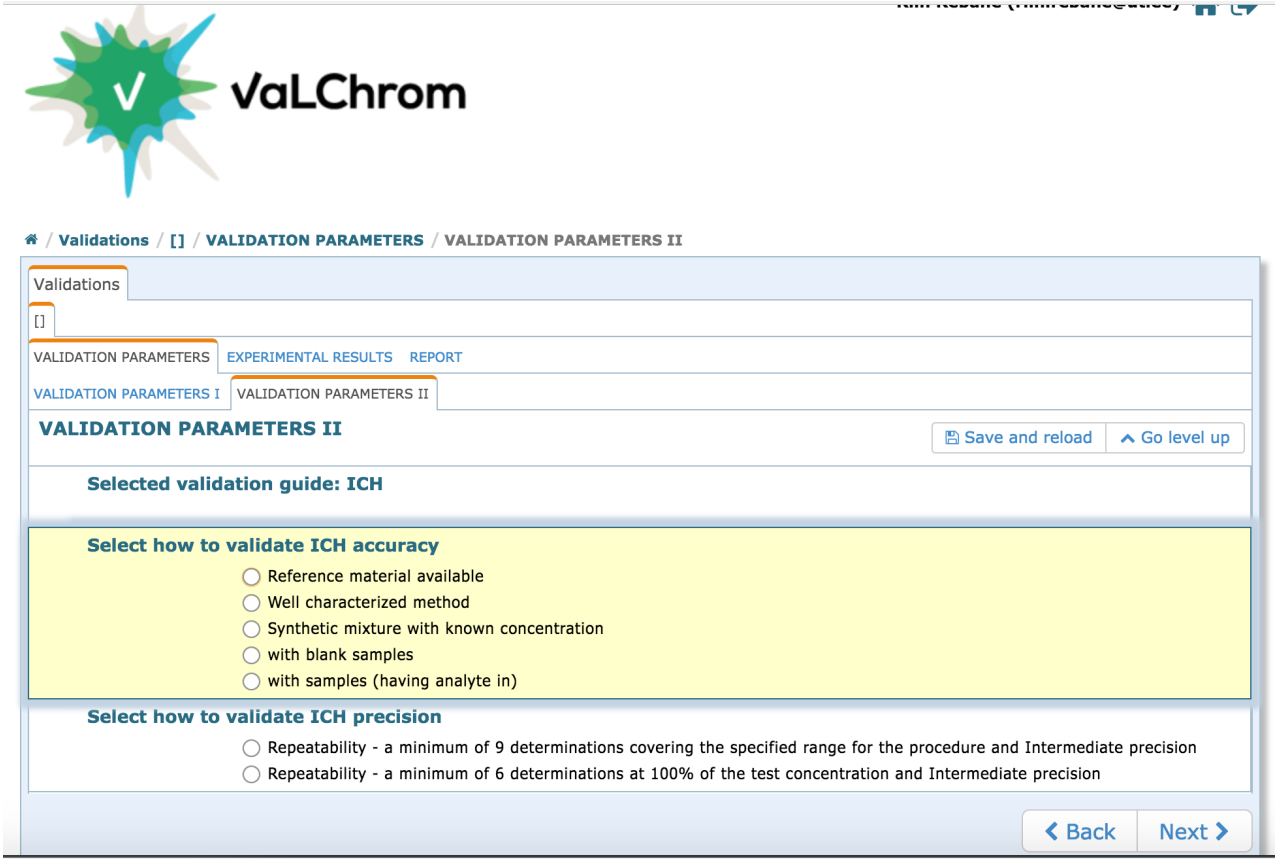
http://www.uttv.ee/naita?id=23633

https://www.youtube.com/watch?v=pObPZYAMdtE

When validating according to specific guidelines, one is offered a selection of way how each parameter is evaluated. Validation software VaLChrom can be used as a guidance. For example, when validating according to EuraChem, it will help in choosing precision and trueness evaluation:



In case of ICH guideline presents choices for accuracy and precision evaluation:



6_practical_aspects.pdf 82 KB

6.1. Choosing sample type, concentration levels, replicates and time range for precision

Most guidelines agree that the used matrices should as well as possible correspond to the matrices encountered in routine analysis. If the precision of the whole method (not just the analytical step) (1) is determined then each replicate determination should be made from a separate subsample and include all steps of the method, including sample preparation.

- The sample matrices should represent those routinely analyzed in the laboratory. The sample(s) can in principle be of any origin: leftovers from large routine samples, leftovers from proficiency testing samples, etc. Accurate reference values are not necessary for precision determination and it is generally not advised to use certified reference materials (CRM) as samples for determining precision since these are often better homogenized than real samples, leading to too optimistic estimates of precision.
- Sufficient amount of sample must be available so that the required number of replicate measurements with subsamples can be carried out.
- The sample should be homogeneous, so that the subsamples would have as similar matrix and analyte concentrations as possible.
- In the case of determining intermediate precision, the sample has to be stable during a longer period of time (generally at least several months).

Planning precision experiment

http://www.uttv.ee/naita?id=23479

https://www.youtube.com/watch?v=L8gwuy41JD0

***

(1) It is almost always useful to determine the precision of the whole method, not just the analytical step. So, throughout this course we generally address the precision of the whole method, unless it is explicitly stated otherwise.

6.2. Obtaining reference value for bias calculation

Three types of samples can be used for obtaining reference value:

1. Certified reference materials (CRMs) with sufficiently small uncertainty of the reference value. CRMs should match both the matrix of the samples routinely analyzed by the validated method as well as the range of the expected concentrations of the analytes in routine samples.
2. Reference materials that do not have to have a certified uncertainty estimate. These can be materials characterized by a reference material producer, but whose values are not accompanied by an uncertainty statement or are otherwise qualified; materials characterized by a manufacturer of the material; materials characterized in the laboratory for use as in-house reference materials; and materials subjected to a restricted proficiency testing. If there is no better possibility then materials distributed in a proficiency test and characterized by consensus values obtained from participant results can also be used, but such consensus values generally have low reliability, leading to unreliable bias estimates. Again, the type of matrix and analyte content range should match that of routinely analyzed samples.
3. When no reference materials are available, bias can be investigated by spiking studies. Sample is split into two aliquots – one is analyzed in its original state and the other is analyzed after a known amount of the analyte has been added. It is important that the spiked analyte must be dispersed into the sample in such a way that its molecules experience the same molecular environment as that of the native analyte. With many matrices (soil, fruits-vegetables, meat, plant leaves, …) this is difficult to achieve. In such cases the molecular interactions experienced by the analyte molecules originating from the spike differ from those experienced by the native analyte. As a result, the spiked analyte molecules behave somewhat differently from the native analyte and the obtained bias value may not accurately reflect the bias operating on the native analyte. In most cases, the native analyte is more strongly bound by the matrix than the spiked analyte, resulting in somewhat optimistic bias estimates from spiking experiments [ref 3, ref 12]. Therefore, bias (recovery) studies by spiking are strongly subject to the observation that while small bias (good recovery) is not a guarantee of good trueness, large bias (poor recovery) is certainly an indication of poor trueness. Strictly speaking, trueness studies of this type only assess bias due to effects influencing the added analyte. The smaller the recovery, i.e. the larger the bias affecting the method, the lower is the trueness of the method.

Some important considerations:

1. It is acceptable to use a particular reference material for only one purpose during a validation study: either for calibration or for evaluation of trueness, but not both at the same time.

2. How good matrix match between CRM and routine sample is good enough? It is not possible to give a universal answer to this question. Depending on the matrix and the analyte(s) the sufficient similarity of matrices may be different and the decision has to be based on the physical and chemical properties of the matrix and the analyte. In the case of food commodity matrices some general directions can be found in SANTE/SANCO [ref 4], which relate to grouping of matrices according to their properties, with the assumption that extraction recovery of an analyte from the same group of matrices should be similar. A comprehensive treatment of the topic of matrix match and usage of surrogate/artificial matrices, together with a large number of examples from the area of bioanalysis (analysis of endogenous compounds) is available in [ref 55].

It is important, however, to note that in the case of LC-MS the situation is additionally complicated by existence of the matrix effect as bias component. This means that the question is not only about extracting the analyte from the sample, but also about possible co-extracting and co-eluting interferences. The following table intends to give some examples of reasoning in such cases. The table is by no means complete and the recommendations given there should be taken with caution, as the table lists only matrices, not analyte-matrix combinations.

Table 1. Example situations of incomplete matrix match between samples and CRMs with comments.

Sample matrix	CRM or RM matrix	Comments	Recommendation
Drinking water	Sea water	Both are homogeneous matrices and drinking water can be regarded as the "limiting case" of sea water by salt content and matrix simplicity.	Sea water CRM can be used for validating a method of drinking water analysis and if only a small bias is detected then the method can be considered good from the bias perspective. If significant bias is detected then the method may in fact still be OK and bias correction certainly should not be applied.
Sea water	Drinking water	Sea water is much richer in components, especially salts, than drinking water and drinking water cannot be used for approximating sea water.	Drinking water CRM should not be used for validation of a method intended for sea water analysis.
A citrus fruit	Another citrus fruit	Citrus fruits are similar by almost all chemical constituents and acidity.	In general citrus fruits can be used interchangeably. However consider that the co-eluting matrix compounds may still be different and matrix effects may vary from one citrus fruit to another. It may be a good idea to determine matrix effects separately.
Apple	Another variety of apple	It turns out that apple varieties can differ significantly between themselves. The differences are not large from the point of view of analyte extraction, but can be large from the point of view of matrix effects (co-eluting compounds) [ref 27].	It is of course logical that if an apple CRM is available then it is used for bias determination. Care should be taken however, and it may be a good idea to determine matrix effects separately, as is explained in section 5.4.
Blood	Blood	Although blood plasma	Again, it is logical that if a blood plasma RM or CRM is available then it is

plasma from an ill patient	plasma from a healthy volunteer	properties of all humans should in principle be similar, it has been found that the blood plasma matrix can differ substantially between individuals [ref 39].	used for bias determination. In addition, since generic blood plasma is easily available for bioanalytical groups and is reasonably homogeneous RMs can conveniently be prepared in the lab. Care should be taken however, as certain specific interferents can be different and therefore the success is greatly dependent on the specific analytes that are determined. It may be a good idea to determine matrix effects separately, as is explained in section 5.4.
Crude rapeseed oil	Refined rapeseed oil	Analyte extraction properties of these two matrices should be very similar, but an important constituent of crude rapeseed oil are phospholipids, which are known to be a major cause of LC-MS matrix effect.	Because of phospholipids in the sample matrix, which are absent in the RM/CRM, every care should be taken, that they will not co-elute with the analyte. Matrix effect profile, as explained in 5.3 can be very useful.

3. It is strongly recommended to determine bias using at least two different reference samples. They can be of the same type or of different types.

Obtaining a reference value

http://www.uttv.ee/naita?id=23478

https://www.youtube.com/watch?v=kURtbQ7ACYs&t=10s

6.3 Avoiding/minimizing bias and correcting for bias

There is an important difference between precision and trueness. Although the repeatability and/or reproducibility standard deviation can be decreased, they cannot be fully eliminated. In contrast, elimination of bias is in principle possible, although care must be exercised. In practice bias correction (1) is frequently done or at least attempted. In very broad terms distinction can be made between avoiding/minimizing bias and correcting for bias.

Minimizing (in favorable cases avoided/eliminated) bias can in the case of LC-MS methods be most effectively done via the isotope-labeled internal standard (ILIS) method. This approach means that at an as early as possible stage of analysis an accurately measured amount of ILIS is added to the sample and instead of just using the analyte signal $S_{Analyte}$ for quantitation, the ratio of analyte and ILIS signals $S_{Analyte}/S_{ILIS}$ is used. If some of the heavy atoms (i.e. not hydrogen atoms) in ILIS are labeled (which is preferred as opposed to replacing hydrogen by deuterium) then the extraction, adsorption, etc properties of analyte and ILIS are almost indistinguishable. As a consequence, extraction losses, retention time, co-eluting interferents and consequently matrix effects for ILIS will be almost indistinguishable from the analyte and will be cancelled out if the above-mentioned ratio will be used for quantification. At the same time the *m/z* values are different, enabling separate MS measurement of the signals $S_{Analyte}$ and S_{ILIS}, even if their retention times are the same. For this reason, ILIS is one of the most powerful approaches of assuring quality of LC-MS results. The two main drawbacks of ILIS are (a) non-availability of ILIS-s for many analytes and (b) difficulties in dispersing ILIS into sample in such a way that it will be in the same molecular environment as the native analyte.

If ILIS is successfully used, then the remaining bias can be negligible, even though process efficiency (PE) and/or recovery (R) may be below 100% (sometimes significantly). In this case PE, R, etc. cannot be called bias constituents anymore. Most of the discussion related to bias in this course is given from a wider point of view and does not assume that ILIS has been used.

Correcting for bias. If ILIS cannot be used then bias can be determined as explained in the previous sections and thereafter results can be corrected for bias. There can in broad terms be three cases with respect to bias correction [ref 22]: (a) correction may be required (e.g. by some regulations), in which case it has to be carried out; (b) correction may be forbidden, in which case it must not be done or (c) correction may be allowed. In the latter case, it is not easy to decide, whether or not to correct and careful consideration is needed. The main reasons for this are:

- In many cases accurate determination of a systematic effect (accurate determination of bias) can involve a very large effort and because of this can be impractical. This means that the determined bias is not accurate (has high uncertainty) and does not enable reliable correction.
- In the case of correcting results with bias, the uncertainty of bias has to be included into the result's uncertainty budget. It can happen that the uncertainty of correction is not much smaller than the uncertainty due to possible bias.
- It is possible that the bias determined using a matrix similar (but not exactly the same) to the sample matrix will be so much different from the real bias that correction makes the result more wrong than it was before correction.

For this reason it is useful to check the following four criteria and carry out bias correction only if all of them are fulfilled [ref 21, ref 22]:

- There is evidence of a significant effect (i.e. no need to correct when bias is close to 1).
- The cause of bias can be identified (i.e. bias correction should never be used for solving problems of unknown origin).
- Bias can be reliably determined for the particular analyte-matrix combination (otherwise the result can become even more wrong!).
- Useful reduction of combined uncertainty is achieved (as compared to inclusion of the possible bias into uncertainty)

Figure 6 in ref 22 presents the decision tree regarding bias correction.

If bias is not eliminated and is not corrected for then the uncertainty due to possible existence of bias has to be taken into account in the uncertainty estimate of the result!

***

(1) In many sources bias correction is termed recovery correction. For reasons outlined in section 5.1 in LC-MS it is reasonable to treat recovery as one of bias constituents.

7. Accuracy

Accuracy - relations to different concepts

http://www.uttv.ee/naita?id=23668

https://www.youtube.com/watch?v=iunClGAivzo

Measurement result **accuracy** indicates its closeness to the true value [ref 6]. Accuracy differs from trueness: accuracy can be used to chracterize an individual result, but trueness always refers to the mean value of a large number of results. Because of that and because accuracy can characterize an individual result, accuracy involves also precision. So, accuracy of a method (i.e. accuracy of the results delivered by the method) is affected by systematic (bias) as well as random (precision) error components and is, therefore, studied as two components: trueness and precision [ref 2].

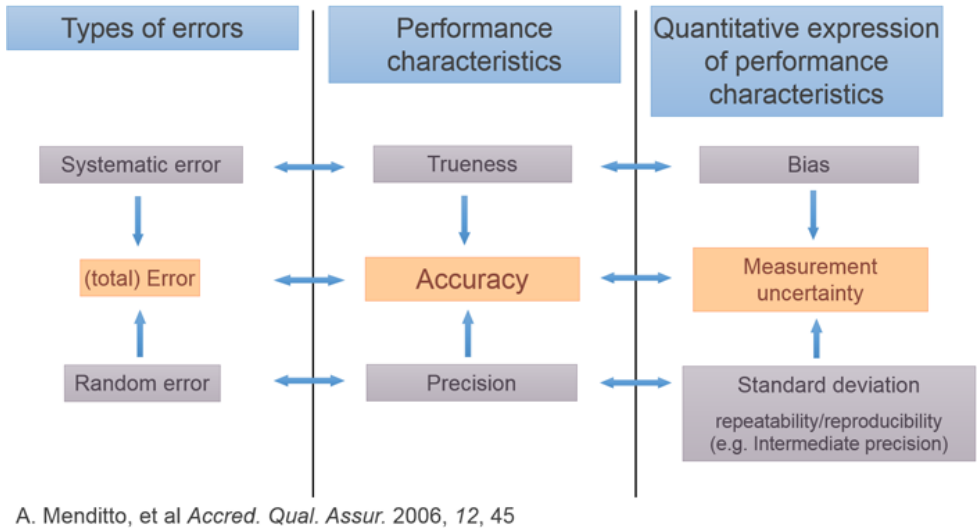


Figure 1. Interrelations between the different error types, the performance characteristics used to estimate them and the ways of expressing the estimates quantitatively. This type of scheme was originally published in ref 56.

Accuracy, trueness, precision and measurement uncertainty

http://www.uttv.ee/naita?id=23345

https://www.youtube.com/watch?v=NfEsN1Gaq5k

A number of regulatory bodies (ICH, FDA, EMA) define accuracy as the degree of agreement between the experimental value, obtained by replicate measurements, and the accepted reference value. This definition is identical to the currently accepted definition of trueness. For the evaluation of acceptability of measurement accuracy, different evaluation criteria can be used: E_n -numbers, *z*-scores or zeta-scores.

Accuracy is often considered a qualitative term [ref 6]. However, in practice it is useful to consider that accuracy is quantitatively expressed as measurement uncertainty. There are different approaches for measurement uncertainty estimation, but in practice the approach based on validation data is often the most convenient. The following video explains the basics of this approach. In-depth treatment of the topic of measurement uncertainty, together with numerous examples and self-tests can be found in the on-line course Estimation of Measurement Uncertainty in Chemical Analysis. Uncertainty estimation on the basis of validation and quality control data is covered in section 10. The Single-lab validation approach.

Measurement uncertainty estimation approaches

http://www.uttv.ee/naita?id=23667

https://www.youtube.com/watch?v=syB2RKAEeMs&t=46s

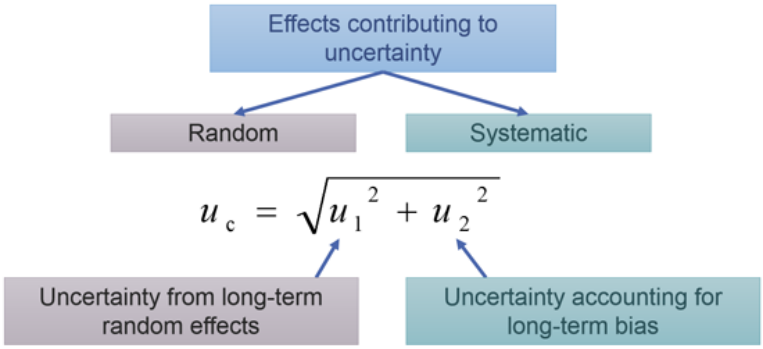


Figure 2. The main idea of estimating measurement uncertainty using validation and quality control data.

7_accuracy_and_measurement_uncertainty.pdf614 KB

8. Stability

In this part of the course the issues related to the analyte's (insufficient) stability are introduced and explained and approaches are proposed how to deal with unstable analytes.

Stability introduction

http://www.uttv.ee/naita?id=23632

https://www.youtube.com/watch?v=1XaPGinpkfQ&feature=youtu.be

Analyte stability is not universally included in validation guidelines as a validation parameter (it is included only in SANTE/SANCO, EMA, FDA and AOAC). The reason is that if the analyte is unstable its decomposition influences the trueness and precision of the procedure and is, thus, accounted for by these two parameters. However, stability is a very important parameter in bioanalytics and this importance justifies handling of it as a separate performance parameter. In bioanalytics the instability of an analyte is rather a rule, not an exception. Therefore, it is necessary to find out whether the analyte is unstable, how unstable it is, could we change conditions in our method to improve the analytes' stability and, finally, how can we still get acceptable results if the analyte is unstable.

Analyte stability must be ensured during **sample collection, processing, storage, extraction and duration of the analysis** to generate reliable (bio)analytical data. Therefore, stability tests can be among the most time-consuming tests in the validation procedure.

Stability is the lowering of analyte content in the sample over the period of time.

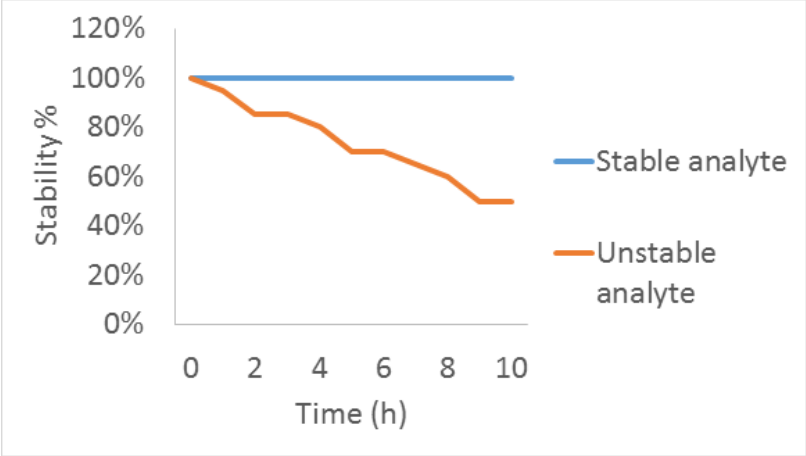


Figure 1. Analyte stability over the period of time.

If the analyte is stable, then the concentration remains the same in time e. g. 100%. If the analyte degrades with time, then its concentration is decreased and also the stability is lower than 100%.

Decomposition usually leads to lowering of the analyte content. In the specific case of analysis of decomposition products decomposition can actually lead to increase of the analyte content.

If the analyte is unstable its decomposition influences the trueness and precision (since both systematic and random effects are usually involved) of the procedure and is, thus, indirectly accounted for by these two parameters. It is nevertheless useful to handle analyte stability separately from trueness and precision.

The rate of decomposition can be strongly dependent on minor experimental details (type of matrix, access of oxygen, temperature, light etc.).

Furthermore, besides the analyte in the samples, analyte in the standards can also decompose. If both occur at the same rate then decomposition only affects precision. If not, then both trueness and precision are affected. In addition, the EMA guide [ref 7] stresses that analyte stability cannot be proven by literature data, further outlining the importance of analyte stability testing.

8_stability_introduction.pdf 40 KB

8.1 Different types of stability

SANTE/SANCO [ref 4] specifies briefly that analyte stability in prepared sample extracts has to be evaluated. Especially in the bioanalytical field the possible analyte decomposition is of very high importance for the quality of the results and therefore deserves special attention. For this reason, the EMA [ref 7], FDA [ref 8] and AOAC [ref 9] validation guidelines specifically address analyte stability as a separate validation parameter.

The FDA guide [ref 8] distinguishes the following types of stability:

- Freeze and Thaw Stability,
- Bench-Top Stability,
- Long-Term Stability,
- Stock Solution Stability,
- Processed Sample Stability.

This guide distinguishes between the analyte stability in calibration and stock solutions and stability in sample matrix and stresses the influence of storage conditions, matrix and container system on stability, besides the intrinsic properties of the analyte itself.

According to EMA [ref 7], stability of the analyte is evaluated using both low- and high-level QC samples. The investigation of stability should cover:

- short-term stability at room temperature or sample processing temperature;
- freeze-thaw stability;
- long-term freezer stability.

Different types of stability

http://www.uttv.ee/naita?id=23669

https://www.youtube.com/watch?v=po-ZhVfSW8o&feature=youtu.be

The emphasis in the EMA [ref 7] guide is not as much on the intrinsic stability of the analyte, as on its stability in the specific matrix.

AOAC [ref 9] is less specific on the experiments that have to be carried out and recommends checking the stability of the stock and initial diluted solutions, stored at room or lower temperatures, by repeating their measurements several days or weeks later.

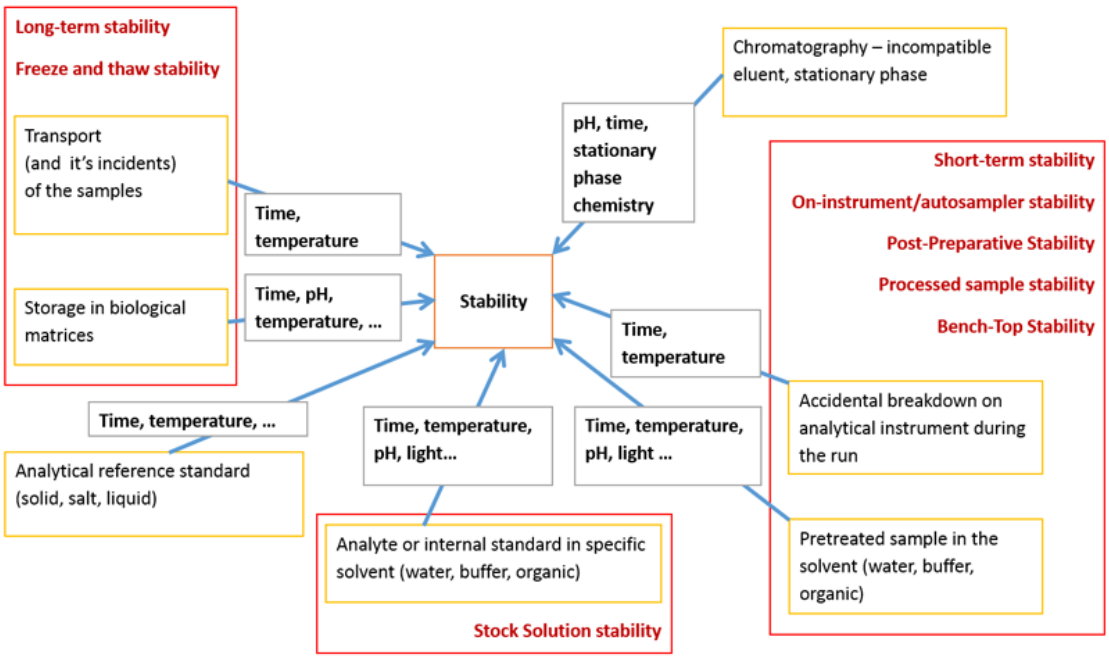


Figure 2. Different types of stability in LC–MS analysis and the chemical/physical parameters influencing stability.

Software  **helps to calculate stability in different conditions.**

8.1_different_types_of_stability.pdf 46 KB

8.2 Examples of unstable analytes

Examples of unstable analytes

http://www.uttv.ee/naita?id=23671

https://www.youtube.com/watch?v=Gn8Wde6lD-w&feature=youtu.be

Table 1. Examples of unstable analytes

Compound	Typical cause of instability
Acyl glucoronide metabolites	Hydrolysis to parent drug, acyl migration [ref 48]
Lactones	Reversible hydrolysis to hydroxyl acid [ref 49]
Prodrugs	Conversion into active drug by ester bond cleavage [ref 50]
Amides	Hydrolysis to acid [ref 51]
Oxidizable compounds	Oxidative degradation (e.g. alcohol to ketone or aldehyde [ref 40], adrenaline to adrenochrome [ref 41] 5-aminosalicylic acid to 5-aminosalicylic quinoneimine [ref 42] cholesterol to 7-beta–hydroperoxy isomer [ref 43] thiols dimerization [ref 44])
Cytostatic nucleosides	Enzymatic deamination [ref 45]
Enantiomers	Racemization [ref 46]
Cis/trans isomers	Cis-trans interconversion [ref 47]
Beta-lactams [ref 10, ref 11] (all penicillin derivatives)	Hydrolysis, enzymatic degradation, oxidation, presence of metal ions

The bench-top/short term stability of a beta-lactam antibiotic – meropenem – in processed blood plasma samples at the room temperature (+20 °C) was investigated during 24 h using overnight injections and after 12 h the concentration of meropenem in the samples was only 17% of the original concentration at both concentration levels 1 µg/mL and 25 µg/mL. After keeping the autosampler at 4 °C and samples at -20 °C until loading into the autosampler, the *ST*% remained around 95% after 3 h and 92 % after 12 h storage in autosampler (+4 °C) [ref 10]. Thus, storage of samples and solutions at low temperature is a powerful means of increasing stability.

As another example [ref 11], the Penicillin G stability was examined over the 16 h time-period and compared in different pH conditions. Penicillin G is well known for its instability in solutions. Stability of penicillins is also highly dependent on the pH of the solution. The bench-top/short term stability of processed blood plasma samples in an autosampler thermostated at +4 °C showed rapid degradation: *ST*% was 20% after 16 h for samples with pH 6.2. After pH adjustment to 7.0 processed blood plasma samples showed significant improvement on *ST*%, resulting in 78% after 16 h at the +4 °C. This example convincingly demonstrates how small changes in conditions can strongly influence analyte stability.

8.2_examples_of_unstable_analytes.pdf 62 KB

8.3 Evaluation of stability

The time during which stability is tested, is important, but is usually not specified in the guidelines, only brief recommendations are given. The reason is that depending on the particular situation, the suitable **duration of the stability study can be very different** and **should be decided by the analyst** based on the situation at hand. Usually when stability of the analyzed sample solutions or standard solutions is evaluated, the injections of the sample solutions and standard solutions are carried out overnight (i.e. the duration of the study is 10-20 h). Testing of (not yet prepared) samples should be carried out during at least a 48 h period, after which the components should be quantified and the terms for storage and conditions that improve the stability over the time, should be identified.

The experimental design of stability testing should take into account the most important parameters - storage **time** and **temperature. Stability results should be used as the indicator to adjust one or the other in order to improve the stability of analyte or internal standard.**

Stability should be studied at least at two concentration levels – low and high concentration level and with matrix, matching the "real-life" matrix.

For doing this the native blank biological matrix should be used and the analyte should be spiked into the matrix at these concentration levels.

Bench-top stability or short-term stability at room temperature or sample processing temperature will indicate analyte stability under sample preparation conditions. Freeze-thaw stability is evaluated usually during three thawing cycles to predict the influence of possible delays and glitches on sample handling.

Stability should be evaluated at different time points and the samples should be analyzed in six replicates [ref 2, ref 7, ref 8].

The analyte or internal standard stability in test or reference solutions, *ST*%, expresses the part of the analyte or internal standard in a sample that does not decompose before the actual LC-MS analysis of the sample. In the absence of decomposition *ST*% = 100%.

Evaluation of stability

http://www.uttv.ee/naita?id=23670

https://www.youtube.com/watch?v=ClUMsRPgmOY&feature=youtu.be

Stability can be evaluated either via chromatographic peak areas (a) or via concentrations (b)

(a) Stability can be evaluated via peak areas as follows:

$$ST\% = \frac{S_t}{S_0} \cdot 100\% \qquad \textbf{(Eq 1)}$$

where S_0 is the initial peak area, determined without introducing any extra pauses in the analysis process; S_t is the peak area obtained when analysis is carried out with making a pause with duration t in the analysis.

(b) Stability can be evaluated via concentrations as follows:

$$ST\% = \frac{C_t}{C_0} \cdot 100\% \qquad \textbf{(Eq 2)}$$

where C_0 is the initial concentration, determined without introducing any extra pauses in the analysis process; C_t is the concentration obtained when analysis is carried out with making a pause with duration t in the analysis.

Stability can be eventually expressed as the average *ST*% value of analyte found in the sample under specific conditions. The freshly prepared calibration standards (at similar concentration levels as the stability study samples) are considered as containing 100% of the initial analyte content.

Software ValChrom **helps to calculate stability in different conditions.**

8.3_evaluation_of_stability_and_overcoming_stability_issues.pdf 52 KB

8.4 Aspects to be considered if analytes are unstable

There are many ways to improve the stability of the analyte. Let us look at some of them.

Table 1. Aspects to be considered while analytes are unstable.

Measure	Challenge
pH control	Extract pH may be crucial and therefore has to be monitored carefully. A possible problem may be esterification of alcohols. If low pH is used for sample preservation then alcohols in the sample may partially be esterified if there are also carboxylic acids present. This can be a problem if some of the analytes are alcohols.
Addition of stabilizers	Some additives are themselves not stable and therefore their concentration and thus also stabilizing efficiency changes in time. Accurate pipetting of stabilizers is required.
Protecting samples and standards from light	Keeping samples and standards on the bench for minimum time possible. Working in the dark or under yellow/red light.
Reducing sample processing time	Time-critical processes have to be managed – e.g. optimization of time during each sample preparation step.
Cold storage and handling of samples and standards	If analyte requires cooling then all steps in the processing have to be carried out under cooling.
Long-term storage in cold (e.g. -70 °C; -80 °C)	Availability of -70 °C or -80 °C freezers.
Derivatization in an as early as possible stage (e.g. already at sample collection)	Time-critical and time-consuming. Achieving completeness of reaction. Different matrix effects when derivatization reagents are used compared to the initial method. Reproducibility issues during the derivatization that can be taken into account with usage of IS or ILIS. Derivatization often increases s_r and s_{RW}. IS and ILIS help reducing s_r and s_{RW}.

Overcoming stability issues

http://www.uttv.ee/naita?id=23685

https://www.youtube.com/watch?v=OgqkQEGVJc0

9. LoD and LoQ

This section addresses the performance parameters that are related to the ability of the method to detect, identify and quantify low analyte levels in samples:

- Limit of detection (LoD) (also called detection limit) – the smallest amount or concentration of analyte in the test sample that can be reliably distinguished from zero [ref 12].
- Decision limit (CC_α) – the concentration level, as determined by the method, at which there is probability α (usually defined as 0.05 or 5%) that a blank sample will give a signal at this level or higher.
- Detection capability (CC_β) – the concentration level of the analyte in sample at which there is probability β (again usually defined as 0.05 or 5%) that the method will give a result lower than CC_α, meaning that the analyte will be declared as undetected.
- Limit of quantitation (LoQ) – the lowest concentration of analyte that can be determined with an acceptable repeatability and trueness.

All section 9 slides are downloadable in one file:

9_lod_loq.pdf 878 KB

9.1. Definitions and important aspects

Limit of Detection

http://www.uttv.ee/naita?id=23290

https://www.youtube.com/watch?v=PnYWCM8lkzs

In this course we define limit of detection (LoD) (also called detection limit) as the smallest amount or concentration of analyte in the test sample that can be reliably distinguished from zero [ref 12]. It should be noted, however, that there are different approaches for defining limit of detection which can cause confusion [ref 12]. One of these approaches (examined more thoroughly in section 9.2) is by defining two separate limits by probability of false positive and false negative errors. To avoid confusion in this course these definitions will be called decision limit (CCα) and detection capability (CCβ).

LoD is used for two purposes:

- To estimate whether the analyte is detected with the used measurement procedure in **the sample** by comparing the LoD value and the result obtained from analysis of the sample.
- To characterize the **analytical method** in terms of its ability of detecting low levels of analytes and compare it to other methods, laboratories or standards.

As can be seen, in the first case LoD is used for making decisions on analyzed samples (whether we can claim that the analyte is detected or not), in the second case LoD is used for making decisions on analytical methods (how low analyte levels is the method able to detect).

In Table 1 (of Section 9.3) approaches to estimate LoD with detailed information are presented. Only widespread approaches, taken from prominent guidelines and articles, are presented, thus it is not an exhaustive list of all approaches. The approaches can be based on different statistical assumptions and therefore the obtained LoD values often cannot be strictly compared. In some cases even results from the same approach should not be compared if important differences exist in the estimation procedure. Choosing a suitable approach to estimate LoD is discussed in following sections.

It is also possible to calculate LoD and LoQ in 

software. This software allows the user to choose the guideline that must be followed, specific estimation approaches that the user wants to apply and to input the data. All calculations (including plotting the residuals) are done automatically.

Not all LoD estimation approaches are fit for all analytical methods. Therefore, as LoD is an important validation parameter, these problems must be understood by the analyst. LoD can then be estimated and used correctly for the particular method in use. Section 9.3 gives guidance relevant for LC-MS methods.

In many cases regulations set maximum allowed limits for analytes (e.g. pesticides) in certain matrices (e.g. different food products). This limit is called the maximum residue limit (MRL).

Two aspects must be considered when carrying out analysis in the framework of a regulation with the aim of establishing whether the analyte level in the sample exceeds the MRL: (a) is the analytical method capable of detecting the presence of the analyte in the matrix at levels lower than MRL, and (b) starting from which *determined* analyte content in the sample we can with high probability conclude that the *actual* analyte content in the sample is over the MRL. To answer the first question the analyst must compare the MRL value to LoD. For example by the European Commission standards the method LoD to estimate cadmium in drinking water should be 10 times lower than the MRL value [ref 13]. The second question can be answered by using decision limit (CC_α), which is discussed further in section 9.2.

Limit of Quantitation

http://www.uttv.ee/naita?id=23629

https://www.youtube.com/watch?v=SHZ4Zik3uFg

Limit of quantitation (LoQ) is defined as the lowest analyte concentration in the sample of analyte that can be determined with an acceptable repeatability and trueness. LoQ can also be called lower limit of quantification, limit of quantification, quantification limit, quantitation limit or limit of determination.
LoQ is determined by the required trueness and precision. Therefore, LoQ is not a limit set by nature and quantitative information at analyte concentrations below the LoQ exists. However, such quantitative information should be used with care as the relative uncertainty of results below LoQ can be too high for further use.

Instrumental LoD and method LoD

http://www.uttv.ee/naita?id=23294

https://www.youtube.com/watch?v=PSHiS56U2bo

Distinction has to be made between instrumental LoD and method LoD. Instrumental LoD is found from the analysis of the analyte in pure solvent. Instrumental LoD therefore only shows the capability of the instrument to detect the analyte and can be used only for comparing different instruments. Method LoD also takes into account the effects that sample preparation and measurement procedure have on the analysis result. The samples used to estimate method LoD must be matrix-matched and must go through the whole sample preparation procedure. Any conclusions related to detection ability of a method have to be made using method LoD (not instrumental LoD). This course addresses only method LoD.

Interpretation of analysis results with LoD and LoQ

http://www.uttv.ee/naita?id=23246

https://www.youtube.com/watch?v=oEIG2cPy5fU

In simplified cases – i.e. if it is not critically important to accurately state whether the analyte can be detected or not (1) – the interpretation of the results obtained with a specific sample should be done in the following way:

- when the analyte cannot be detected or its concentration in the sample is found to be below LoD it can be said that the analyte content is below LoD;
- when the analyte content is above LoD, but below LoQ, the analyte can be said to be present at trace level, but usually no quantitative data are reported;
- when analyte content is at or above LoQ, then the result can be quantified.

It is also important to state the LoD value of the analysis method as a result below LoD does not indicate that there is no analyte in the sample, but only that the used analysis method is not capable of detecting the analyte in this sample. If the analyte is found to be above LoD, but below LoQ, there is more information in the results than is presented by only stating the analyte's presence. These data are not always presented due to the set limits of necessary repeatability and trueness for LoQ. However, it is often still suggested that the measured value and its uncertainty be reported even if the results are below LoQ or LoD, so that further calculations could be made if necessary.

It must be understood that the LoD value might not account for false positive and false negative results. Therefore more correct interpretation of the results can be given with CC_α and CC_β values which are discussed further in the following section.

***

(1) This is in most cases when analysis is performed to establish whether the analyte level in the sample is above or below MRL. In such case, if the analysis result indicates that the analyte content is near LoD then it is by far below MRL and it is thus not very important whether it can be claimed with high confidence that the analyte was detected.

9.2. Decision limit and Detection capability

Decision limit (CC_α) and detection capability (CC_β)

http://www.uttv.ee/naita?id=23306

https://www.youtube.com/watch?v=posQ05DUCIc&t=1s

As mentioned in the previous section, LoD by definition does not take into account both false positive (results where the analyte is declared to be present although actually it is below LoD) and false negative (results where analyte is declared to be below LoD although it is not) errors at the same time.

To account for both of these errors Currie [ref 14] suggested to use two different quantities: decision limit (CC_α) and detection capability (CC_β). CC_α is defined as the concentration level, as determined by the method, at which there is probability α (usually defined as 0.05 or 5%) that a blank sample will give a signal at this level or higher. CC_β is defined as the concentration level of the analyte in sample at which there is probability β (again usually defined as 0.05 or 5%) that the method will give a result lower than CC_α, meaning that the analyte will be declared as undetected (although the analyte content in the sample is in fact higher than CC_α, see Figure 1 for illustration of CCα and CCβ definitions). It must be noted that in some cases the name "limit of detection" is used in place of CC_β. When dealing with LoD the precise definition under question must be clarified.

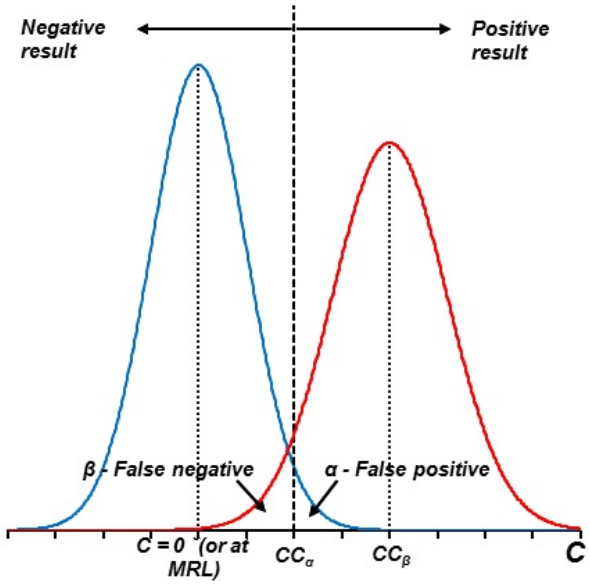


Figure 1. Relation between permitted limit, CC_α and CC_β

Although in most cases LoD can be considered equal to CC_β this is not always the case. The approaches to calculate LoD make many assumptions about the properties of the system – e.g. linearity and scedasticity. The approaches given by standards that suggest using CC_α and CC_β usually provide more complex approaches to estimate them that do not make many of the assumptions made when estimating LoD. The result is that although in principle LoD and CC_β could be considered equal then their values cannot always be compared in a meaningful way.

If an MRL value has been set for the analyte in the particular sample then the following should be considered: (a) the MRL value should be above the CC_β or LoD value (see Section 9.1), and (b) the concentration at which it can be said with known confidence (taking into account both false positive and false negative probabilities) that the analyte level is above MRL should be estimated. For the second part CC_α and CC_β must be evaluated not for 0 concentration but at MRL level (see Figure 1). It should be noted that for example Council Directive 96/23/EC has defined CC_α and CC_β so that if a MRL is set these values are not used to estimate the minimum amount that the method can detect but for estimating when the result is over or under the MRL value. The same approaches to estimate CC_α and CC_β can therefore be used.

It has been shown that often results from very low concentrations and blank samples do not have normal distribution. Moreover, it is more correct to use *t*-distribution rather than normal distribution in case if only a small number of replicate measurements are made. In this example normal distribution is used only to clearly and simply explain the concept of CC_α and CC_β.

Interpreting results with CCα and CCβ

http://www.uttv.ee/naita?id=23349

https://www.youtube.com/watch?v=DP1wAdmXkIg

The results must be interpreted with CC_α and CC_β in the following way: (a) in case the result is below CC_α it must be reported that the analyte concentration in the sample is below CC_β (this is so because it is not known whether our result is falsely negative or truly under CC_α), (b) in case the result is over CC_α, but below CC_β, the sample can be said to contain the analyte with good probability, and (c) in case the value is over CC_β, then the same statement can be made as in the point (b).

However, only the CC_β value can be used for the characterization of the method as at CC_β we take into account the possibility of both false positive and false negative error.

9.3. Estimation of LoD

Different approaches to estimate LoD

http://www.uttv.ee/naita?id=23291

https://www.youtube.com/watch?v=xk2Ou3jaovg

There are a multitude of different approaches that can be used to estimate LoD. LoD approaches from most prominent guidelines can be found in Table 1. These approaches can result in widely varying LoD estimates. Different guidelines often suggest different approaches and it is up to the analyst to choose which approach to use. If a specific approach is not demanded by the guideline this choice must be made based on the necessities and properties of the analytical method.

Table 1. Different approaches for determining LoD, CC_α and CC_β.

Group	Ref.	What is obtained?	Equation
1	[ref 3, ref 8, ref 9, ref 12, ref 15]	LoD (considers false positive and negative results – the probability of false positive and negative values depends on choice of *t*)	$LoD = \bar{y}_0 + t \times S(y)$ **(Eq 1)** $\bar{y}_0$ is mean value of blank samples or 0; t is Student's Coefficient; S(y) is standard deviation of blank or fortified samples. Equation gives LoD in intensity scale.

Read more:

Description: Concentration of fortified samples in LoD range (e.g. lowest level where S/N > 3) or at maximum residue limit (MRL);
t is taken 3 or 4.65;
6 to 10 repeated measurements for blank and fortified samples;
all signal intensities and standard deviations have to be over 0;

Assumptions, simplifications: Homoscedasticity;
normal distribution of replicates;
variability of slope and intercept are not taken into account;
linearity of calibration data;
t value is rounded and does not take into account the degrees of freedom;
only for single sample measurement results.

Notes: Care must be taken when integrating blank samples;
Erroneous calibration function can lead to negative LoD results;
Note that is not necessary (taken as 0) if subtraction with intercept (or with) is done to all results.

2	[ref 16]	LoD essentially equivalent to CC_α (considers only false positive results)	$LoD = t \times S(x)$ **(Eq 2)** S(x) is the standard deviation or pooled standard deviation of analyte concentrations from replicate measurements.

Read more:

Description: A detailed procedure is given to choose fortified sample concentration (incl. estimating an approximate LoD first, measuring only 2 of the needed repeated samples before measuring the rest of the 7 samples);
t is taken depending on degrees of freedom;
Recommended analyte concentration range in fortified samples is 1-5 times LoD.

Assumptions, simplifications: Normal distribution of replicates; variability of slope and intercept are not taken into account;
linearity of calibration data;
heteroscedasticity is somewhat considered by careful choice of fortification concentration;
only for single sample measurement results.

Notes: LoD as equivalent to CC_α (false negative results are not accounted for);
the background (mean of blank values or the intercept value) is subtracted from all other results.
It is then suggested to iteratively check the LoD by estimating it again.

3	[ref 3]	LoD (considers false positive and negative results – the probability of false positive and negative values depends on choice of t)	$LoD = \bar{a} + t \times \frac{S(y)}{\sqrt{n}}$ **(Eq 3)** $LoD = t \times S(y) \times \sqrt{\frac{1}{n} + \frac{1}{n_b}}$ **(Eq 4)** where $\bar{a}$ is the average intercept; n is the number of repeated measurements of the sample; $S(y)$ is standard deviation of blank or fortified samples; n_b is the number of repeated measurements of blank samples. Equations give LoD in intensity scale.

Read more:

Description: Second equation is used if LoD is estimated from single day measurement results and blank values are used for correction;
t is taken as 3.

Assumptions, simplifications: Homoscedasticity; normal distribution of replicates;linearity of calibration data; variability of slope and intercept are not taken into account.
t value is rounded and does not take into account the degrees of freedom.
Allows taking into account the averaging of sample measurement results.

Notes: Using intermediate precision (not repeatability standard deviation) to estimate LoD is suggested.
Monitoring of precision and regular recalculation of LoD values is suggested if LoD is used for making decisions.

4	[ref 17]	LoD (considers false positive and negative results)	$LoD = 3.3 \times \frac{S_d}{b}$ **(Eq 5)** b is the slope of the calibration function, S_d can be chosen as standard deviation of blank samples, residuals ($S_{y.x}$) or intercept. Instructions to calculate standard deviation of residuals and the intercept can be found in the SI.

Read more:

Description: Regression line must be in the range of LoD.
Calibration function is used to estimate slope and standard deviation of residuals and intercept.
Number of repeated measurements not specified.

Assumptions, simplifications: Homoscedasticity;
normal distribution of replicates;
linearity of calibration data; variability of slope and intercept are not taken into account.
If repeated results at each calibration level are averaged and standard deviation of residuals is used for estimate LoD then the number of repeated measurements must be the same as repeated measurements for each calibration level.

Notes: The standard deviation of intercept underestimates the variance of results at 0 concentration and should not be used.
Due to conservative LoD estimates, simple calculation procedure and reasonable workload (S_d is taken from residual values), this is the suggested approach if a rigorous LoD estimate is not needed [ref 19, ref 20].

5	[ref 3, ref 15]	LoD (considers false positive and negative results)	Cut-off approach: number of repeated measurements (usually 10) are made at different concentrations near LoD. The lowest concentration at which all the samples are „detected" is used as the LoD. The detection threshold can be established for example based on S/N, visual evaluation or automatic integration for chromatographic methods.

Read more:

Assumptions, simplifications: Uses robust statistics.

This approach does not assume normal distribution.
Visual evaluation of presence of a peak depends on the analyst.

Notes: This approach is very work intensive;
If repeated estimations of LoD need to be made this approach is not recommended for LC-MS/MS methods;
It has also been suggested to plot the portion of positive responses against concentration to find the lowest concentration at which necessary number of samples give the decision „detected";
Each sample should be independent of the others.

6	[ref 5, ref 15]	CCα and CCβ	**CC_α:** 1. Calculated as $CC_\alpha = \bar{a} + 2.33 \times S_{lab}$ **(Eq 6)**

			$\bar{a}$ is the average intercept and S_{lab} is the within-laboratory reproducibility of the intercept (found from data at and above minimum required limit) 2. Blank matrices are analyzed to estimate noise in the analyte time window. S/N > 3 can be used as CC_α. **CC_β:** 1. Calculated as $$CC_\beta = CC_\alpha + 1.64 \times S_{lab}$$ **(Eq 7)** S_{lab} is the within-laboratory reproducibility of the mean value at CC_α; 2. Calculated as $$CC_\beta = CC_\alpha + 1.64 \times S_{CC\alpha}$$ **(Eq 8)** $S_{CC\alpha}$ is the standard deviation found from 20 repeated sample measurements fortified at CC_α; 3. Lowest concentration level where ≤ 5% of samples are compliant is taken as CC_β. Equations give LoD in intensity scale.

Read more:

Description: Some simple approaches suggested to estimate CC_α and CC_β;
Similarly CC_α and CC_β estimation approaches are suggested in case an MRL is set;
After estimating the intensity value corresponding to CC_α and CC_β calibration function should be used to convert them to the concentration scale;
Approach 2 for estimating CC_α and approach 3 estimating CC_β demand at least 20 replicates (at each level for CC_β).

Assumptions, simplifications: Normal distribution of replicates;
linearity of calibration data; variability of slope and intercept are not taken into account.
Possible heteroscedasticity is considered to some extent: CC_α and CC_β are not found using the same variance.
In these approaches the α value is 1 % and the β value is 5 %.
The coefficients in equations do not take into account the degrees of freedom.

Notes: CC_α and CC_β are found for minimum required performance level or MRL.
Identification requirements have to be followed (only after identification of the analyte can the sample be used for CC_α and CC_β evaluation).

7	[ref 18]	CC_α and CC_β	$$CC_\alpha = t_{0.95}(v)\frac{\hat{\sigma}}{\hat{b}}\sqrt{\frac{1}{K}+\frac{1}{I\times J}+\frac{\bar{x}^2}{s_{xx}}}$$ **(Eq 9)** $$CC_\beta = \delta\frac{\hat{\sigma}}{\hat{b}}\sqrt{\frac{1}{K}+\frac{1}{I\times J}+\frac{\bar{x}^2}{s_{xx}}}$$ **(Eq 10)** $$s_{xx} = J\sum_{i=1}^{I}(x_i-\bar{x})$$ **(Eq 11)** $\hat{b}$ is the estimated slope, $\hat{\sigma}$ is the estimated residual standard deviation, $t_{0.95}$ is the 95% one-sided quantile of *t*-distribution (where $v = I*J - 2$), δ is non-centrality parameter of the non-central *t*-distribution (similar to $t_{0.95}$), *K* is the number of repeated preparations

			of the (unknown) sample, I is the number of calibration levels, J is the number of separate sample preparations at each concentration level, $\bar{x}$ is the mean value of the concentration levels, x_i is the concentration of ith calibration level.

Read more:

Description: Given equations are for homoscedastic data;
iterative approach to estimate CC_α and CC_β, suggested for heteroscedastic data, is also given in the guideline;
Requirements of the approaches:

1. K should equal J
2. I should be at least 3 (5 is recommended)
3. J should be at least 2
4. Number of measurements per sample (L) should be at least 2 and identical for all samples.

The blank measurements are required to also be included in the calibration points.

Assumptions, simplifications: Normal distribution of replicates;
linearity of calibration data;
It is suggested to estimate whether the data are heteroscedastic based on prior knowledge and visual evaluation of the data;
In heteroscedastic approach standard deviation of results is assumed to increase linearly with concentration

Notes: In this guideline the concentration scale is called the net state variable and the intensity scale is called the response variable.
Notice that 2 measurements are recommended for each preparation and the mean of these measurements is then used in the following calculations.

Calculating CC_α and CC_β

http://www.uttv.ee/naita?id=23348

https://www.youtube.com/watch?v=BQ_dOEMDoDs

Estimating CC_α and CC_β

The approaches that are usually suggested to estimate CC_α and CC_β are more complex than the approaches suggested for LoD. This is so because their definition is statistically more rigorous (demanding a known probability level of false positive and negative results) but the results are also more reliable. Some approaches to estimate CC_α and CC_β suggested in guidelines and articles can be found in Table 1.

The CC_α and CC_β calculations [ref 18] take into account the standard deviation of the used linear regression line parameters (slope and intercept). This variance is propagated into the concentration values that are calculated by using these parameters. As CC_α and CC_β are used in concentration scale (similarly to LoD) the variance of slope and intercept must be taken into account when estimating them.

Another property that must be taken into account is homo- and heteroscedasticity. Homoscedasticity means that the variance of signal is constant in case the concentration changes and heteroscedasticity therefore means that the variance changes with concentration. Analytical methods are often heteroscedastic – the signal variance increases as the concentration of the sample increases. If it is shown that the calibration data collected is heteroscedastic then for being fully rigorous, weighted linear regression (WLS) should be used in order to take more accurately into account the variance of the slope and intercept. A simplified approach that usually works sufficiently well is presented below.

With WLS the propagated errors of slope and intercept to the concentration value significantly decrease at lower concentration levels. Therefore the CC_α and CC_β values are also significantly influenced. Using WLS can be complex and a possibility to avoid this is to select a narrow concentration range at lower concentrations from the calibration data that can be shown to be reasonably homoscedastic. These data can then be used to estimate slope and intercept with ordinary linear regression (OLS) which assumes that the data are homoscedastic. As a result calculating the CC_α and CC_β estimates also becomes more simple.

Important aspects of estimating LoD and CCα, CCβ

http://www.uttv.ee/naita?id=23350

https://www.youtube.com/watch?v=9GFMa0AYkdA

It should be considered how important the LoD value for a given analytical method is. Based on this knowledge it can be chosen whether a simple approach to estimate LoD is enough or a more complex approach that makes less assumptions (e.g. about homoscedasticity) and therefore gives more accurate results should be used. The important assumptions made by different approaches are summarized in Table 1. Further details about how to evaluate whether these assumptions can be made is discussed in the following references [ref 19, ref 20]. If the analyte concentration will never come close to LoD value then LoD does not have to be estimated at all. However, often LoD is still estimated in these cases just for proving that the samples are significantly above the LoD of the method. For example when measuring calcium in milk by complexonometric titration then we do not have to worry that in some samples the concentration of calcium might be so low that it would be below LoD for a reasonable titration procedure. However, if the LoD estimate is an important parameter used to interpret the results of analysis more complex but accurate approaches must be used to estimate LoD. For example, when analyzing blood samples of athletes for doping that are not allowed to be used the method must interpret the results correctly even if only very small amounts of analyte is detected. Therefore CC_α and CC_β values estimated with complex approaches that make less assumptions (e.g. ISO [ref 18]) must be used.

In some cases the analytical method can have properties that do not allow the use of some LoD estimation approaches. For example it can be difficult to estimate the standard deviation of the blank for LC-MS/MS methods as the noise can be zero due to signal processing. As the blank values all give intensity of 0 the LoD value cannot be calculated from them but the standard deviation at 0 can be still estimated by other approaches: from the standard deviation of intercept value or from standard deviation of residuals. A more thorough discussion about the problems of processing chromatograms of samples at low concentrations can be found in the following references [ref 19, ref 20]. In conclusion, in general the analyst must understand which approaches cannot be used for a given analytical method.

It should always be kept in mind that LoD is an estimated value and never represents the true LoD as it is calculated from parameters that deviate randomly from their true value between measurements. Moreover, the true value around which the results deviates can change randomly between days. For example the slope of the LC-MS/MS changes significantly between days – this means that the true intensity value given by a concentration changes between days. For this reason the within-day standard deviation is lower than the standard deviation of results collected on different days (see section 4.1). Therefore LoD also changes between days. To take this fact into account LoD should be estimated over a long period of time (e.g. a month) and the median LoD value can then be used [ref 18]. If it can be seen that the LoD estimate changes significantly between days (meaning the variation of LoD value within a day is significantly smaller than between days) and the estimate is important for the correct interpretation of results on that day then LoD should be estimated on that day and that value should be used for the interpretation. However it can also be noted here that if LoD is used only for simple characterization of the method and not used further (see above) then LoD does not have to be estimated on multiple days. It must also be noted that the previous discussion also applies for CC_α and CC_β.

As the different approaches can give differently biased values it should be always stated which approach is used to evaluate LoD. If different approaches are used (to characterize the lab or the new method) then the comparison should be made with caution.

A different concept for estimating LoD is by using the signal-to-noise ratio (S/N). This approach is mostly used in chromatographic methods. Modern chromatography programs determine this value automatically. The signal value for this is found from the height of the peak and noise values are found from either the standard deviation of the noise or from so called peak-to-peak value (meaning the difference between the highest and lowest points in the noise). From this it can be seen that S/N can be found for only one measurement of a sample. A single measurement however does not take into account the variability between measurements and therefore LoD should not be evaluated from this result. A scheme has been suggested by Eurachem where 10 samples are measured on different concentration levels and the lowest concentration where all 10 are detected is taken as LoD. Here the decision that an analyte has been detected can be made from the fact that the S/N is equal to or over 3. However, this means that many measurements have to be made to estimate LoD and due to S/N being conceptually different from other approaches it will be difficult to compare the LoD estimates found with other approaches.

Calculating LoD

http://www.uttv.ee/naita?id=24440

https://www.youtube.com/watch?v=u7LCGkFuUFE

9.4. Estimation of LoQ

Different approaches to estimate LoQ

http://www.uttv.ee/naita?id=23684

https://www.youtube.com/watch?v=G8avqFKe0ds

Approaches to estimate LoQ can generally be divided into two groups: (a) based on the estimation of trueness and precision (or uncertainty) at different concentration levels, or (b) based on similar approaches that are used to estimate LoD. However, in the case when a specific guideline must be followed then LoQ must be estimated by following the guideline.

First let's discuss approaches to determine LoQ by estimating precision and trueness of the method at multiple concentration levels. LoQ can then be taken as the lowest concentration where these parameters are fit for purpose (set by e.g. the analyst or client) or meet the requirements of the necessary standards or guidelines. For example, FDA requires that the intensities at LoQ must have precision of 20% and trueness of 80–120%, and SANTE/SANCO requires that mean recovery (recovery is meant here as trueness) is in range of 70–120% and relative standard deviation (which indicates precision) of at most 20%.

See section 4, 5 and 6 for estimating trueness and precision and section 7 for further information about uncertainty.

Moreover, it has been suggested that LoQ should be found by expressing the uncertainty of the measurement as a function of concentration and comparing the results to the uncertainty levels demanded of that method.

Although the definitions of LoD and LoQ differ significantly, the second group of approaches to estimate LoQ use similar approaches that are used for estimating LoD. However, in case of LoQ a greater multiplication coefficient *k* (used in place of Student's Coefficient) in the equation is used or other higher demands are set on the same parameters. For example, S/N value of at least 10 is required at the LoQ concentration level.

The following equation can be used to estimate LoQ:

$$LoQ = \overline{Y} + k \times S(Y) \quad \textbf{(Eq 1)}$$

Here the same variables can be used as with LoD (see Table 1) and the same experimental design (e.g. concentration levels and number of replicate measurements) is required. However, the coefficient k is required to have values of 5, 6 or 10 (depending on standard or guideline that is being followed). In case of the ICH approach (using the calibration function to estimate the standard deviation) the following equation can be used [ref 17]:

$$LoQ = \frac{10 \times S(cal)}{Slope} \quad \textbf{(Eq 2)}$$

Here again all the variables can be taken from the same datasets for both LoD and LoQ. Visual evaluation can also be used to estimate LoQ: the LoQ is taken as the lowest concentration level where the analyte can be quantified with acceptable level of precision and trueness. It has also been suggested that LoQ can be found by simply multiplying LoD by 2.

LoQ is determined in most approaches from the same data as LoD or is based on LoD and therefore in principle the same issues occur. However, these approaches can over- or underestimate the LoQ because they do not look at the trueness and precision values at specific concentrations. It is assumed that the properties of analytical methods are relatively similar and above the defined point over LoD the accuracy of the results will be fit for purpose.

We recommend using the first group of approaches: precision and trueness should be estimated at different concentration levels and the lowest level where both parameters are in agreement with the demanded value is taken as LoQ. Although labor-intensive, this approach estimates LoQ by its exact definition.

However, if LoQ is not critically important, then using the approach given by ICH (Eq 2, using the standard deviation of residuals) is suggested.

It has also been stated that the LoQ value can be taken as the lowest concentration level used for calibration. This can be convenient if the LoQ of the analytical method value is significantly below the working range. In this case extra measurements do not have to be made at lower concentrations specifically for estimating LoQ. However, it should still be shown, that the method is capable of fit-for-purpose trueness and precision at the lowest concentration level.

About some important aspects of LoQ

The aspects of LoD and LoQ estimation are often similar.

LoQ is also used for two purposes: (a) to determine whether the measured concentration of the sample is in the range that allows it to be quantified with fit for purpose accuracy with the given analytical method, and (b) to characterize the analytical method.

When estimating LoQ, the data used to estimate it should be in range of LoQ. If a calibration function is used to estimate it, then linearity and scedasticity have similar effects on LoQ estimation as on LoD estimation.

The samples used to estimate LoQ should be matrix matched.

As LoD, trueness and precision vary between days (and between measurement series due to random variance of results), LoQ values also have to be estimated on different days.

Due to variance between days, we recommend that LoQ should be determined 5 times over a longer period and the most conservative result should be stated as the methods' performance level to increase its reliability. Moreover, the methods' performance at the LoQ level can be monitored with regular analysis of samples (either real contaminated samples or spiked blank samples) with concentrations close to LoQ.

The exact way of determining LoQ should be specified as with LoD due to the differences of the results when different approaches are used. The difference in LoQ estimates can, however, also come from the fact that different precision and trueness limits have been set.

Calculating LoQ

http://www.uttv.ee/naita?id=24441

https://www.youtube.com/watch?v=DXiGL72twow

10. Ruggedness, robustness

In this part of the course the robustness and ruggedness are introduced and explained.

The terms robustness and ruggedness refer to the ability of an analytical method to remain unaffected by small variations in method parameters (mobile phase composition, column age, column temperature, etc.) and influential environmental factors (room temperature, air humidity, etc.) and characterize its reliability during normal usage.

The notion of remaining unaffected by varying a method parameter has two possible interpretations – it can be interpreted as:

(a) no change of the detected amount of the analyte in a certain sample in spite of the variation of the method parameter or

(b) no change of the critical performance characteristics (e.g. limit of quantitation) by the variation of the method parameter.

In experimental evaluation of robustness either one of these interpretations can be used.

Robustness and ruggedness definitions in the guidelines as well as review articles are very similar. Some guidelines use the term robustness and some use ruggedness. When used together these terms are treated as synonyms in most cases.

Robustness and ruggedness introduction

http://www.uttv.ee/naita?id=23686

https://www.youtube.com/watch?v=8Fpo71pUTR4&feature=youtu.be

These definitions refer to changes made to the method **within the same laboratory.** However, robustness can also be described as the feasibility to reproduce the analytical method in different laboratories or under different circumstances without the occurrence of unexpected differences in the obtained results or performance parameters.

Along the similar lines it has been suggested that **ruggedness** should be used as a parameter evaluating constancy of the results when external factors such as analyst, laboratory, instrument, reagents and days are varied and **robustness** should be used as a parameter characterizing the stability of the method with respect to variations of the internal factors of the method. This involves the parameters related to sample preparation, mobile phase composition, mobile phase flow rate, injection volume, column temperature etc. In addition, an important aspect of robustness is stability of the method against variability of sample matrix.

There are studies where the terms robustness/ruggedness are misinterpreted and actually decision threshold, detection capability or measurement uncertainty is evaluated.

In this course we use the term **robustness** for expressing the **stability of the method against small variations of the intrinsic method parameters** (section 10.1) **and variability of sample matrix** (section 10.2).

In this course we use the term **ruggedness** for **expressing the stability of the method against extraneous influencing factors**.

We address changes of the method parameters (i.e. within-laboratory assessment of robustness) and variability of sample matrices. We do not explicitly address changes occurring when a method is transferred from one laboratory to another.

10_robustness_ruggedness_introduction.pdf 52 KB

10.1 Robustness and ruggedness relation to LC-MS method development

Evaluation of ruggedness and robustness is very important in the case of LC-MS, as there are a large number of system parameters, some of which are difficult to control. As a result, some of the key performance characteristics are also difficult to control or are sensitive to small changes in system parameters (or sample properties), resulting in poor instrument reproducibility between the runs (between samples with formally identical matrix).

Furthermore, LC-MS is very often used for the determination of very low levels of analytes in highly complex matrices. Analysis of complex matrices often requires complex multi-step sample preparation procedures, which contributes to the overall complexity of the method.

Robustness and ruggedness relation to LC-MS method development

http://www.uttv.ee/naita?id=23687

https://www.youtube.com/watch?v=qM0R4OGYI5E&feature=youtu.be

As to which parameters should be varied during robustness studies, the guidelines vary in their recommendations [ref 1]. In this course we recommend the following list of method parameters and changes in these parameters, that should be investigated during robustness studies:

Table 1. Recommended method parameters to be investigated during robustness studies.

Parameter (1)	Likelihood of uncontrollable change (2)	Possible extent of variation (3)	Comments
Liquid chromatography			
Mobile phase pH	Medium	± 0.5 units	pH will have a strong effect on retention time (and possibly resolution) if the analyte's pK_a value is within ± 1.5 units of the mobile phase pH
Concentration of additives in eluent	Medium	± 10% (relative)	Salts, ion-pair reagents, modifiers can suppress/enhance analyte's ionization in the ion source and change its retention time and possibly resolution from other compounds
Organic solvent content in the eluent	Low to Medium	± 2% (relative)	Organic solvent content influences retention time (and possibly resolution) and analyte signal in LC-MS
Column temperature	Low	± 5 °C	Column temperature influences the retention time (and possibly resolution)
Eluent flow rate	Low	± 20%	Eluent flow rate influences the retention time (and possibly resolution)
Column batch and age	Medium	-	Changes in column can influence the retention time (and possibly resolution) and peak shape. There can be also influence on ionization: because of different retention time there can be different co-eluting compounds and, consequently, different ionization suppression/enhancement.
Samples and sample preparation			
Analyte extraction time; solvent amount and composition (in liquid/liquid and solid phase extraction, etc)	High	± 20%	Influences recovery and LoQ/LoD
Injection solvent composition	Low/High	± 10% (relative)	This is the solvent in which analyte is taken up during the last stage of sample preparation. This composition can influence retention time and recovery and therefore also the matrix effect (in broad sense), LoQ/LoD and stability. The effect can be very serious in the case of UHPLC (4) and is usually not that critical in the conventional HPLC.
Matrix effect in broad sense (sample matrix source)	High	6 different sources	Can be assessed under selectivity studies. Influences trueness (recovery and ionization suppression), LoQ/LoD
Mass spectrometry			
Drying gas temp	Low	± 10 °C	Drying gas temperature can influence analyte ionization efficiency in the ion source
Nebulizer gas pressure/flow rate	Low	± 5 psi / ± 1 L/min	Nebulizer gas pressure/flow rate can influence analyte ionization efficiency in the ion source
Ion source configuration	High (if configurations	According to the ion source	Ion source configuration can influence spray and ionization efficiency in the ion source

(Nebulizer position)	can vary) Not applicable (if fixed source)	design. Should be varied if source is used in different configurations.	
Ion source condition (Nebulizer aging, ion source contamination)	High	After analysis of samples versus cleaned system	Contamination can spontaneously accumulate when analyzing a series of samples.

***

(1) This list is not exhaustive, there can be different other parameters that, depending on situation, could be changed.

(2) This grading is generic, corresponding to most common situations. In specific cases the grading can be different.

(3) These extents of parameter variation can be considered as possible defaults. In practice different extents can be used, but they should be realistic.

(4) In the case of UHPLC the smaller changes in the injection solvent have stronger influence, especially if larger injection volume is used.

10.1_robustness_and_ruggedness_relation_to_lc-ms_method_development.pdf 57 KB

10.2 Matrix robustness

The term robustness is in most cases understood in terms of influence of variations of method parameters on results. Our experience suggests, however, that an additional dimension – robustness in terms of variability of sample matrix – is beneficial: different matrices can lead to different matrix effects (either in the narrow or broad sense) and thus different responses and different results with the same analyte concentration in the sample.

On the example of blood plasma: depending on the personal variations in metabolism, diet, possible diseases, e.g. the composition (first of all but not limited to the content of proteins, phospholipids and polyunsaturated fatty acids) of blood plasma can vary significantly, even though formally the matrix is the same – blood plasma [ref 24].

The sample preparation procedure that is suitable for blood plasma of low protein or phospholipid content may give different results for blood plasma with high protein or phospholipid content. This effect is closely related to and is occasionally addressed under selectivity in validation guidelines.

However, the possible effects of this kind of variability are not limited to the loss of selectivity, but can also influence recovery (and hence trueness), ionization suppression/enhancement as well as limit of detection (LoD)/limit of quantitation (LoQ). It is thus useful to investigate the effect of sample matrix variability (in the case of formally identical matrices) more broadly than just for selectivity.

Matrix robustness

http://www.uttv.ee/naita?id=23733

https://www.youtube.com/watch?v=v5aj8PMbQMY&feature=youtu.be

10.2_matrix_robustness.pdf 36 KB

10.3 Different ways to evaluate robustness

In the lecture 10.1 Robustness and ruggedness relation to LC-MS method development we saw different LC-MS parameters that influence robustness and ruggedness, as well as the influence of these parameters to method and recommended changes. The influence of these parameters on the method and recommended changes in these parameters. As a reminder, there was a list of LC parameters, sample and sample preparation parameters and mass spectrometry parameters.

Because of the very large number of potentially variable parameters it is reasonable to divide assessment of ruggedness into separate parts. A very logical division would be to test ruggedness separately for sample preparation and for the LC-MS analytical part.

Before starting investigation of robustness it is crucial to find out, what are the most important performance characteristics of the method.

For example, if method's LoQ is very close to the LoQ required by legislation, then changes in LoQ value have to be monitored against small changes in method parameters. The most influential method parameters influencing LoQ could be MS parameters, mobile phase pH and sample preparation parameters.

The main criteria for choosing parameters are (a) how much a given method parameter can influence the critical characteristic and (b) how likely it is that this parameter will change uncontrollably.

If those parameters are chosen, then we should use one of two options to evaluate method robustness and ruggedness – to use the experimental design or the One Variable At a Time approach.

Different ways to evaluate robustness

http://www.uttv.ee/naita?id=23735

https://www.youtube.com/watch?v=U1nchnq8TZE&feature=youtu.be

In most cases experiments with one-by-one variations (One Variable At a Time approach) of the most important parameters are carried out. Experimental design approaches are somewhat less used, especially at routine laboratories, because these approaches require knowledge and experience with mathematical statistics. In this course we will give an overview of both – One Variable At a Time approach and the Experimental Design approach.

Based on the common practice in literature and on our own experience we recommend the following:

1. Change parameters one by one (One Variable At a Time approach) in both directions from the nominal (optimal) value. Changes in the parameters should be realistic in the context of normal use of the method.
2. "Do not stop there!" Often parameters may be mutually unrelated (uncorrelated), but in some cases this does not hold. For example: change in mobile phase pH can decrease resolution between two adjacent peaks. Likewise, increase of mobile phase flow rate can also lead to decrease of resolution. While separately either of these two changes can still lead to insignificant loss of resolution, their occurrence together may lead to peak overlap. Whether this is the case, can often be determined by educated inspection of the effects of the changes (without additional experiments) and noting potential problems.
3. Effects from the change of parameters should be recorded and if necessary, graphical or statistical analysis of the effects should be done.
4. Regarding the robustness tests results, if necessary, measures to improve the performance of the method should be taken.

10.3_different_ways_to_evaluate_robustness.pdf 46 KB

10.4 Experimental design

Experimental design

http://www.uttv.ee/naita?id=23346

https://www.youtube.com/watch?v=XLsVkO-wlF4

In addition to the One Variable At a Time approach (explained in 10.3) robustness can be estimated by varying multiple parameters simultaneously. Such approaches are strongly recommended if it is expected that different method parameters have *interactions*: meaning that the value of one parameter influences the optimal value of another parameter (e.g. retention behaviour of acidic compounds in RP is influenced by pH, but the effect of pH is strongly influenced by the organic solvent content). These effects are also called non-linear effects. Such effects can remain unnoticed with the One Variable At a Time approaches.

A most straightforward way to study this is called *full factorial design*, carried out as in the following example:

a) Decide all parameter levels. For example in case of studying injector influence on trueness you could wish to test autosampler temperature and sample drawing speed. Usually your autosampler is thermostated at 20 °C and the drawing speed is 10 µL/s. So you decide to test if anything changes when the autosampler temperature is 18 °C or 22 °C and if drawing speed is 5 µL/s or 15 µL/s.

b) Construct a measurement plan that would contain all possible combinations of the parameter levels:

Experiment	Parameter 1	Parameter 2	Temperature (°C)	Drawing speed (µL/s)
1	+	+	22	15
2	+	-	22	5
3	-	+	18	15
4	-	-	18	5

So for 2 parameters to be studied at two levels there are 4 combinations.

c) Carry out analysis with all of the parameter value combinations in random order. If possible make replicates, otherwise you can use repeatability determined previously for comparison.

d) Calculate the effects and compare with random variability (repeatability). For obtaining the parameter effects the results obtained at the higher parameter value need to be averaged and compared to the averaged results obtained on lower parameter value.

Experiment	Parameter 1	Parameter 2	Temperature (°C)	Drawing speed (µL/s)	Trueness
1	+	+	22	15	89%
2	+	-	22	5	87%
3	-	+	18	15	100%
4	-	-	18	5	101%

In this case for example:

$$Temperature\ effect = \frac{Trueness_1 + Trueness_2}{2} - \frac{Trueness_3 + Trueness_4}{2} = -12.5\% \quad \textbf{(Eq 1)}$$

Also the combined effect of the parameters is worth studying. For this the product or the two parameters should be viewed:

Experiment	Parameter 1	Parameter 2	Product	Temperature (°C)	Drawing speed (µL/s)	Trueness
1	+	+	+	22	15	89%
2	+	-	-	22	5	87%
3	-	+	-	18	15	100%
4	-	-	+	18	5	101%

$$Combined\ effect = \frac{Trueness_1 + Trueness_4}{2} - \frac{Trueness_2 + Trueness_3}{2} = 1.5\% \quad \textbf{(Eq 2)}$$

e) Comparing the effect obtained with repeatability. The relative repeatability at similar concentration level for this method has been previously determined to be 3.7%. Therefore we can see that the temperature effect is significant and autosampler temperature definitely needs to be accurately controlled for this method. However the combined effect of autosampler temperature and sample drawing speed is not significant.

The main practical difficulty in applying the full factorial design approach is the large number of combinations if many parameters are varied simultaneously. There are, however, also numerous reduced designs available to do this kind of studies, which can be used even if the number of parameters is very high. Also the calculations can easily be made with different computer programs, which automatically assess the statistical significance.

10.4_experimental_design.pdf 2.06 MB

VaLChrom

Introduction

VaLChrom is an on-line software tool for validation of chromatographic analysis methods, which implements validation guidelines from 7 international organizations (NordVal, EMEA, FDA, ICH, AOAC, EuraChem, IUPAC). VaLChrom lets the user choose validation guideline, specify validation parameters and ways of assessment – thereby assisting in creation of validation plan. After performing planned experiments, user enters the experimental results into VaLChrom, which calculates required validation parameters. In addition to conventional validation parameters, VaLChrom also assists in measurement uncertainty assessment. See VaLChrom homepage for more information.

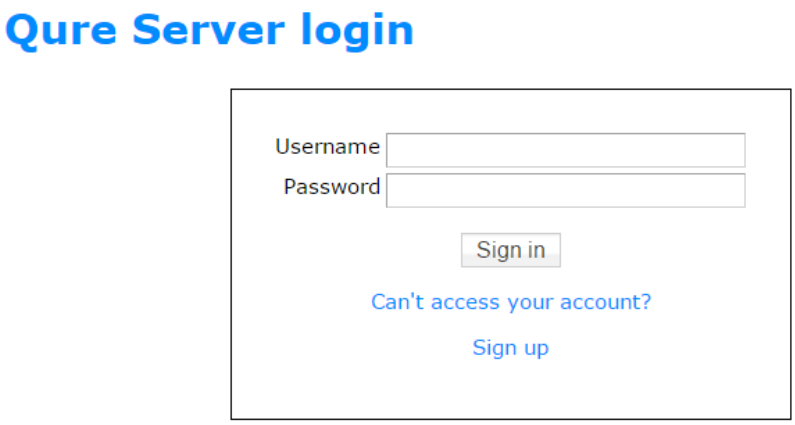
Link to VaLChrom.

VaLChrom is available at https://mooc.valchrom.ut.ee/

First steps in VaLChrom

Signing up/signing in/lost password.

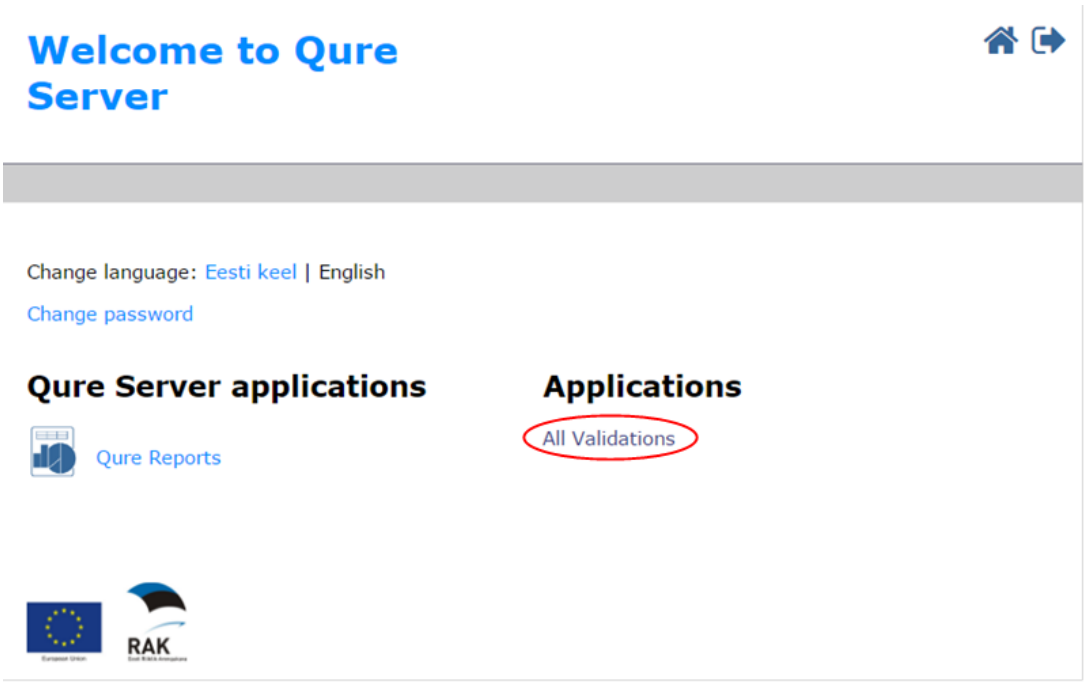
Navigate to https://mooc.valchrom.ut.ee/



On your first visit use "Sign up" link to register as a user. Please enter your functional e-mail address, as the login information is sent by e-mail.

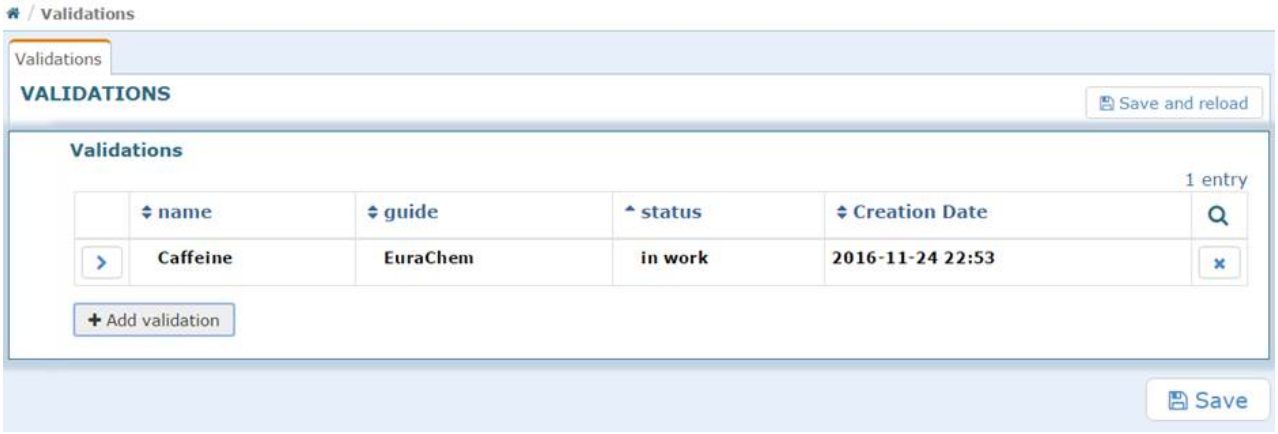
Use your e-mail address as "Username" and enter your password to sign into the system. If you forget your password, use "Can't access your account?" link.

When signed in to VaLChrom, click "All validation" (see the image below) to proceed to validation planning.



Opening page [Validations].

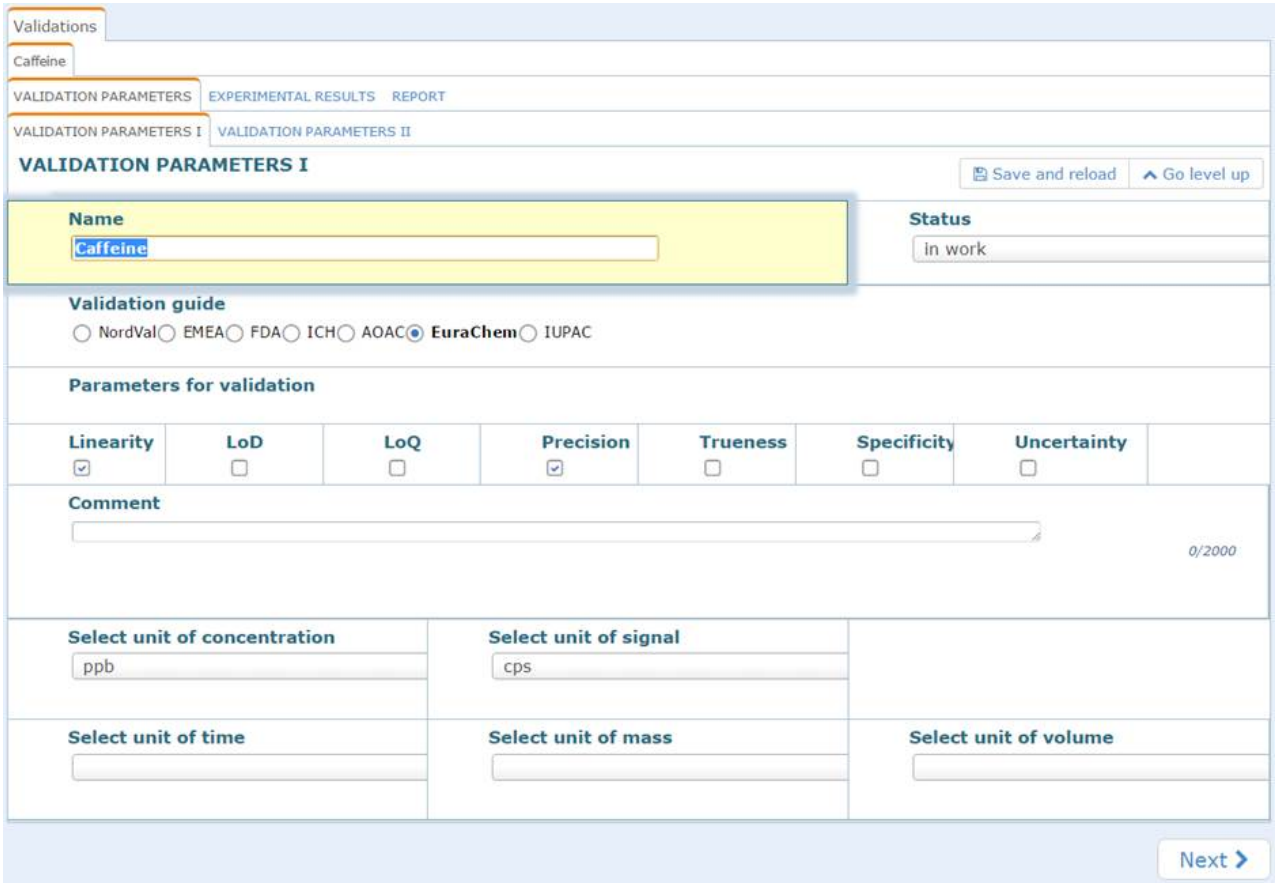
The opening page lists all your validations and allows to add new validations (using "+ Add validation" button).



Use the ">" button to open existing validation for editing and "x" button to delete the validation.

Adding new or editing existing validation.

When adding new validation or opening existing validation for editing, the following screen appears. Required fields are marked accordingly.



"Name" – name of the validation; appears on the opening screen.

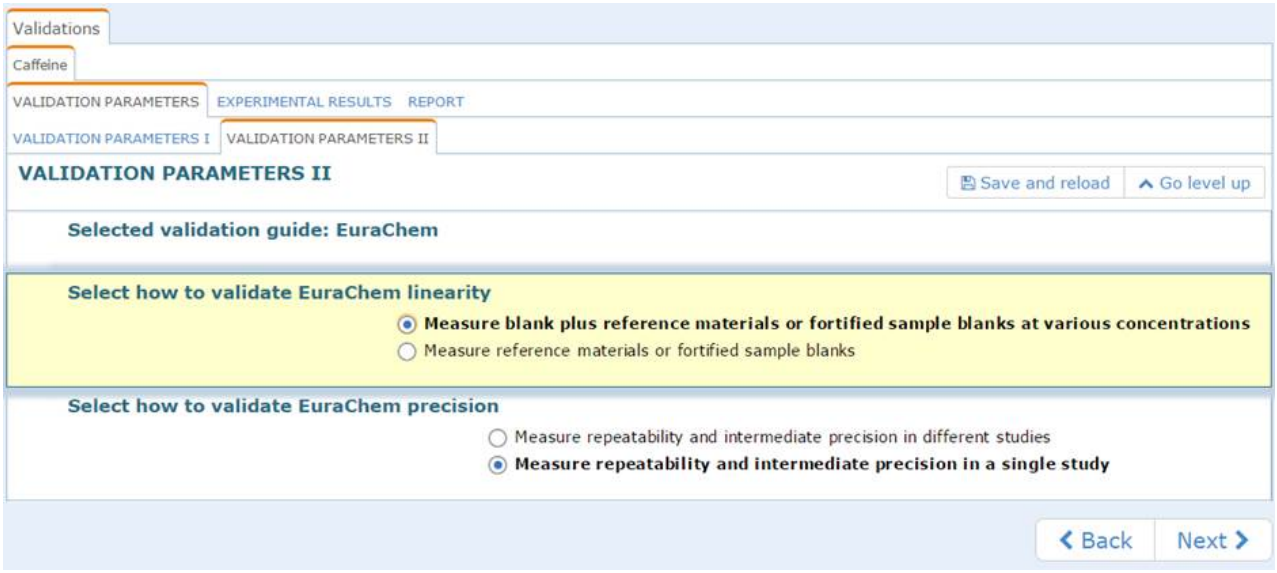
"Validation guide" – validation guide to use in the project.

"Parameter for validation" – select parameters to validate.

Note that concentration and signal units must also be selected.

To save changes and continue, press "Next" button. This brings you to …

"Validation parameters II" pane.



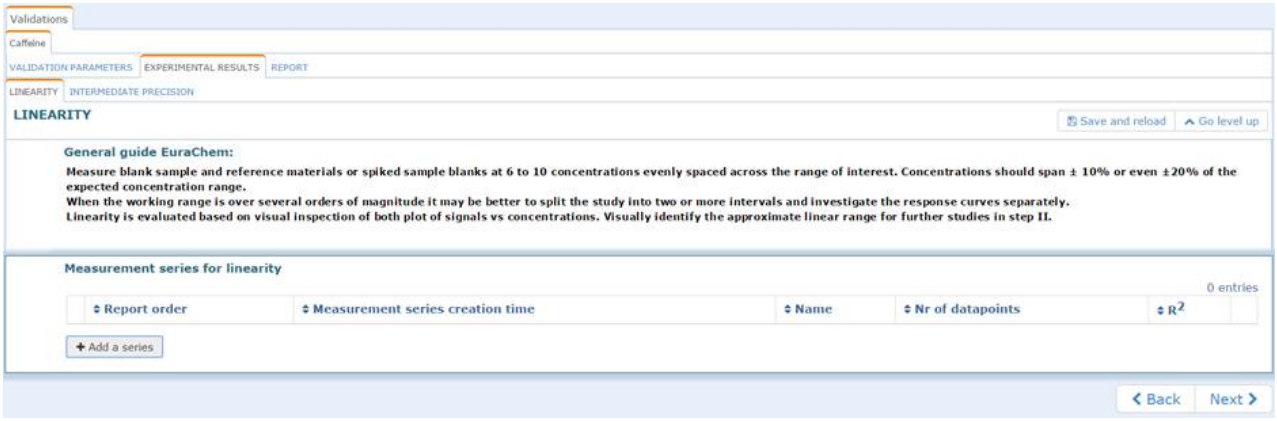
As demonstrated in the picture, EuraChem guidelines offers two ways for evaluation of linearity and precision. Here you can choose specific ways for evaluation of validation parameters.

Pressing "Next" button brings you to …

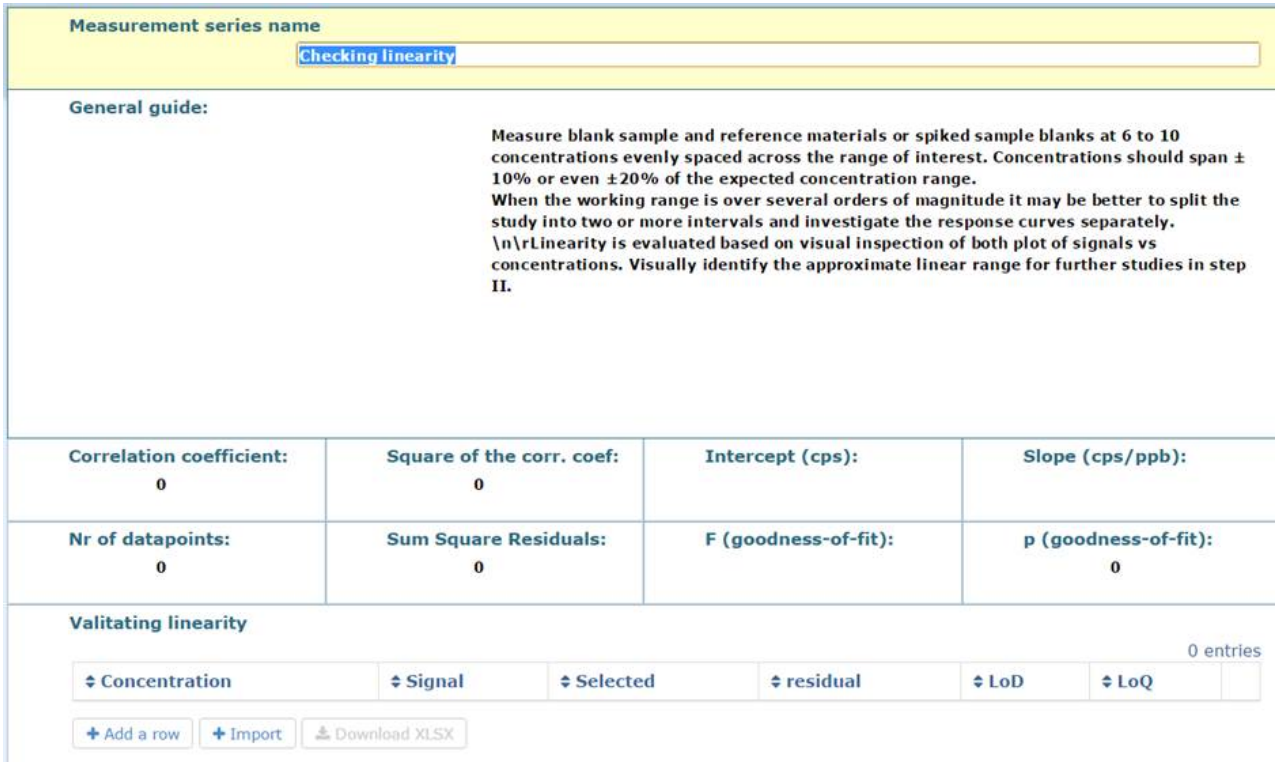
"Experimental results" pane.

Here guidance is given, which experiments to carry out for evaluation of the validation parameter. Depending on the validation parameter the experimental data may be entered as series or row-by-row by directly filling in the fields.

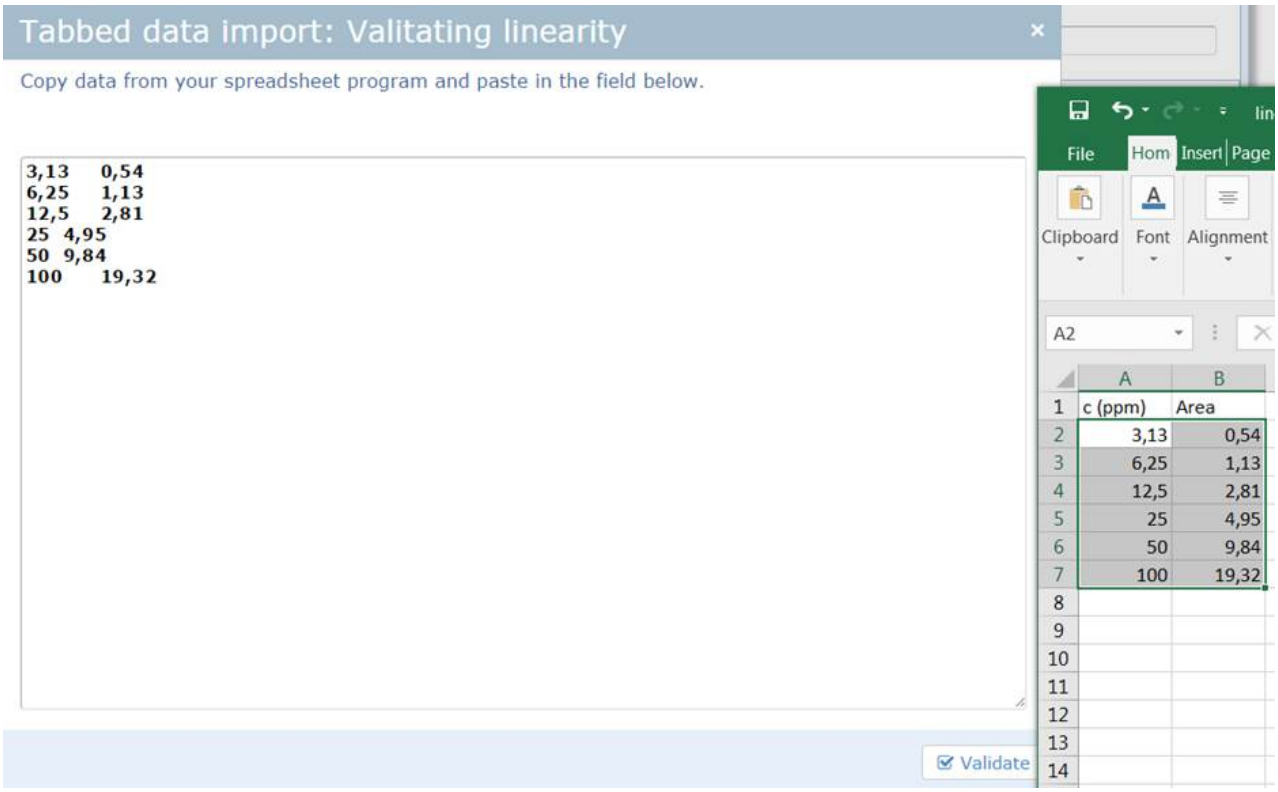
Entering experimental data as series.



Press "+ Add a series" button to add new series.



Data within the series can be entered row by row using "+Add a row" button, but it's easier to "+Import" data. The following picture illustrates how data was copy-pasted from spreadsheet program into VaLChrom.



Clicking "Validate" button checks whether the data can be accepted by the system.

Tabbed data import: Valitating linearity

Concentration	Signal	Selected	residual	LoD	LoQ
3,13	0,54				
6,25	1,13				
12,5	2,81				
25	4,95				
50	9,84				
100	19,32				

Clear Save

Click "Save" to complete importing the data. The following picture illustrates that the data was imported and parameters characterizing linearity were calculated.

Correlation coefficient:	Square of the corr. coef:	Intercept (cps):	Slope (cps/ppb):
0.999684	0.999369	0.101149	0.192925
Nr of datapoints:	**Sum Square Residuals:**	**F (goodness-of-fit):**	**p (goodness-of-fit):**
6	0.161572	6330.183322	0

Valitating linearity

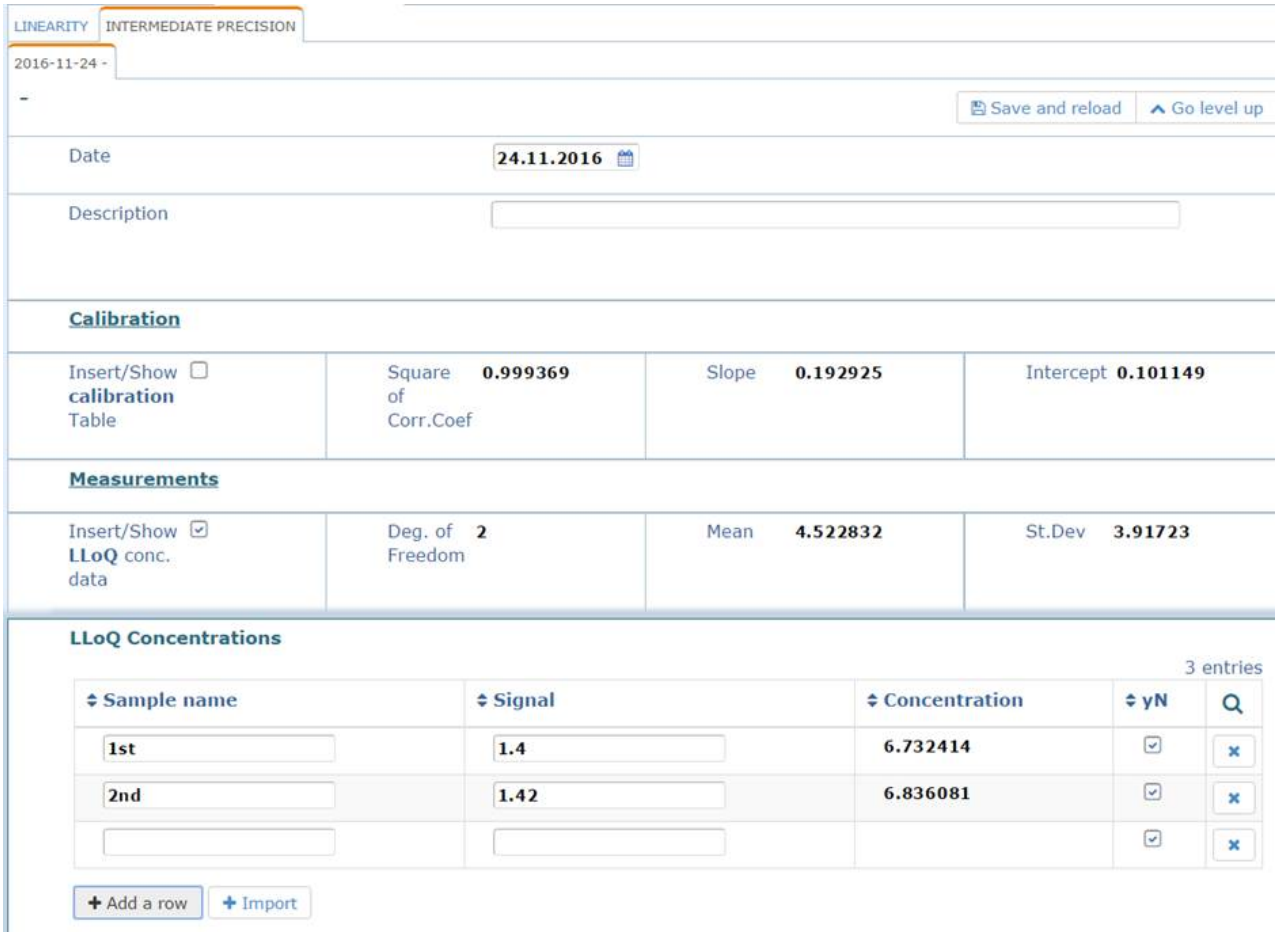
6 entries

Concentration	Signal	Selected	residual	LoD	LoQ	
3.13	0.54	☑	-0.165004	☑	☑	×
6.25	1.13	☑	-0.17693	☑	☑	×
12.5	2.81	☑	0.297288	☑	☑	×
25	4.95	☑	0.025726	☑	☑	×
50	9.84	☑	0.092601	☑	☑	×
100	19.32	☑	-0.073649	☑	☑	×

+ Add a row + Import Download XLSX

Entering experimental data row-by-row.

Data in "Intermediate precision" pane can also be imported (as was done with linearity data), but the picture below demonstrates row-by-row approach. In this example, calibration data was imported, but is hidden by unchecking "Insert/Show calibration table" box.



Calculation results in VaLChrom.

VaLChrom calculates results automatically and displays usually above the data matrix. Graphical results (calibration graphs, residuals plots) are displayed if respective links are clicked. You may discover that sometimes results are not automatically recalculated after editing data. Pressing "Save and reload" button resolves this issue.

Acknowledgements

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 602453.

Clonidine for Sedation of Paediatric Patients in the Intensive Care unit

The validation software VaLChrom was developed at University of Tartu as an EU Project (Development of software for validation of chromatographic methods, No. 3.2.1201.13-0020).

VaLChrom

Euroopa Liit
Euroopa Regionaalarengu Fond

Eesti tuleviku heaks

Dr. Karin Kipper has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° 608765 and her work is also supported by PUTJD22 grant from the Estonian Research Council.

WHRI ACADEMY

MARIE CURIE ACTIONS

Eesti Teadusagentuur
Estonian Research Council

References

ref 1

A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio, I. Leito. Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part I. *Anal. Chim. Acta* **2015**, 870, 29-44

ref 2

A. Kruve, R. Rebane, K. Kipper, M.-L. Oldekop, H. Evard, K. Herodes, P. Ravio, I. Leito. Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part II. *Anal. Chim. Acta* **2015**, 870, 8-28

ref 3

B. Magnusson and U. Örnemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods - A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014).

ref 4

SANTE/11945/2015 (Until 01.01.2016: SANCO/12571/2013), Guidance document on analytical quality control and method validation procedures for pesticides residues analysis in food and feed. European Commission, 2015.

ref 5

European Commission Decision 2002/657/EC implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Off. J. Eur. Commun. L221 (2002) 8-36.

ref 6

JCGM 200:2008, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*, 3rd edition. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, 2008.

ref 7

Guidance on bioanalytical method validation, European Medicines Agency, 2011.

ref 8

U.S. Department of Health and Human Services Food and Drug Administration, Guidance for Industry: Bioanalytical Method Validation, Draft Guidance, U.S. Department of Health and Human Services Food and Drug Administration, 2013.

ref 9

AOAC, AOAC Guidelines for Single-laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals, Official Methods of Analysis, 19th ed., AOAC, INTERNATIONAL, Gaithersburg, MD, 2012 Appendix K.

ref 10

Rapid determination of meropenem in biological fluids by LC: comparison of various methods for sample preparation and investigation of meropenem stability. K. Kipper, K. Anier, I. Leito, J. Karjagin, K. Oselin, K. Herodes, *Chromatographia*, **2009**, 70, 1423–1427.

ref 11

Bioanalytical UPLC-MS/MS method development and validation for measuring penicillins in human blood plasma – analyte stability issues. K. Kipper, C. Barker, A. Johnston. *Mass spectrometry: Applications to the Clinical Lab (MSACL 2015). 8th Annual Conference & Exhibits*, **2015**, San Diego, CA, USA, 28.03.2015-01.04.2015.

ref 12

Harmonized Guidelines for single-laboratory validation of method of analyses (IUPAC Technical Report). M. Thompson, S.L.R. Ellison, R. Wood, *Pure Appl. Chem.* **2002**, 74(5), 835–855.

ref 13

Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption, *Off. J. Eur. Commun.* **2014**, L221, 1-101.

ref 14

Limits for qualitative detection and quantitative determination. L. A. Currie, *Anal. Chem.* **1968,** 40, 586–593.

ref 15

NordVal Protocol No. 2, Guide in validation of alternative proprietary chemical methods, **2010**.

ref 16

40 CFR Part 136, Appendix B to Part 136 - Definition and Procedure for the Determination of the Method Detection Limit-Revision 1.11, US Environmental Protection Agency **2012**

ref 17

ICH harmonized tripartite guideline: validation ofanalytical procedures: text and methodology Q2(R1), International Conference of harmonization of technical requirements for registration of pharmaceuticals for human use **2005**.

ref 18

ISO 11843-2:2000 Capability of detection - Part 2: Methodology in the linear calibration case, International Organization of Standardization **2000**.

ref 19

H. Evard, A. Kruve, I. Leito, Tutorial on estimating the limit of detection using LC-MS analysis, part I: Theoretical review. *Anal. Chim. Acta* **2016**, 942, 23-39.

ref 20

H. Evard, A. Kruve, I. Leito, Tutorial on estimating the limit of detection using LC-MS analysis, part II: Practical aspects. *Anal. Chim. Acta* **2016**, 942, 40-49.

ref 21

B. Magnusson, S. L. R. Ellison *Anal. Bioanal. Chem*. **2008**, 390, 201–213

ref 22

E. Theodorsson, B. Magnusson, I. Leito, *Bioanalysis* **2014**, 6, 2855–2875

ref 23

R. Rosen *False Positives in LC-MS/MS: to what Extant Do We Have to Live with Them?* Presented at IsrAnalytica, Tel Aviv, 2010.

ref 24

P. Bastos-Amador, F. Royo, E. Gonzalez, J. Conde-Vancells, L. Palomo-Diez, F.E. Borras, J.M. Falcon-Perez, Proteomic analysis of microvesicles from plasma of healthy donors reveals high individual variability, *J. Proteomics* **2012**, 75, 3574–3584.

ref 25

P.J. Taylor "Matrix effects: The Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry" *J Chromatogr A* 38 **2005**, 328-334.

ref 26

S.J. Lehotay, K. Mastovska, A.R. Lightfield, R.A. Gates, Comparison of QuEChERS sample preparation methods for the analysis of pesticide residues in fruits and vegetables, J. AOAC Int. 93 (2010) 355–367.

ref 27

A. Kruve, A. Künnapas, K. Herodes , I. Leito, Matrix effects in pesticide multi-residue analysis by liquid chromatography–mass spectrometry, *J. Chromatogr. A.* 1187, 1-2, **2008**, 58-66

ref 28

I.R. Pizzutti, A. de Kok, M. Hiemstra, C. Wickert, O.D. Prestes, Method validation and comparison of acetonitrile and acetone extraction for the analysis of 169 pesticides in soya grain by liquid chromatography–tandem mass spectrometry, J. Chromatogr. A 1216 (2009) 4539–4552

ref 29

A. Kruve, K. Herodes, I. Leito, Electrospray ionization matrix effect as an uncertainty source in HPLC/ESI-MS pesticide residue analysis, Rapid Commun. Mass Spectrom. 93 (1) (2010) 306–314.

ref 30

R. Bonfiglio, R.C. King, T.V. Olah, K. Merkle, The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds, Rapid Commun. Mass Spectrom. 13 (1999) 1175–1185.

ref 31

R. Dams, M.A. Huestis, W.E. Lambert, C.M. Murphy, Matrix effect in bio-analysis of illicit drugs with LC-MS/MS: influence of ionization type, sample preparation, and biofluid, J. Am. Soc. Mass Spectrom. 14 (2003) 1290–1294.

ref 32

P. Keski-Rahkonen, K. Huhtinen, R. Desai, D.T. Harwood, D.J. Handelsman, M. Poutanen, S. Auriola, LC–MS analysis of estradiol in human serum and endometrial tissue: comparison of electrospray ionization, atmospheric pressure chemical ionization and atmospheric pressure photoionization, J. Mass Spectrom. 48 (2013) 1050–1058.

ref 33

O.A. Ismaiel, M.S. Halquist, M.Y. Elmamly, A. Shalaby, H.T. Karnes, Monitoring phospholipids for assessment of ion enhancement and ion suppression in ESI and APCI LC/MS/MS for chlorpheniramine in human plasma and the importance of multiple source matrix effect evaluations, J. Chromatogr. B 875 (2008) 333–343.

ref 34

E. Beltrán, M. Ibánez, J.V. Sancho, F. Hernández, Determination of patulin in apple and derived products by UHPLC-MS/MS. Study of matrix effects with atmospheric pressure ionisation sources, Food Chem. 142 (2014) 400–407.

ref 35

E.M. Thurman, I. Ferrer, D. Barcelo, Choosing between atmospheric pressure chemical ionization and electrospray ionization interfaces for the HPLC/MS analysis of pesticides, Anal. Chem. 73 (2001) 5441–5449.

ref 36

Asko Laaniste, *Comparison and optimisation of novel mass spectrometry ionisation sources*, PhD dissertation, University of Tartu, 2016

ref 37

A. Kruve, I. Leito, K. Herodes, Combating matrix effects in LC/ESI/MS: The extrapolative dilution approach, Anal. Chim. Acta, 651, 1, 2009, 75-80

ref 38

A. Kruve, I. Leito, Comparison of different methods aiming to account for/overcome matrix effects in LC/ESI/MS on the example of pesticide analyses, *Anal. Methods*, 2013, **5**, 3035-3044

ref 39

P. Bastos-Amador, F. Royo, E. Gonzalez, J. Conde-Vancells, L. Palomo-Diez, F.E. Borras, J.M. Falcon-Perez, Proteomic analysis of microvesicles from plasma of healthy donors reveals high individual variability, *J Proteomics*. **2012**, 75, 12, 3574-84

ref 40

Q. Sun, W. Zhang, W. Zhong, X. Sun, Z. Zhou, Pharmacological inhibition of NOX4 ameliorates alcohol-induced liver injury in mice through improving oxidative stress and mitochondrial function. *Biochim Biophys Acta.* **2016**, 1861, 2912-2921.

ref 41

V.M. Costa, R. Silva, L.M. Ferreira, P.S. Branco, F. Carvalho, M.L. Bastos, R.A. Carvalho, M. Carvalho, F. Remião, Oxidation process of adrenaline in freshly isolated rat cardiomyocytes: formation of adrenochrome, quinoproteins, and GSH adduct. *Chem Res Toxicol*. **2007**, 20, 1183-1191.

ref 42

R.K. Palsmeier, D.M. Radzik, C.E. Lunte. Investigation of the degradation mechanism of 5-aminosalicylic acid in aqueous solution. *Pharm Res.* **1992**, 9, 933-938.

ref 43

P.G. Geiger, W. Korytowski, F. Lin, A.W. Girotti, Lipid Peroxidation in Photodynamically Stressed Mammalian Cells: Use of Cholesterol Hydroperoxides as Mechanistic Reporters. *Free Radic Biol Med*. **1997**, 23, 57-68.

ref 44

A. Beuve, Thiol-Based Redox Modulation of Soluble Guanylyl Cyclase, the Nitric Oxide Receptor. *Antioxid Redox Signal*. **2016** Apr 1.

ref 45

J.V. Voorde, S. Sabuncuoğlu, S. Noppen, A. Hofer, F. Ranjbarian, S. Fieuws, J. Balzarini, S. Liekens, Nucleoside-catabolizing enzymes in mycoplasma-infected tumor cell cultures compromise the cytostatic activity of the anticancer drug gemcitabine. *J Biol Chem*. **2014**, 289, 13054-13065.

ref 46

M.M. Musa, R.S. Phillips, M. Laivenieks, C. Vieille, M. Takahashi, S.M. Hamdan, Racemization of enantiopure secondary alcohols by Thermoanaerobacter ethanolicus secondary alcohol dehydrogenase. *Org Biomol Chem*. **2013**, 11, 2911-2915.

ref 47

R.J. Guan, Y. Xiang, X.L. He, C.G. Wang, M. Wang, Y. Zhang, E.J. Sundberg, D.C. Wang, Structural mechanism governing cis and trans isomeric states and an intramolecular switch for cis/trans isomerization of a non-proline peptide bond observed in crystal structures of scorpion toxins. *J Mol Biol*. **2004**, 341, 1189-1204.

ref 48

S. Feng, M.A. ElSohly, D.T. Duckworth, Hydrolysis of conjugated metabolites of buprenorphine. I. The quantitative enzymatic hydrolysis of buprenorphine-3-beta-D-glucuronide in human urine. *J Anal Toxicol*. **2001**, 25, 589-593.

ref 49

M. Caswell, G.L. Schmir, Formation and hydrolysis of lactones of phenolic acids. *J. Am. Chem. Soc*., **1980**, 102, 4815–4821.

ref 50

Y. Yang, H. Aloysius, D. Inoyama, Y. Chen, L. Hu. Enzyme-mediated hydrolytic activation of prodrugs. *Acta Pharm. Sin. B*. **2011**, 1, 143–159.

ref 51

X. Yan, J. Wang, Y. Sun, J. Zhu, S. Wu. Facilitating the evolution of Esterase Activity from a Promiscuous Enzyme Mhg with Catalytic Functions of Amide Hydrolysis and Carboxylic Acid Perhydrolysis by Engineering the Substrate Entrance Tunnel. *Appl Environ Microbiol*. **2016**. doi: 10.1128/AEM.01817-16

ref 52

IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook.

ref 53

H. Sun, F. Wang, L. Ai, C. Guo, R. Chen, Validated method for determination of eight banned nitroimidazole residues in natural casings by LC/MS/MS with solid-phase extraction, *J. AOAC Int*. **2009**, 92, 612–621.

ref 54

R.P. Lopes, D.V. Augusti, L.F. de Souza, F.A. Santos, J.A. Lima, E.A. Vargas, R. Augusti, Development and validation (according to the 2002/657/EC regulation) of a method to quantify sulfonamides in porcine liver by fast partition at very low temperature and LC–MS/MS, *Anal. Methods*, **2011**, 3 606–613.

ref 55

R. Thakare, Y. S. Chhonker, N. Gautam, J. A. Alamoudi, Y. Alnouti, Quantitation of a novel metalloporphyrin drug in plasma by atomic absorption spectroscopy. *J. Pharm. Biomed. Anal*., **2016**, 128, 426–437

ref 56

A. Menditto, M. Patriarca, B. Magnusson, Understanding the meaning of accuracy, trueness and precision. *Accred Qual Assur*, **2007**, 12, 1, 45. doi:10.1007/s00769-006-0191-z

ref 57

K.K. Murray, R.K. Boyd, M.N. Eberlin, G.J. Langley, L. Li, Y. Naito. Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure Appl. Chem* **2013**, 85, 7, 1515–1609.

Glossary

AOAC

Association of Official Analytical Chemists

Accuracy

Measurement result **accuracy** indicates its closeness to the true value.

Analytical run

or simply run:

Definition 1: A set of samples that are analysed in one batch, during short time. In LC and LC-MS this term typically refers to an automatized sequential analysis of a set of samples, calibrants and QC samples that have been loaded into autosampler.

Definition 2 (defined by Clinical Laboratory Improvement Amendments (CLIA)): An interval (i.e., a period of time or series of measurements) within which the accuracy and precision of the measuring system is expected to be stable. In laboratory operations, control samples are analyzed during each analytical run to evaluate method performance, therefore the analytical run defines the interval (period of time or number of specimens) between evaluations of control results. Between quality control evaluations, events may occur causing the measurement process to be susceptible to variations that are important to detect.

Bias

Quantitative estimate of trueness, can be found as measured value difference from the reference value.

CCα

Decision limit: the concentration level, as determined by the method, at which there is probability α (usually defined as 0.05 or 5%) that a blank sample will give a signal at this level or higher.

CCβ

Detection capability: the concentration level of the analyte in sample at which there is probability β (again usually defined as 0.05 or 5%) that the method will give a result lower than CC_α, meaning that the analyte will be declared as undetected.

CRM

Certified Reference Material

Dynamic range

The range where the response changes when the analyte concentration is changed but the relationship may be non-linear. If the response is linear it can be specified as a dynamic linear range.

FDA

The US Food and Drug Administration

False negative

Results where analyte is declared to be below LoD although it is not.

False positive

Results where the analyte is declared to be present although actually it is below LoD.

Heteroscedasticity

Difference of variance; describes the data where the standard deviation of the signal (y-value) depends on the concentration of the anaylte (x-value).

Homoscedasticity

Homogeneity of variance; describes the data where the standard deviation of the signal (y-value) does not depend on the concentration of the anaylte (x-value).

Identity confirmation

Providing evidence that the analytical signal registered during sample analysis is due to analyte and not any other (interfering) compound.

Instrumental LoD

LoD estimated for the analysis instrument by measuring the analyte from pure solvent without any sample pretreatment.

Intermediate precision

The precision obtained within a single laboratory over a longer period of time (generally at least several months).

Linearity

Methods ability to obtain signals, which are directly proportional to the concentration of analyte in the sample.

Linear range

Range of concentrations where the signals are directly proportional to the concentration of the analyte in the sample.

LoD

Limit of detection: the smallest amount or concentration of analyte in the test sample that can be reliably distinguished from zero.

LoQ

Limit of quantitation: the lowest concentration of analyte that can be determined with an acceptable repeatability and trueness.

MRL

Maximum residue limit: the maximum allowed concentration limit set for the compound in certain matrices.

Measurement uncertainty

Measurement uncertainty (or simply uncertainty) defines an interval around the measured value $C_{MEASURED}$, where the true value C_{TRUE} lies with some predefined probability. The measurement uncertainty U itself is the half-width of that interval and is always non-negative. This definition differs from the VIM [ref 6] definition (but does not contradict it) and we use it here as it is generally easier to understand and to relate to practice.

Method LoD

LoD estimated by using matrix matched samples that are put through the whole analysis method (including the sample preparation).

Precision

Characterizes the closeness of agreement between the measured values obtained by replicate measurements on the same or similar objects under specified conditions.

Prodrug

A prodrug is a medication or compound that, after administration, is metabolized (i.e., converted within the body) into a pharmacologically active drug (e.g. by ester bond cleavage within prodrug). Inactive prodrugs are pharmacologically inactive medications that are metabolized into an active form within the body.

Quasimolecular ion

Ion formed by protonation or deprotonation of a molecule [ref 52]. Note that another IUPAC terminology source [ref 57] discourages usage of this term and instead, recommends terms "anionized, cationized, deprotonated and protonated molecule".

Recovery

The recovery of an analyte in an assay is the detector response obtained from an amount of the analyte added to and extracted from the biological matrix, compared to the detector response obtained for the true concentration of the analyte in solvent. Recovery pertains to the extraction efficiency of an analytical method within the limits of variability.

Repeatability

Expresses the closeness of the results obtained with the same sample using the same measurement procedure, same operators, same measuring system, same operating conditions and same location over a short period of time.

Reproducibility

Expresses the precision between measurement results obtained at different laboratories.

Residual

The difference between experimental signal and signal calculated according to the calibration function.

Robustness

A parameter used to evaluate constancy of the results to variations of the internal factors of the method such as sample preparation, mobile phase composition, mobile phase flow rate, injection volume, column temperature etc.

Ruggedness

A parameter used to evaluate constancy of the results when external factors such as analyst, laboratory, instrument, reagents and days are varied.

Run

See: analytical run

Scedasticity

The distribution of error terms.

Selectivity

The extent to which other substances interfere with the determination of a substance according to a given procedure.

Sensitivity

The change in instrument response, which corresponds to a change in the measured quantity; the gradient of the response curve.

Stability

Stability (*ST%*) characterizes the change in the analyte content in the given matrix under the specific conditions and over the period of time.

Trueness

Closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value.

Working range

The working range is the range where the method gives results with acceptable uncertainty. Working range can be wider than linear range.

What our participants say?

"Very happy to have completed the course successfully. It was an excellent course and helped me to learn a lot. Special thanks to the UT team for providing this opportunity to learn about method validation and for the assistance all along. Looking forward to similar courses in future."

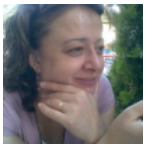
Dinesh Payingoli Kunnenoth,
India
January 2017

"Thank you for all your support and dedication for helping us to successfully understand and pass the course. It is a wonderful educational course and an excellent on-line learning environment."

Ann Yang,
USA
January 2017

"I would like to express my thanks for the opportunity to participate in the LC-MS Validation Course. The course was very informative and relevant to my field. Congratulations to you and your team for a job well done!"

Aaron Dacuya,
Philippines
February 2017



"Everything about method validation (that I am struggling with) is enclosed professionally under a single course title. I am currently not making LC MS/MS measurements but from time to time the missing pieces in the big picture of method validation disturb me. Details presented [in this course] were more than enough and this makes me comfortable that I can turn back and read more if I need it in the future. There is one more thing that I want to underline. The topic about accuracy trueness and precision is not clear among the metrologists. I am happy that you clearly define the relation between these 3 terms."

Kevser Topal,
Turkey
February 2017