MOOC: 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 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, this course intends to offer sufficient knowledge and skills for carrying out validation for the most 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. 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 only acquired through practice.

Running as MOOC

The MOOC is run in the UT Moodle environment. You can have a preview of the MOOC in Moodle environment as a guest without registration. There (course introduction) you will find detailed information about the course organization, graded tests, formation of the grade and awarding of the certificate. Guest access also allows you to view the course contents in Moodle. As a guest you cannot read forums and take quizzes.

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 videos from YouTube, using the links below each video window.
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 to some extent). It rather gives additional explanations and addresses some additional topics that were not covered in the lecture.
4. Participants are introduced to the ValChrom 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. Self-tests contain questions, as well as calculation problems. Self-tests are on the 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 the validation in different practical situations. Therefore, self-tests are intrinsic components of the course and participants are strongly encouraged to take all of them.

The printout of the current version of the course materials (including links to the lecture slides) can be downloaded from here.

Self-testing

Throughout the course there are numerous self-tests for enabling a participant to test his/her knowledge and skills in specific topic. 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 the score 100% and only then move to the 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).

If you would like to learn more ...

This course is part of the Excellence in Analytical Chemistry (https://www.analyticalchemistry.eu/) *Erasmus Mundus* master's programme, which offers education in all aspects of Analytical chemistry, including metrology in chemistry (measurement uncertainty, method validation, reference materials, etc), as well as economic and legal aspects of chemical analysis.

This course is run within the framework of the Estonian Center of Analytical Chemistry with the aim of offering easily accessible knowledge in analytical chemistry to labs and industries in Estonia and elsewhere.

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. The European Commission: Guidance document on analytical quality control and method validation procedures for pesticide residues and analysis in food and feed. Also known as SANTE (formerly SANCO). The link leads to a page where the most recent version, as well as the older versions can be found.
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.
   Note that this directive is superseded by Commission Implementing Regulation (EU) 2021/808 of 22 March 2021 on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling.
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. 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. 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.

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 ten 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 the head of Research and Development department in the Estonian Environmental Research Centre (EERC) and responsible for transferring LC-MS methods for environmental samples from development to routine analysis, including accrediation of methods according to ISO 17025. Riin also works as a technical assessor for the Estonian Accreditation Centre and is also an Eurachem executive committee elected member. In 2020 she obtained a master's degree in law from University of Tartu with a focus on intellectual property (mostly patents).

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 with accreditation 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 the 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. Karin Kipper carried out her post-doctoral fellowship focussing on ABDose clinical trial on antibiotics dosage for critically ill patients in the St George's University of London between 2014 to 2016. Karin worked as Study Director and Head of Clinical Department in Analytical Services International Ltd between 2016 to 2019 supporting clinical trial work. Since 2019, Karin works as Head of Therapeutic Drug Monitoring Unit in Epilepsy Society. Her main research fields are pharmaceutical bioanalysis (pharmacokinetic/pharmacodynamic studies), therapeutic drug monitoring, pharmaceuticals' pathways in the environment and development of novel eluent additives for LC-MS in order to improve separation, peak shape and LLoQ 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.

Asko Laaniste obtained his PhD from University of Tartu (UT) in 2016. His PhD work was about the comparison and optimization of different LC-MS ionization sources. Since 2016 he has worked on several collaboration projects between UT and chemistry industry regarding development and validation of chromatographic methods. Since 2018 his main goal has been the development of validation automation software ValChrom, which is integrated into this MOOC. He is currently studying software development in Tallinn University of Technology (TalTech).

Maico Lechner graduated his bachelor's degree 2020 at Aalen University in Germany. The thesis work has been performed in the Testing Center of Tartu University on the topic of "Development of LC-MS method for the analysis of pharmaceutical residues in environmental samples using Design of Experiments approach". Since September 2021 he is studying the Erasmus Mundus master's degree Excellence in Analytical Chemistry (EACH) at Tartu University. Previous working experiences include a 6-month

internship at the Methrom-AG headquarters in Germany in the field of development and application of automated titration methods and several small research project at Aalen University.

Former team members

Anneli Kruve graduated in 2011 from the University of Tartu and continued my studies as a post-doc in Techion, Israel. She was a Humboldt fellow at Freie Universität Berlin and has supervised several doctoral students at the University of Tartu. She joined Stockholm University in 2019 and is in charge of the mass spectrometry laboratory. In 2018 Anneli Kruve was selected to be among the Top 40 under 40 power list by the Analytical Scientist.

The fields of study of the group of Anneli Kruve are focused on the fundamentals and applications of liquid chromatography mass spectrometry. Specifically, on establishing (semi-)quantitative non-targeted analysis and applying ion mobility separation for complex structural characterization. Her group uses modelling and machine learning to understand ionization processes in electrospray (ESI) and developing semi-quantitative non-targeted analysis methodology. Go to her group website for more.

Maarja-Liisa Oldekop obtained her Ph.D. from the University of Tarty (UT) in 2017 and is now working as a quality manager in private sector. Her main field of expertise is development of LC-MS methods using derivatization.

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 five subsections:

1.1. The purpose of validation

Method validation is a key activity in chemical analysis and 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 that 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 the method if the performance is not as expected) as well as to 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 many 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.

Note that all terminology, abbreviations used in the course can be found in Glossary.

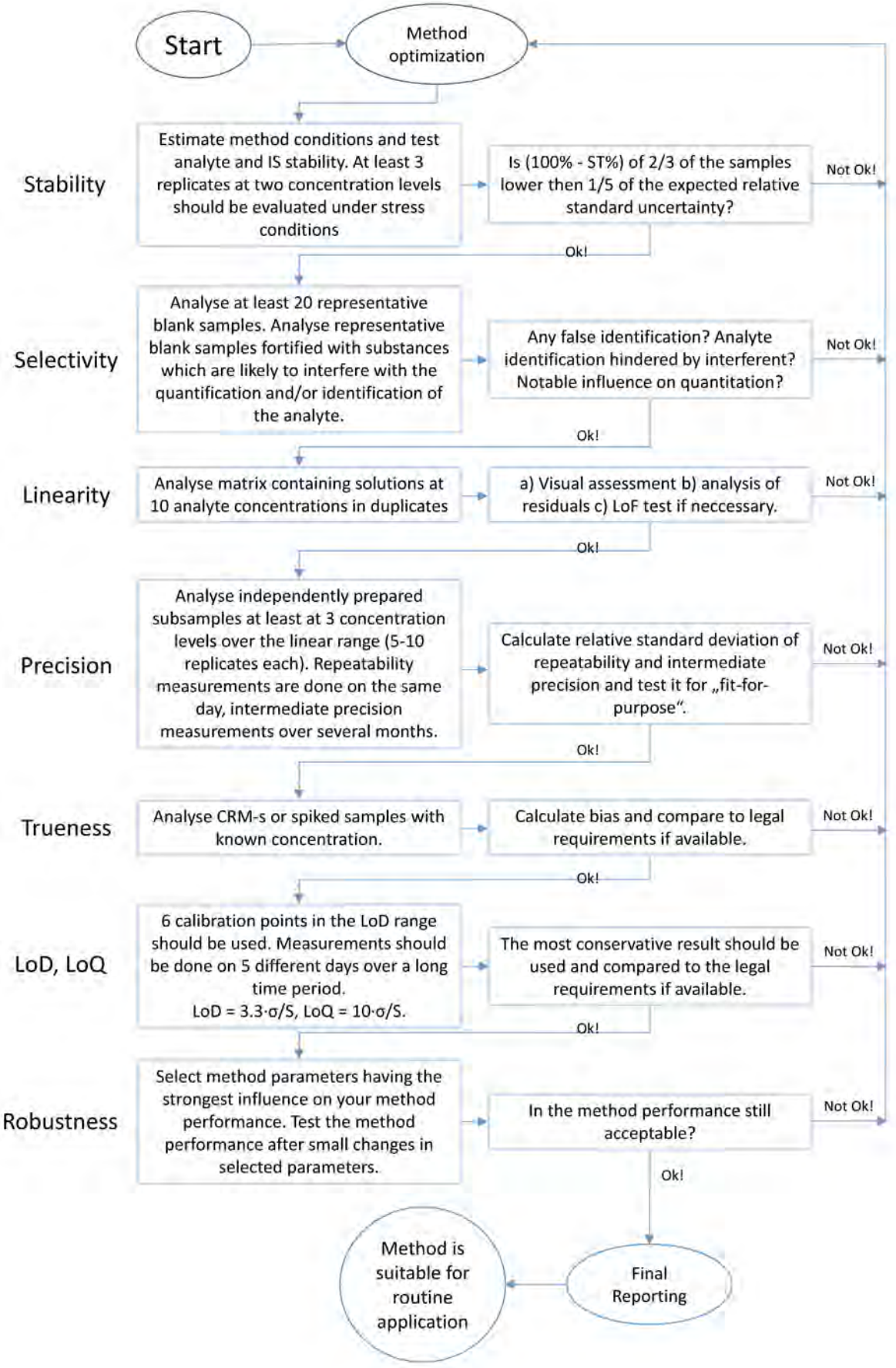


Figure 1. A possible sequence of operations and decisions in LC–MS method validation. All steps are explained in detail in upcoming chapters. *ST*% refers to stability in per cents, explained in section 8.3. LoF stands for Lack-of-Fit test, explained in 3.3. CRM stands for certified reference material.

Before starting a validation, a clear validation 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. Requirements often result from guidelines or from other regulatory documents (directives, standards etc.). A validation plan depends on each of the different method under development and it 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 how to evaluate each performance parameter separately. At the same time, guidance on how to decide about the whole method's validation is usually very general: validation has to demonstrate that the values of all evaluated parameters are satisfactory. Few different cases arise.

- (a) When methods are applied in the scope of standards, laws or directives, then requirements from those documents, if present, must be followed and a decision on validation should be based on these documents. When a decision on validation suitability is based on the guidelines, then the decision for each parameter must be given separately according to the requirements.
- (b) Sometimes the client can specify the requirements, then 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.

Validation should start with evaluating analyte stability and method selectivity as all other parameters strongly depend on these. For example, if an analyte extensively decomposes in the autosampler, no linear relation can be achieved. In that case, non-linear calibration models need to be considered. Consequently, we propose an 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.

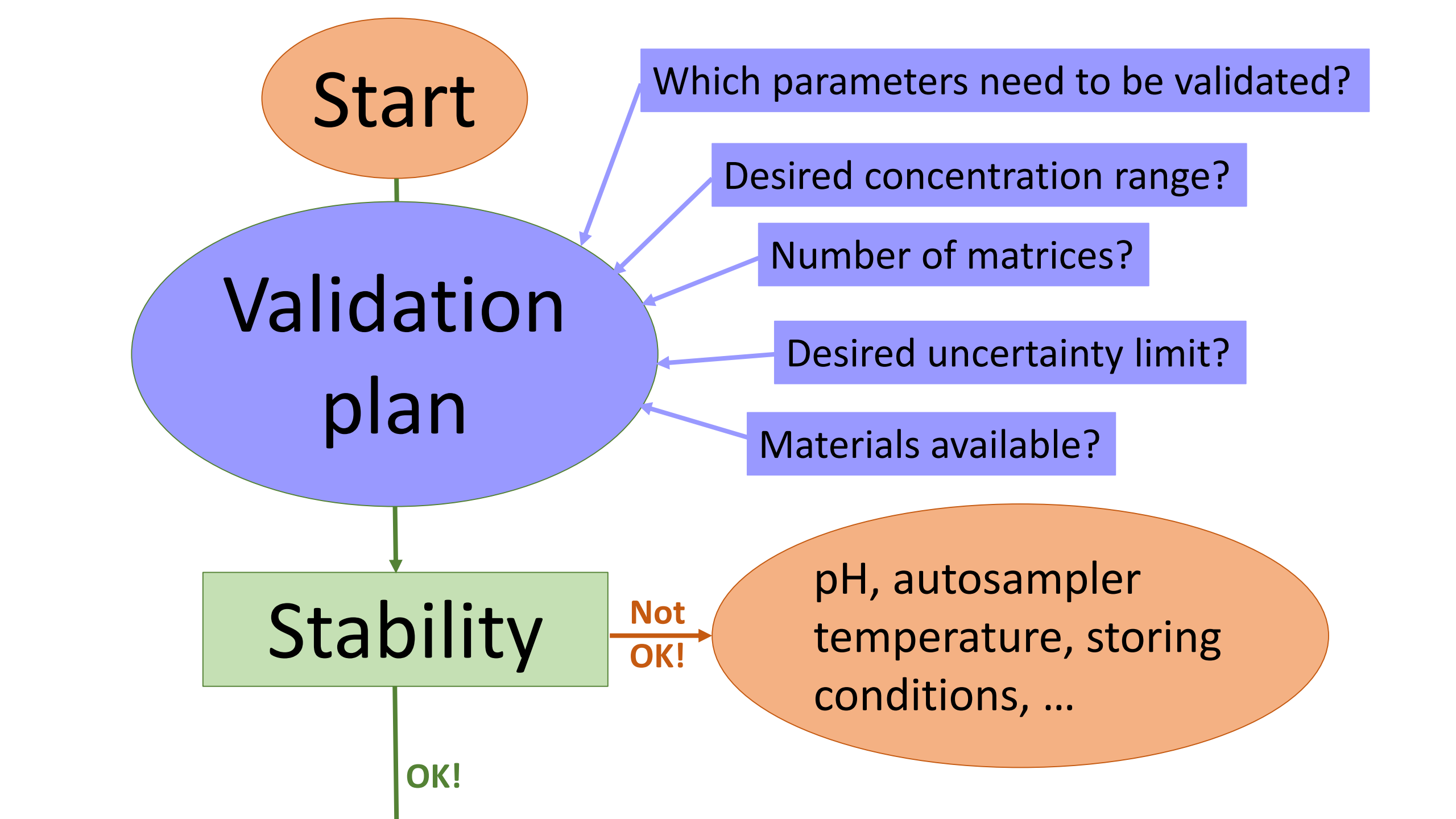
Not only validation but also an appropriate documentation of the validation is required for an adequate interpretation as well as on the validity of the obtained results. As the last stage of validation, an assessment of validity (fitness for the intended purpose) of the method should be given, based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

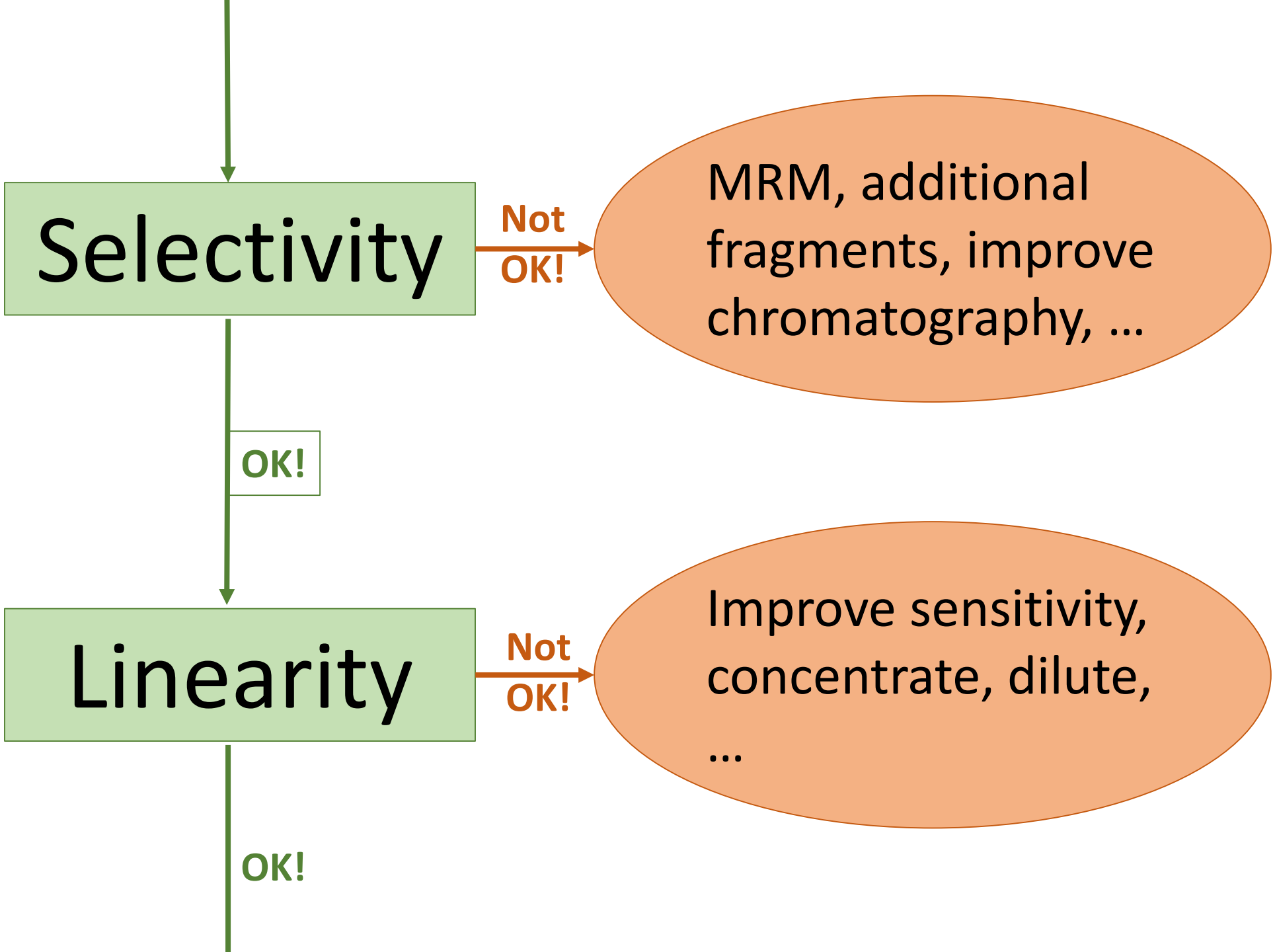
Overview of validation

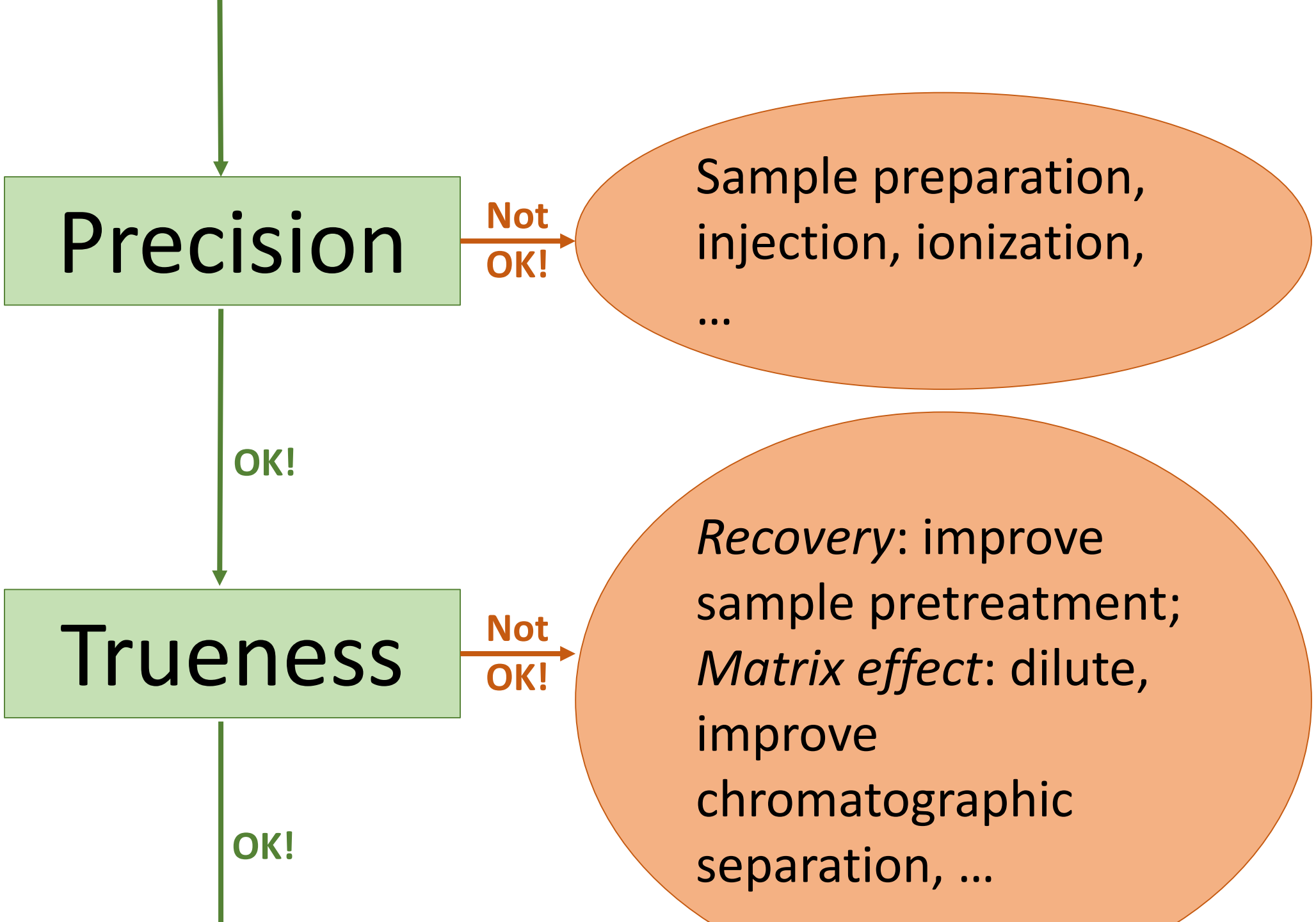
http://www.uttv.ee/naita?id=23287

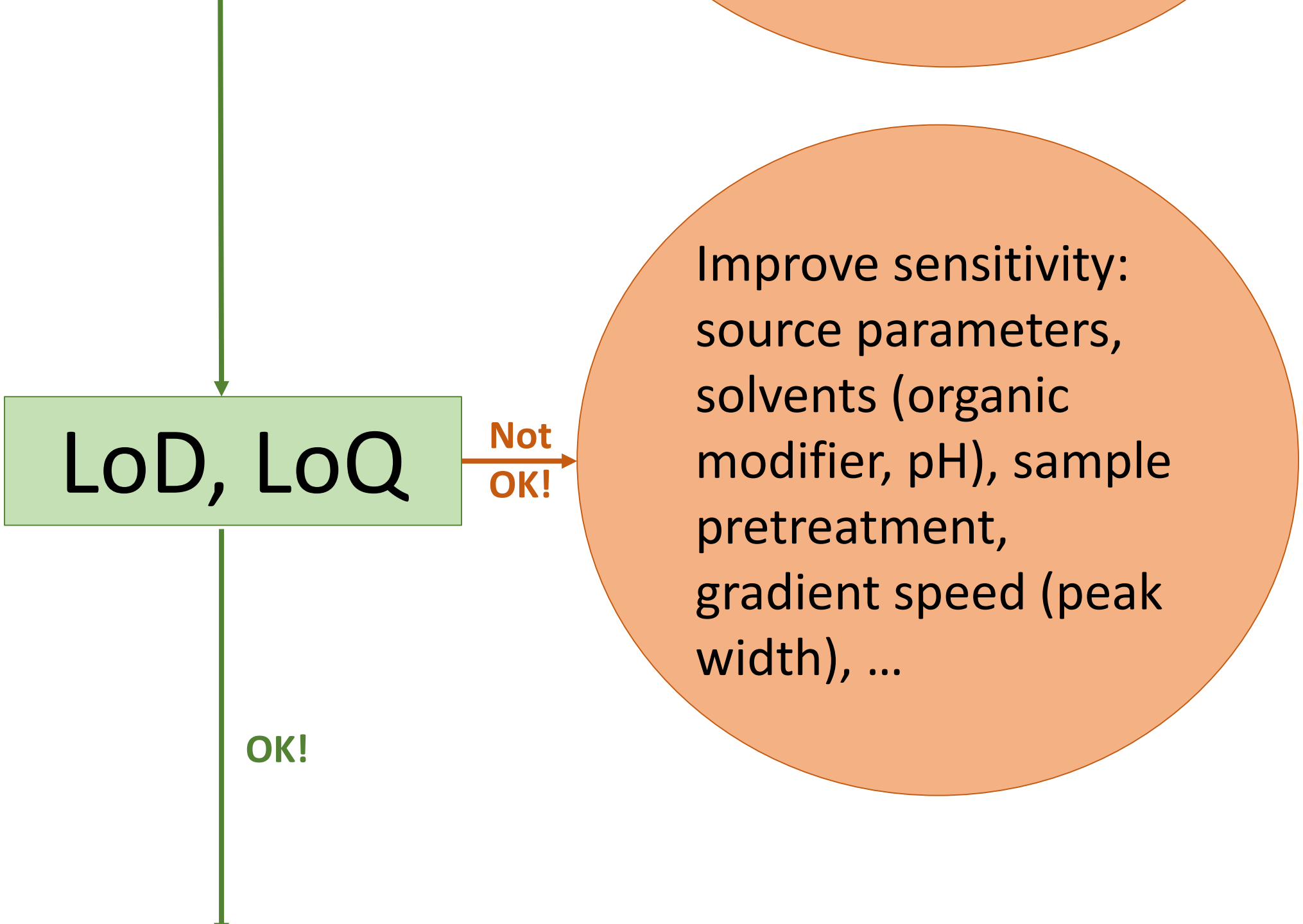
https://www.youtube.com/watch?v=K312lRGTJFk

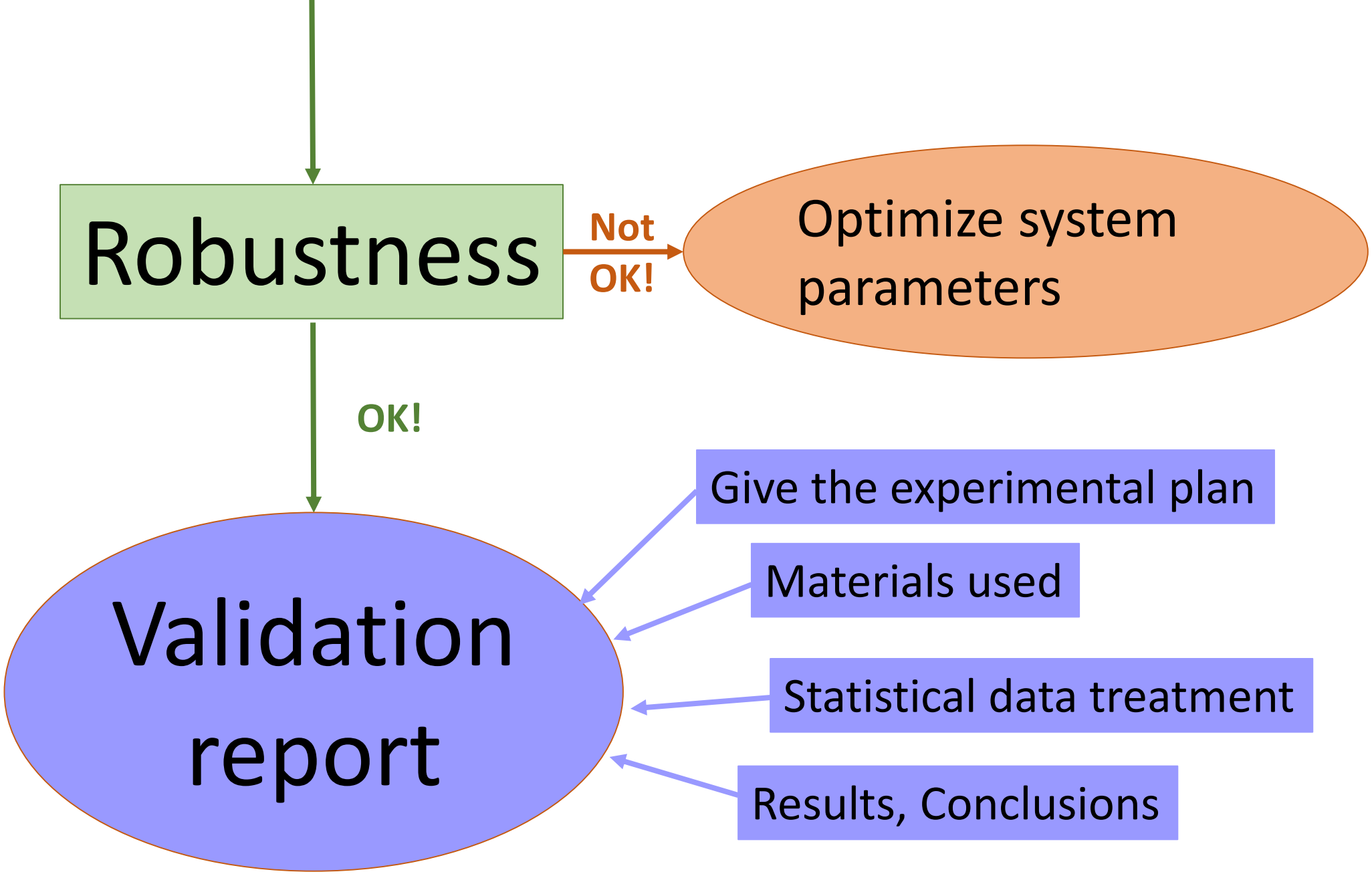
1.2_overview_of_validation.pdf 267 KB


Start
Validation plan
Which parameters need to be validated?
Desired concentration range?
Number of matrices?
Desired uncertainty limit?
Materials available?
Stability
Not OK!
pH, autosampler temperature, storing conditions, …
OK!


Selectivity
Not OK!
MRM, additional fragments, improve chromatography, …
OK!
Linearity
Not OK!
Improve sensitivity, concentrate, dilute, …
OK!


Precision
Not OK!
Sample preparation, injection, ionization, …
OK!
Trueness
Not OK!
Recovery: improve sample pretreatment; *Matrix effect*: dilute, improve chromatographic separation, …
OK!


LoD, LoQ
Not OK!
Improve sensitivity: source parameters, solvents (organic modifier, pH), sample pretreatment, gradient speed (peak width), …
OK!


Robustness
Not OK!
Optimize system parameters
OK!
Validation report
Give the experimental plan
Materials used
Statistical data treatment
Results, Conclusions

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 a so-called verification: verify that a method 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 when methods with a narrow application range are validated. Small changes can include transfer of the method to another laboratory, adding a new matrix, implementing new reagents in sample the preparation, etc. FDA states that in the case of bioanalytical methods, it is sometimes sufficient to determine only trueness and repeatability. The ICH guideline 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

Working range

Trueness

Recovery

Selectivity

Precision

Intermediate precision

LoD

LoQ

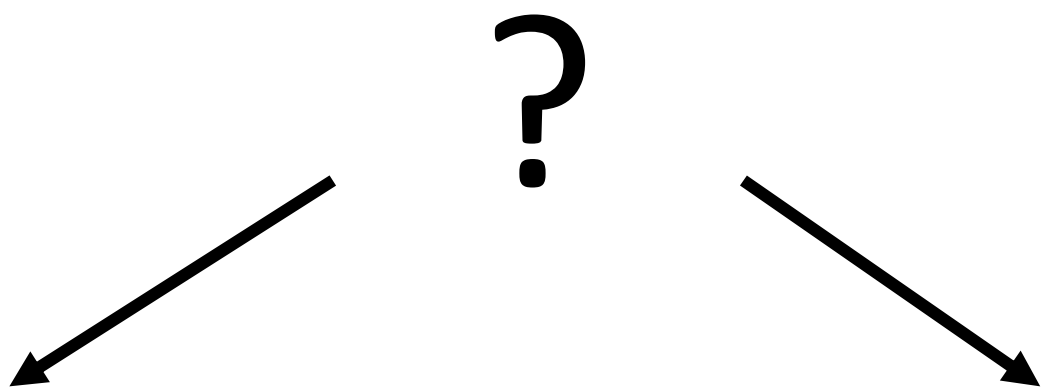
Linearity

Repeatability

Measurement uncertainty

Robustness

Matrix effect



Full validation Partial validation

Validation is always a balance between costs, risks and technical possibilities

ISO 17025:2005 General requirements for the competence of testing and calibration laboratories

	Identification	Impurities (quantitative)	Impurities (limit test)	Assay - Dissolution - Content
Accuracy		√		√
Precision		√		√
Selectivity	√	√	√	√
LoD			√	
LoQ		√		
Linearity		√		√
Range		√		√

ICH Harmonised Tripartite Guideline: Validation of Analytical Procedures: Text and Methodology Q2(R1) 2005

Initial validation

All required parameters for all analytes in all matrices

Ongoing performance validation

Determine some parameters for some analytes in some matrices during routine analyses

New analytes or matrices

Determine as in initial full validation

OR

Integrate to the ongoing performance validation

SANCO/12571/2013 Guidance Document on analytical quality control and validation procedures for pesticide residues analysis in food and feed

Adopting a standard method

The laboratory shall confirm that it can properly operate standard methods before introducing the tests or calibrations

ISO 17025:2005 General requirements for the competence of testing and calibration laboratories

Working range

Trueness

Recovery

Selectivity

Precision

Intermediate precision

LoD

LoQ

Linearity

Repeatability

Measurement uncertainty

Robustness

Matrix effect

1.4. ValChrom

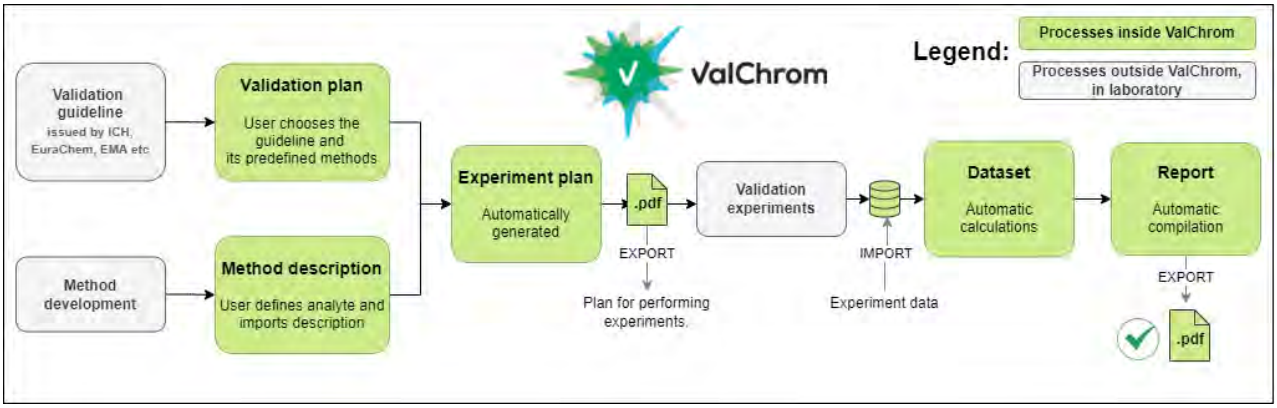
Introduction

ValChrom is an online software tool for the validation of chromatographic analysis methods, which currently implements validation guidelines from 3 international organizations (Eurachem [ref 3], EMA [ref 7], ICH [ref 17]). ValChrom lets the user choose between validation guidelines, specify validation parameters and the ways how the validation parameters are assessed. This information is used by the system for creating the **Validation Plan Template**, where the way the calculations should be done is determined. The **Experimental Plan** is a detailed description of the experiments that have to be performed in the laboratory, built based on validation plan template. After performing the planned experiments, the user enters the experimental data into **Experimental Dataset**. Using this data, ValChrom calculates the required validation parameters and creates a validation report.

In this course ValChrom can be used (not mandatory) as a learning aid in parallel to or instead of spreadsheet programs. There are also some bonus self-tests (not graded tests) that are meant to be solved using only ValChrom (e.g. CC_α and CC_β calculation).

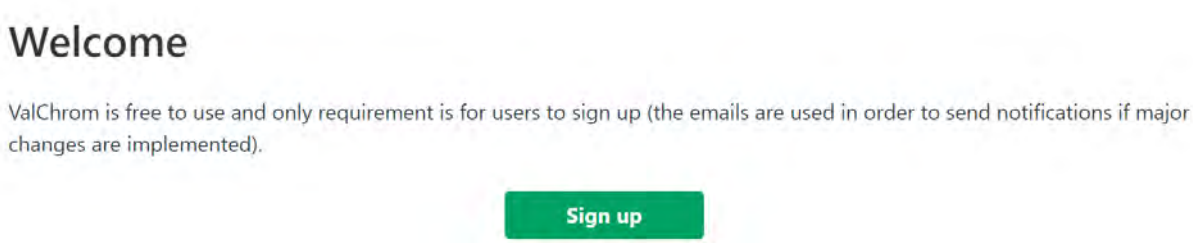
ValChrom is still under development. Any feedback on how to improve the software or any suggestions for additional functionality is greatly appreciated at valchrom@ut.ee or +372 737 6030.

The flowchart of ValChrom is presented below.



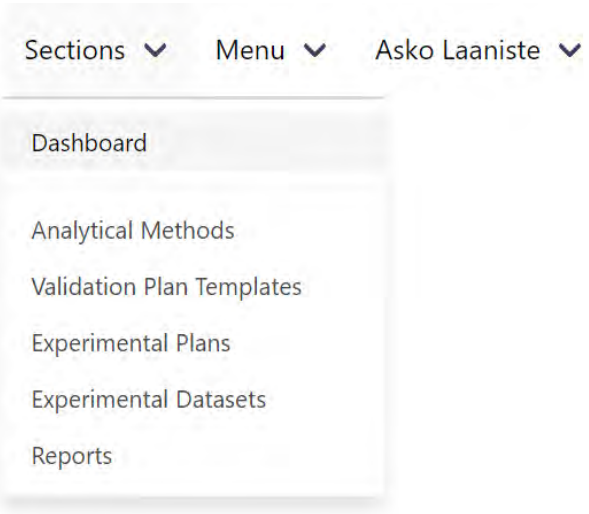
Signing up/logging in

1. Go to https://valchrom.ut.ee
2. Create a user account using the Sign up option. Username is an email address. Please use your real e-mail address, so we can inform you about major updates.

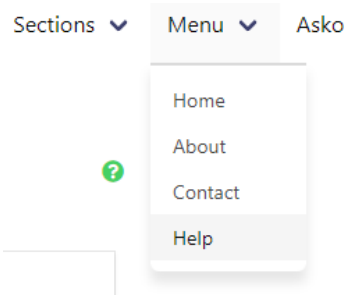


Inside ValChrom

When logged in to ValChrom you are directed to dashboard with the flowchart of ValChrom.



The sections depicted in the flowchart are accessible in the upper right corner, under "Sections". In this tutorial we will visit 4 sections: Analytical Methods, Validation Plan Templates, Experimental Plan and Experimental dataset. For help look in the Help menu or specific view´s green questionmark.



Analytical Methods

This part of the software is used to describe the analytical method that is validated. You are welcome to create as many method descriptions as you like, but **in order to follow this tutorial and solve ValChrom-based self-tests in this course, please follow the instructions below precisely**.

1. Sections > Analytical Methods.

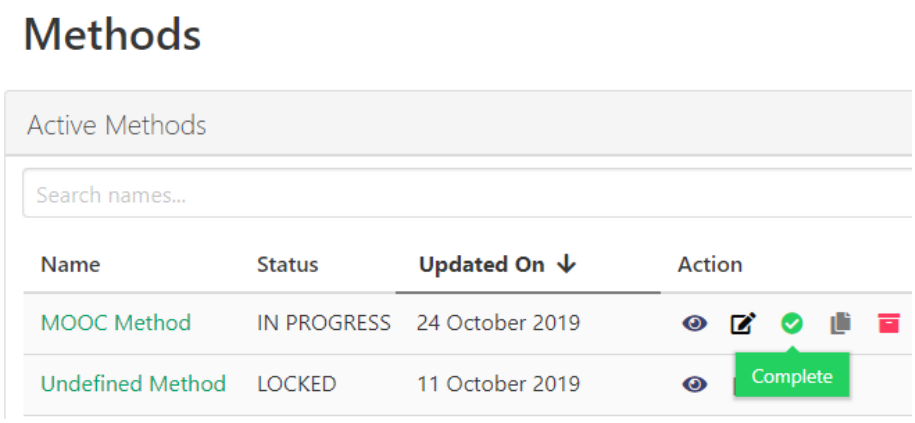
2. Click on the large + sign at the lower right corner of the window. 
3. Insert Method name: **MOOC method**
4. Insert Concentration unit: **ppm**
5. Insert Area unit: **kcount*s**
6. Insert Retention time unit: **min**
7. Insert Description: **Sample method used in MOOC self-tests**.
8. In section "Add Analyte" insert:

- Short Name: **universalite**
- Official Name: **universal analyte**
- Abbreviation: **UA**
    - **NB!** This abbreviation creates link between the method and experimental data (see "Uploading Instructions" under section 1.5).
- Target Concentration: **5** (not used in the course; refers to the expected concentration, e.g. active substance concentration in tablet)
- Click "Add"

9. Under Method File (.zip) and Method Description file (.pdf) one could upload method file from the instrument software (zipped) and longer description of the method (as pdf file). These purely descriptive sections are not used in the MOOC and can be left blank.
10. Click "Submit".
11. In the Methods table (Sections > Analytical Methods) change the status of the method to "COMPLETE" by clicking respective action button.

- Method with "IN PROGRESS" status cannot be used downstream of the program flow (i.e. to make a new experimental plan).

- If Analytical method is used downstream of the program flow, then its status will change to "LOCKED" and cannot be edited anymore.



12. **Well done, we now have a method with universal analyte to use!**

Validation plan template

Validation plan template is where the user selects which validation parameters (e.g. linear range) to assess and how the assessment (e.g. via back-calculated concentration as described in EMA guideline) should be carried out. Once a template has been created, it can be used for as many different experimental plans as needed.

The user can currently choose between 3 guidelines, ICH [ref 17], Eurachem [ref 3], EMA BA [ref 7].

1. Sections > Validation plan templates.

2. Click on the large + sign at the right bottom of the window.
3. Insert Template Name: **MOOC validation plan template**
4. Insert Comment: **Validation plan for MOOC.**
5. In section "Select a guideline for your validation plan" choose "Mixed guidelines".
6. Click Continue and choose the following **assessment methods**:

- LoD and LoQ
    - Based on calibration graph | Slope and standard deviation of **residuals**
    - CC_α and CC_β for heteroscedastic data
- Linear range
    - 1 series / 5 concentration levels / 1 parallel per level
- Precision
    - Repeatability | 6 injections at target concentration
    - Intermediate precision | 6 injections at target concentration
    - **NB!** Pay attention that there are several assessment methods with similar name (e.g. "Based on calibration graph | Slope of calibration graph and standard deviation of **y-intercept**" AND "Based on calibration graph | Slope of calibration graph and standard deviation of **residuals**"). Each assessment method has its own distinct logic and they mostly cannot be used interchangeably.

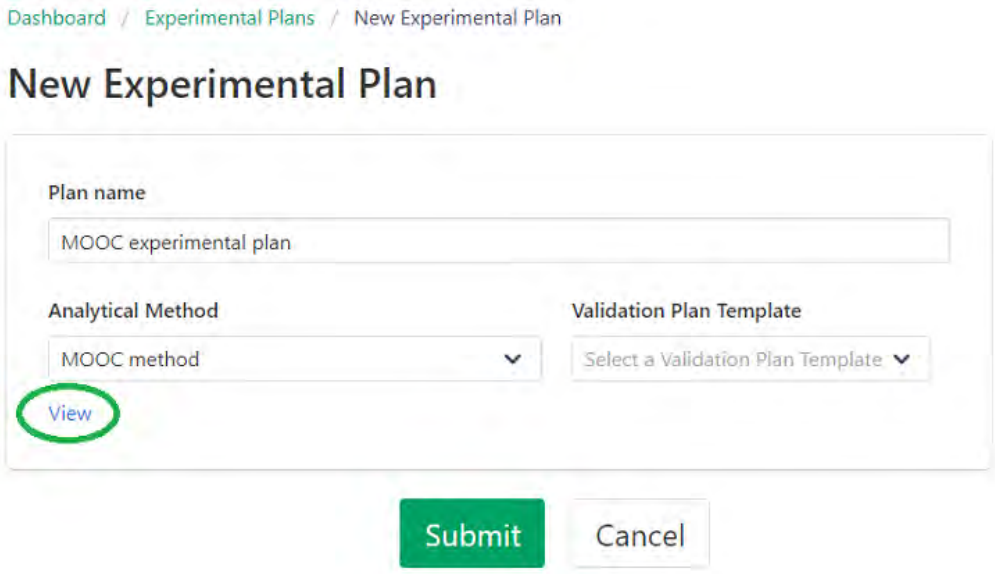
7. Assessment methods define how the results are calculated and what is the minimum data required to perform these calculations.
8.Click "Submit".
9. In the Validation Plan Templates table (Sections > Validation Plan Templates) change the status of the method to "COMPLETED" by clicking the green button.
10. Excellent job! Now we have a template we can use in all cases where ValChrom is used in the course.

Experimental plan

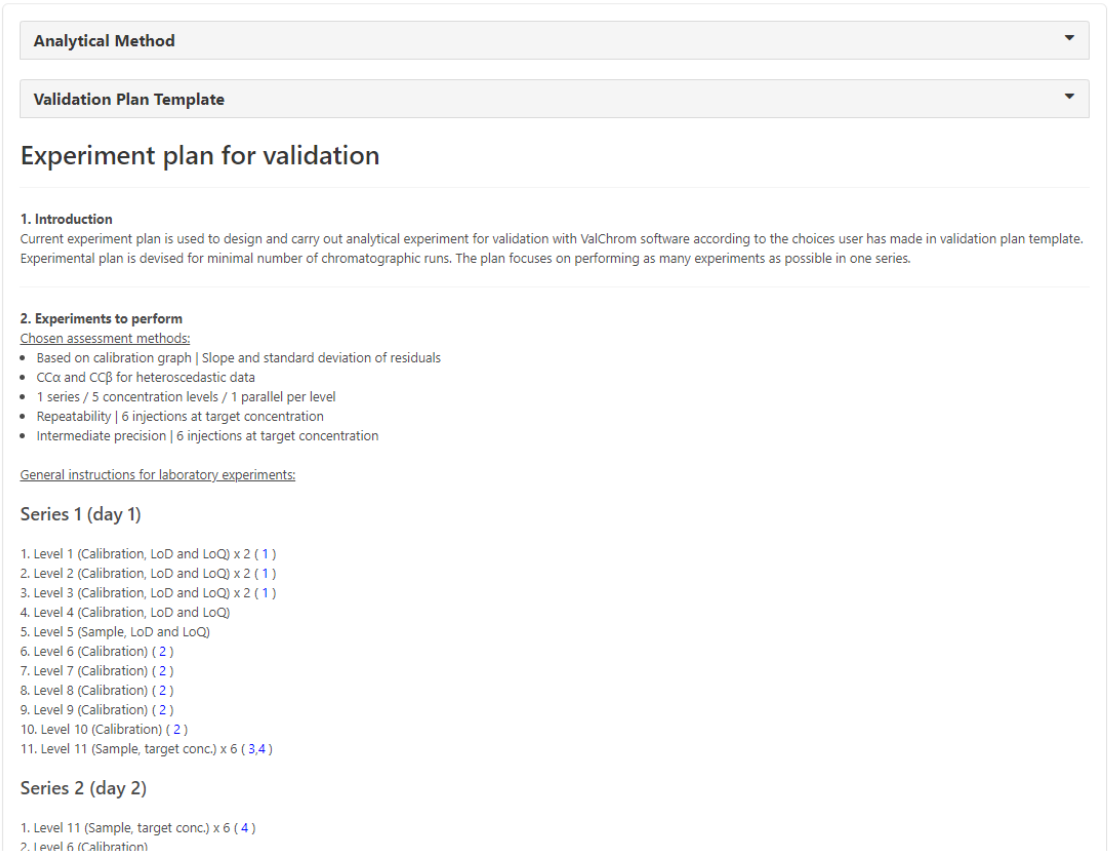
Experiment plan combines an Analytical Method with Validation Plan Template to create an optimized sequence of laboratory experiments for method validation.

1. Sections > Experimental plans.

2. Click on the large + sign at the right bottom of the window.
3. Insert Plan name: **MOOC experimental plan**
4. Select Analytical Method: **MOOC method** (If MOOC method is not available in the drop-down list, then most probably its status was not changed to "Completed")
5. You can check if you have the right method by clicking on "View" button (e.g. check if the analyte abbreviation is UA).



6. Select Validation Plan Template: **MOOC validation plan template**. Here as well, you can use "View" button to check the data.
7. Click "Submit".
8. In the Experimental Plans table (Sections > Experimental Plans) click the Complete button to make the plan available downstream of the software.
9. Open the experiment plan to display a list of experiments needed to complete in the laboratory to acquire data necessary for calculations. ValChrom combines experiments based on Validation plan template to get the optimal plan. However, analytical chemists can go over these plans with critical mind and adapt it to their specific needs. In this MOOC, the plans are not required, as raw data is provided to you when needed, but take a moment to get an idea of the experimental work required.



This concludes our initial steps required for using ValChrom for use in this validation MOOC. In next sections we will cover creating datasets and how to upload data. **See you there!**

1.5. Uploading data in ValChrom

In this section we will upload our first dataset into ValChrom. Data can be uploaded in two main ways:

- Copy from a file and paste into ValChrom
- Uploading a file (.xlsx, .csv, etc.)

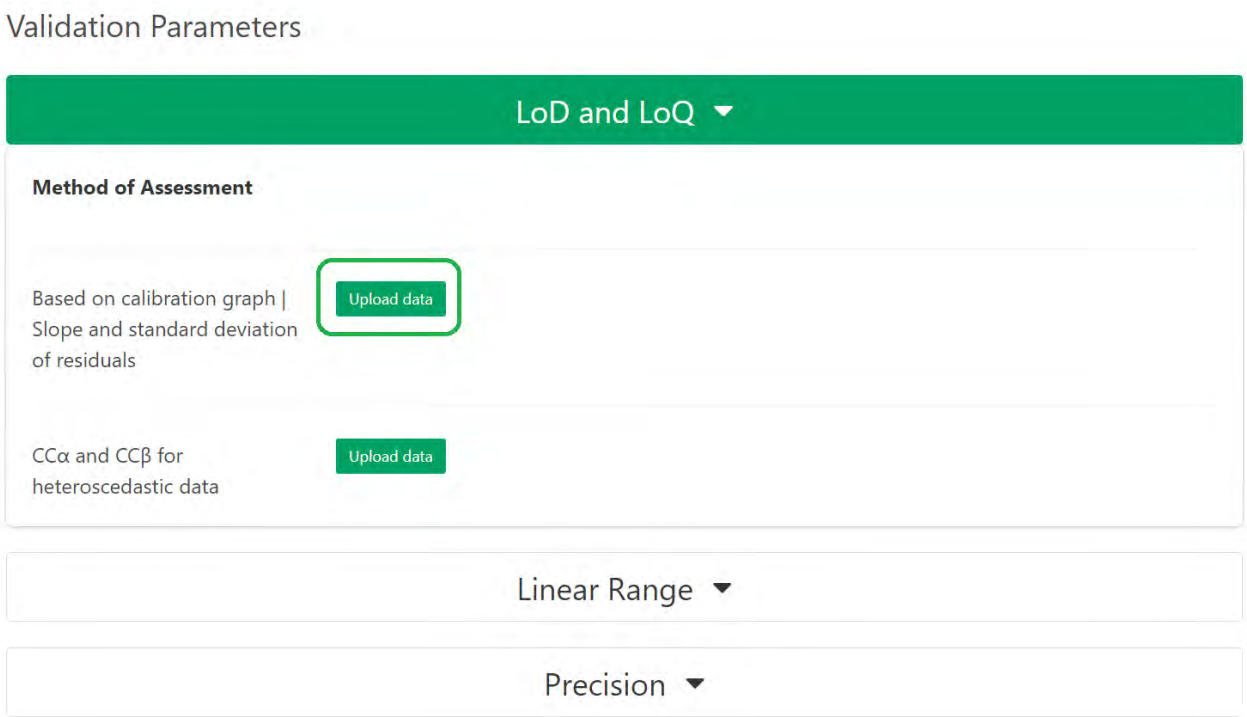
In this MOOC you can use ready-made files for the problems presented in ValChrom-based self tests. Links to these files can be found within each such test.

Experimental Datasets

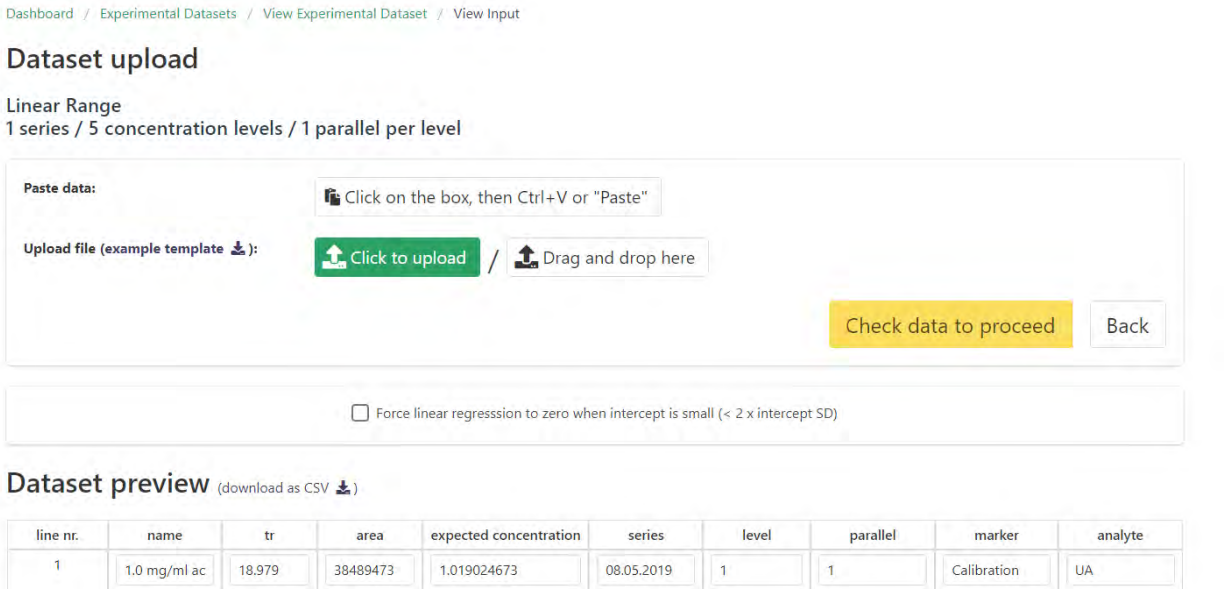
Experimental Datasets are used for holding results of experiments, which were carried out according to the Experimental Plan. The user can start populating a dataset when experiments for some validation parameter have been carried out, e.g. one can start assessing linearity when respective measurements are completed without the need to wait for all the rest of data to become available.

1. Sections > Experimental Datasets.
2. Click on the large + sign at the right bottom of the window.
3. Insert Dataset name: **MOOC dataset**
4. Select an Experimental Plan: **MOOC experimental plan**
   (If not available, then check that Experiment plan status is "COMPLETE")
5. Click "Next". You will be directed to upload view, where you can choose under which specific assessment method you want to upload data.

Data must to be uploaded for each assessment method separately. Click on "Click to upload" button or drag the file onto the button. Next we'll follow this with example data.



Choose assessment method "Based on calibration graph | Slope and standard deviation of residuals" by clicking "Upload data" button. You will be directed to data upload view.

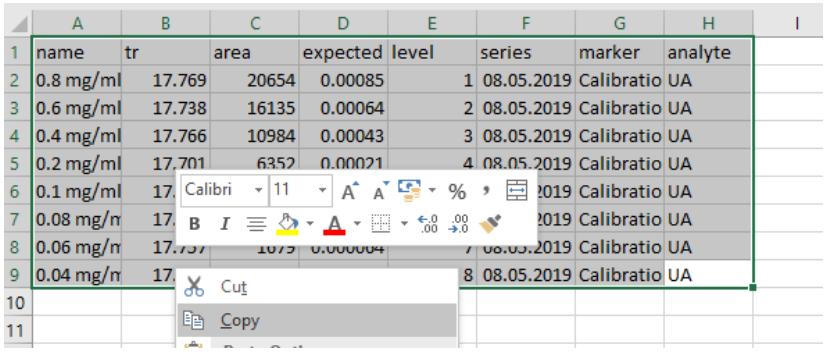


Next we will go over the two ways to upload data.

Uploading data by copying from file

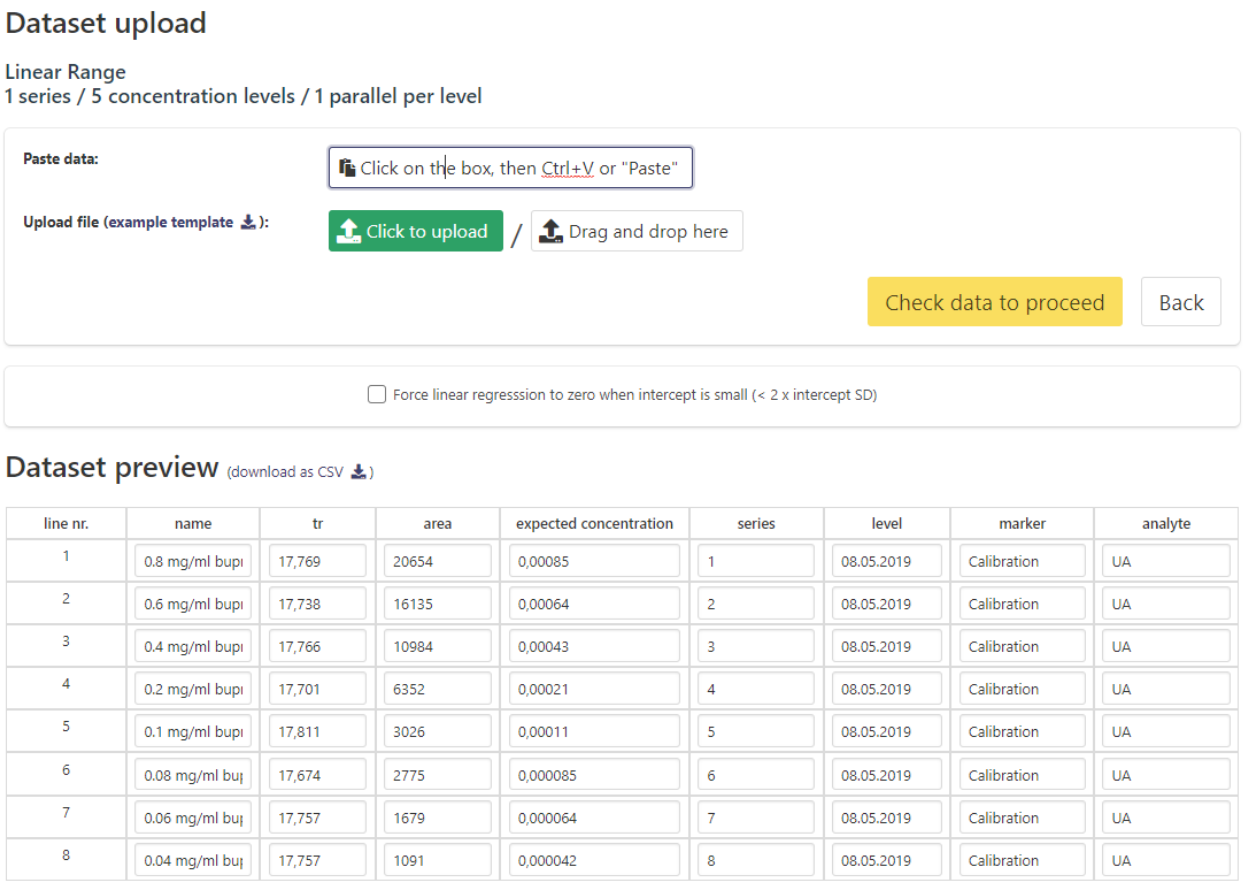
In order to copy data:

1. Download the "test_data.xls" file with raw data: test_data.xls and open the file.
2. Copy the data from the file.



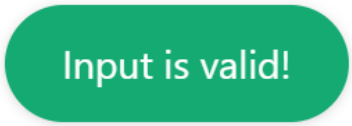
3. Click on the field with "Click on the box, then Ctrl+V or "Paste"", so that a cursor appears there.

4. Paste your data (Ctrl + V).
5. You will see your data appear below at "Dataset preview". Here a quick glance at the data should be made in order to fix obvious errors immediately.



line nr.	name	tr	area	expected concentration	series	level	marker	analyte
1	0.8 mg/ml bupi	17,769	20654	0,00085	1	08.05.2019	Calibration	UA
2	0.6 mg/ml bupi	17,738	16135	0,00064	2	08.05.2019	Calibration	UA
3	0.4 mg/ml bupi	17,766	10984	0,00043	3	08.05.2019	Calibration	UA
4	0.2 mg/ml bupi	17,701	6352	0,00021	4	08.05.2019	Calibration	UA
5	0.1 mg/ml bupi	17,811	3026	0,00011	5	08.05.2019	Calibration	UA
6	0.08 mg/ml bu	17,674	2775	0,000085	6	08.05.2019	Calibration	UA
7	0.06 mg/ml bu	17,757	1679	0,000064	7	08.05.2019	Calibration	UA
8	0.04 mg/ml bu	17,757	1091	0,000042	8	08.05.2019	Calibration	UA

6. Click "Check data to proceed" to check if all data is present.
7. Also, a button "Show results" appears. Clicking this button will direct straight to the calculated results (see section "Overview" below). Clicking "Back" takes you back to "Upload Experimental Dataset" page.



Uploading a file

In order to upload a file:
1. Download the "test_data.xls" file with raw data: test_data.xls.
2. Click on the "Click to upload" button to find the file with the data.
3. Alternatively, you can drag and drop the file on the button.

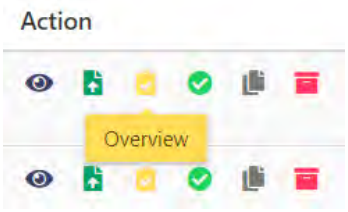
4. From here on follow point 5. and onward in previous section "Uploading data by copying from file".

Overview

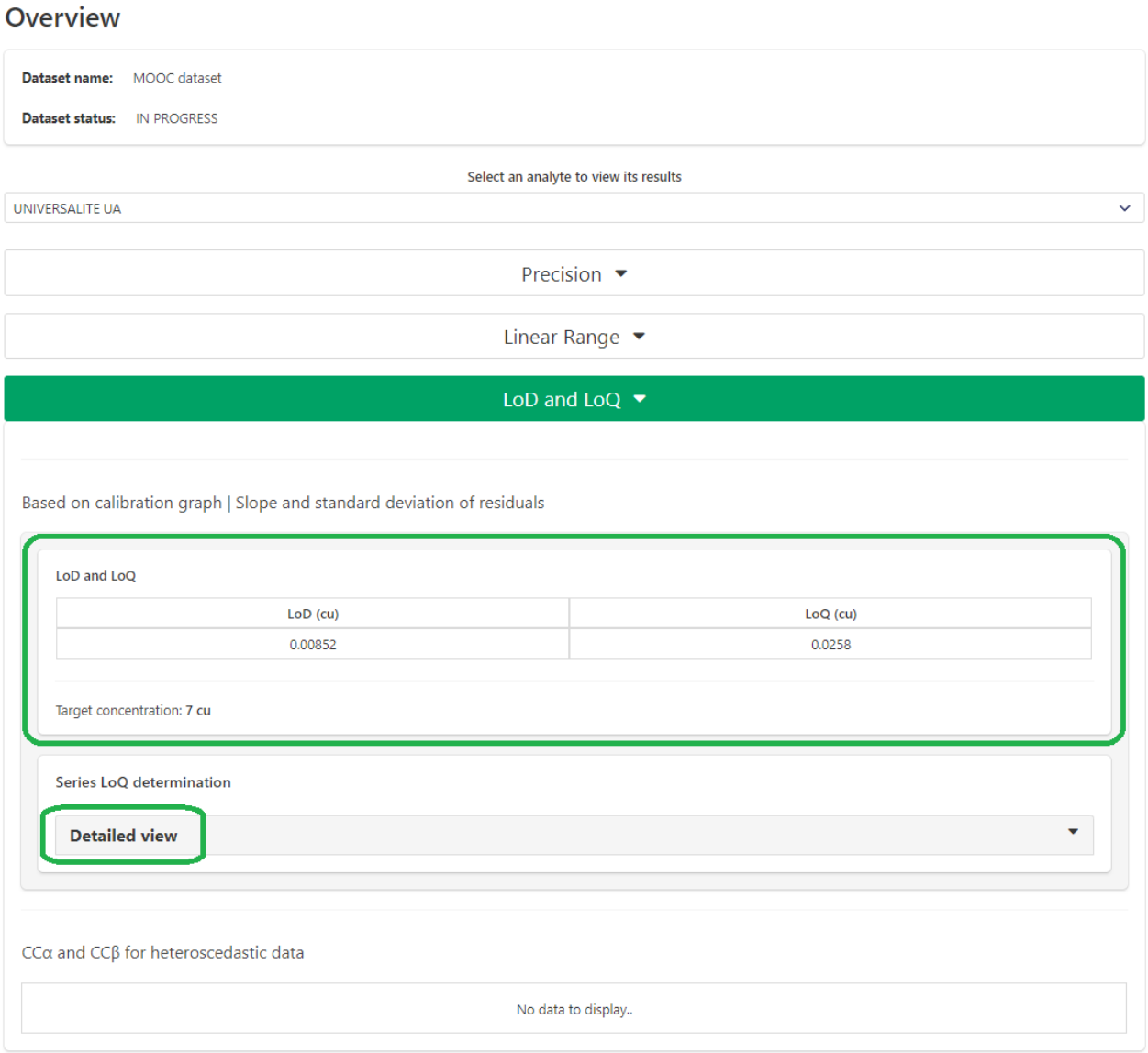
Overview is divided by validation parameter. Every validation parameter is presented in two sections: Overview and Detailed view. At once results for one analyte are shown. Clicking on a specific validation parameter will open Overview for it.

In Overview only very few data are shown, generally representative of all series or the outcome of the assessment method. For example, in case of LoD and LoQ this would be the overall LoD and LoQ values of the assessment method (taken as the maximum (conservative) values, if there are several series to choose from). In case of Accuracy, overview would have pooled standard deviation calculated over all series data. **Detailed View** offers all data available: input data and calculated results with graphs, if applicable.

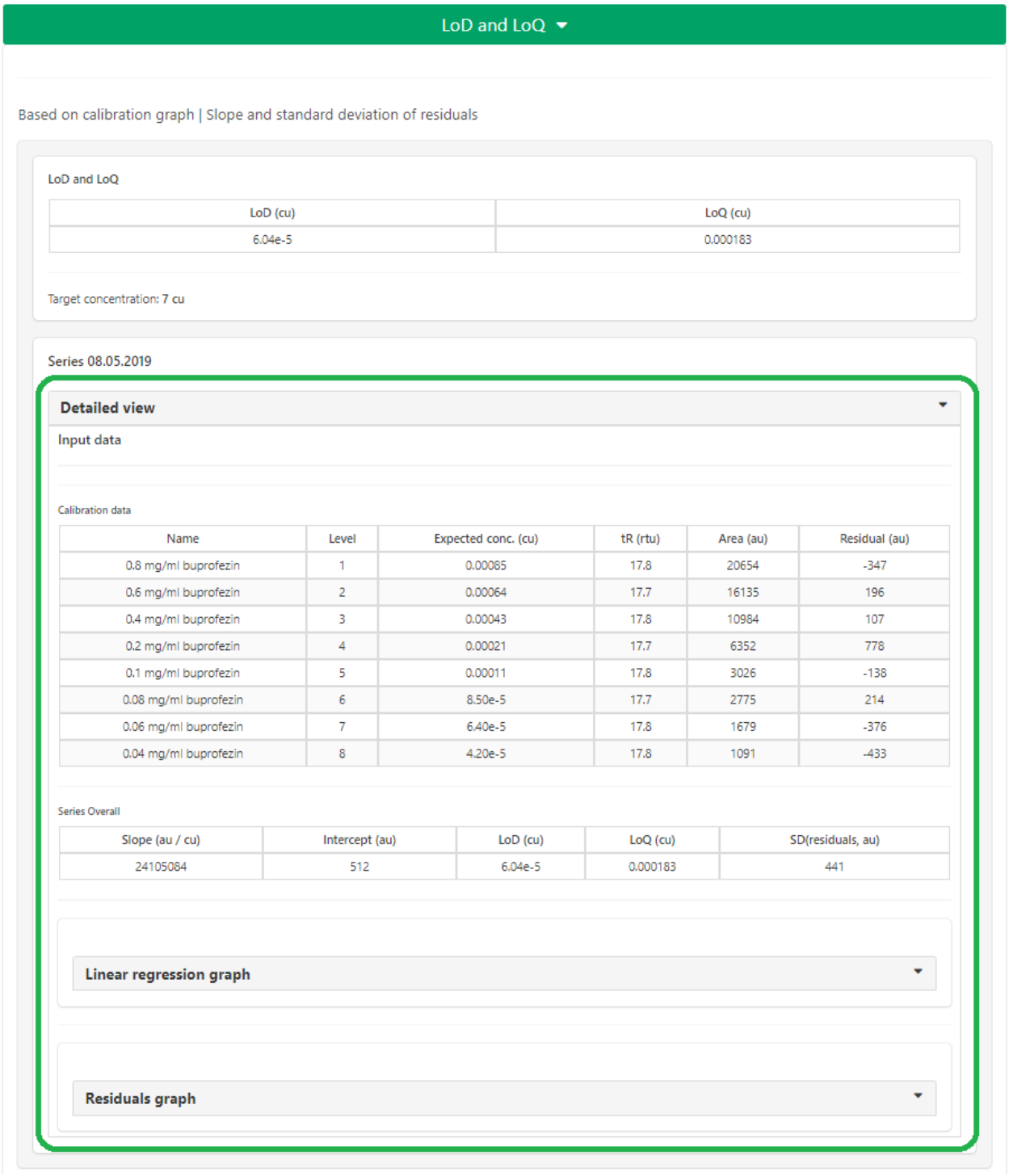
1. Go to Overview either by clicking "Show Results" button after uploading the data.
2. Alternatively, you can go to Sections > Experimental Dataset. Find your dataset and click on Overview button or the Dataset name.



3. You will see overview divided into several sections based on validation parameters. Clicking on a validation parameter will show Overview for assessment methods.



4. Click on “Detailed View”.

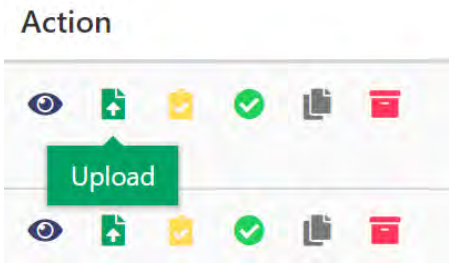


5. **You can now see the automatically calculated results for the test data!** In the "Input data" table the raw data used for the calculations can be seen, together with some calculated results for each table row (e.g. residuals). In "Series overall" table data calculated from the whole series is shown (e.g. LoD and LoQ values, or standard deviation of residuals: SD(residuals)). Try clicking on "Linear regression graph" or "Residuals graph".

6. **Note on MS Excel and programming languages precision**: Excel is following double-precision floating-point format from the IEEE 754 specification. Strictly following the IEEE 754 specification causes some loss of precision, as can be seen from Microsoft´s documentation on the topic [ref: https://docs.microsoft.com/en-us/office/troubleshoot/excel/floating-point-arithmetic-inaccurate-result]. Excel (and other spreadsheets) only store first 15 significant digits and all others are lost. In most practical circumstances this is not noticed nor relevant. However, in cases where raw data has many significant digits this imprecision will carry over to calculated results, like in case of large areas and unrounded theoretical concentrations with a lot of significant digits (e.g. 0.173240255623636). You can test this with the following data in ValChrom (or R/Python/Java): Excel_vs_Python_vs_Java_vs_R_RSS.xslx

In most usual cases the amount of significant digits does not get large enough and/or the differences can mostly be observed in the 8th and beyond significant digit of the answer, so that very rarely in case of large unrounded numbers (slopes, RSS, etc.) can these differences be seen. In the examples of this MOOC the data provided is not meeting any of the extreme criteria, thus Excel and ValChrom will give the same answers.

NB! Later in the course you'll upload data files under the same assessment methods - ValChrom will overwrite previously uploaded data. Use action button Upload under datasets table for this purpose.



Well done! This concludes our introduction into uploading data in ValChrom.

1.6. Comparison of validation guidelines

Comparison can be found on the next page.

'Web course "LC-MS Method Validation"

University of Tartu
https://sisu.ut.ee/lcms_method_validation/

NB! "**This table is meant as a general "big picture" comparison between the different guidelines.**

The wording in this table does not necessarily exactly match the wording in the guidelines.

For the definitive wording the user must consult the original documents."

Validation parameter	FDA 2018 bioanalysis [1]	EMA 2011 bioanalysis [2]	ICH 2005 [3]	Eurachem 2014 [4]
Selectivity	Blank from at least 6 individual sources; **Acceptance Criteria:** Blank and zero: no interference at analyte and IS RTs. Spiked samples: ± 20% LLOQ. Blank: IS response < 5% of Cals and QCs average IS responses.	Blank from at least 6 individual sources; **Acceptance Criteria:** Analyte response < 20% of LLOQ; IS response <5%		Test samples and RMs Candidate and other independent methods Also test samples with suspected interferences
Specificity	Assess for interference by cross-reacting molecules, concomitant medications, bio-transformed species, etc. **Acceptance Criteria:** Same as Selectivity.		Blanks, matrix-matched samples. **Impurities, if applicable:** Spiking blanks with impurities and/or excipients. Degradation experiments: light, heat, humidity, acid/base hydrolysis and oxidation	
Carryover	Assess impact of carryover on accuracy. **Acceptance Criteria:** < 20% of LLOQ	Inject blank after a high concentration sample or calibration; **Acceptance Criteria:** Blank response: < 20% of LLOQ IS response: < 5%		

Validation parameter	FDA 2018 bioanalysis [1]	EMA 2011 bioanalysis [2]	ICH 2005 [3]	Eurachem 2014 [4]
Linearity / Calibration Curve	Matrix-matched, Blank, zero, 6 level (inc. LLOQ), **Acceptance Criteria:** LLOQ ± 20%, others ± 15%, 75% (or min. 6 calibrator levels) should meet the criteria, Cal runs need to be reproducible	Matrix-matched, Blank, zero, 3 runs x 6 levels (inc. LLOQ) (optional, 2 parallels), **Acceptance Criteria:** LLOQ ± 20%, others ± 15%, 75% (min. 6 calibrator levels) should meet the criteria 50% per level should meet the criteria	≥ 5 levels **Range:** assay: from 80 to 120 % of target conc.; content uniformity: from 70 to 130 % of target conc., dissolution testing: +/-20 % over the specified range; impurity: from reporting level to 120% of the specification;	Instrument and method working range. **Range of interest:** Blank, 6 - 10 levels evenly spaced expected range ± 10 % / ± 20 %. **Linear range:** Blank, 2-3 parallels x 6-10 levels evenly spaced **Determine if linear range is fit:** Blank, reference materials or spiked sample blanks, 2-3 parallels x 6-10 levels evenly spaced regression plot, residuals plot, regression statistics
Accuracy and Precision (A & P)	3 runs x 4 levels (LLOQ, L, M, H QC) x 5 parallels, Over several days, **Within-run and between runs** **Accuracy:** LLOQ: ± 20% from nominal conc. others: ± 15% from nominal conc. **Precision:** LLOQ: ± 20% RSD others: ± 15% RSD	**Within-run QCs** 4 levels (LLOQ, L, M, H) x 5 parallels LLOQ: ± 20% others: ± 15% (mean conc. from nominal value or RSD) **Between runs QCs** 3 runs (LLOQ, L, M, H) 2 different days LLOQ: ± 20% others: ± 15% (mean conc. from nominal value or RSD)	**Repeatability:** 3 levels x 3 parallels OR 6 determinations at target conc. **Intermediate precision:** Several days, analysts, equipment, etc. Not necessary to study effects individually. Experiment design is encouraged.	**Accuracy:** Blanks, CRMs and/or spiked samples (if RM not available). 10 parallels per level Alternatively: RM/test sample using candidate method and alternative method. **Precision:** RMs, surplus test samples or spiked sample blanks at various levels **Repeatability:** Same analyst, equipment, laboratory, short timescale. 6-15 parallels **Intermediate precision:** Different analysts, equipment, same laboratory, extended timescale.

Validation parameter	FDA 2018 bioanalysis [1]	EMA 2011 bioanalysis [2]	ICH 2005 [3]	Eurachem 2014 [4]
				6-15 parallels for each material. **Repeatability and intermediate precision in one study:** Different analysts, equipment, same laboratory, extended timescale. 6-15 runs x 2 parallels ANOVA to calculate repeatability standard deviation and intermediate precision standard deviation
Matrix effect	Matrix effects should be assessed and eliminated.	≥ 6 lots of individual blank matrices Pooled matrix should not be used. CV: ≤ 15 % at L and H levels		
Recovery	QC (L, M, H) extracted samples vs blank extracts spiked post extraction			
Stability	Stock solution, freeze-thaw, bench-top, long-term, processed sample, auto-sampler 3 parallels at L and H **Acceptance Criteria:** Accuracy: ± 15% of nominal conc.	Stock and working solution, freeze and thaw, short term, long term, processed sample, on-instrument/auto-sampler, At L and H levels **Acceptance Criteria:** Mean conc. at each level: ±15% of the nominal conc.		
Sensitivity / LLOQ	LLOQ: Lowest non-zero standard **Acceptance Criteria:** Response at LLOQ ≥ 5 x zero response A & P: ± 20% (3 runs x ≥ 5 parallels)	Lowest calibration standard **Acceptance Criteria:** LLOQ response ≥ 5 times of blank response	**LoD and LoQ:** Visual evaluation OR LoD: S/N of 3 or 2:1 LoQ: S/N of 10:1 OR Based on response SD and Slope: $LoD = 3.3 \times \frac{\sigma}{slope}$	**CC_α, CC_β** refer to EU Commission Decision 2002/657/EC and ISO 11843-2:2007 **LoD and LoQ:** Blank samples, test samples or spiked samples, concentrations of analyte close to or below the expected LOD 6 - 15 parallels $LoD = 3 \times s_{0'}$ $LoQ = k_Q \times s_{0'}$

Validation parameter	FDA 2018 bioanalysis [1]	EMA 2011 bioanalysis [2]	ICH 2005 [3]	Eurachem 2014 [4]
			$LoQ = 10 \times \frac{\sigma}{slope}$ σ determined from: 1) standard deviation of several blanks 2) calibration graph in LoD region, residual standard deviation or y-intercept standard deviation	For calculation of s_0'(modified standard deviation) refer to the guide, Other alternatives suggested as well.
Robustness / Ruggedness			System suitability parameters should be established. **Examples for study:** - stability of analytical solutions; - extraction time. - influence of variations of pH in a mobile phase; - influence of variations in mobile phase composition; - different columns (different lots and/or suppliers); - temperature; - flow rate.	Variables with significant effects must be identified, RMs or test samples, Most effective with experimental designs: e.g. Plackett-Burman experimental design for start. Rank the variables in order of the greatest effect on method performance. Significance tests to determine whether observed effects are statistically significant.
Other Validation Runs	3 QCs (L, M, H) in duplicates **Run Acceptance Criteria:** Cals: Same as calibration curve. QCs: ≥ 67% of QCs ± 15% ≥ 50% of QCs per level ± 15%			
Quality Controls (QC)	**Accuracy and Precision:** 4 lvls: LLOQ, L, M, H 3 runs x 5 parallels **Other runs:** At L, M, and H levels QCs in duplicates Nr of QCs: 5% or 6, whichever is higher **Acceptance Criteria**: ≥ 67% of QCs ± 15%	**All runs (also after validation):** Blank, zero Cals: 6 levels QC: ≥ 3 levels (L, M, H) x 2 parallels or ≥ 5%, whichever is higher **Acceptance Criteria:** **Cals:** LLOQ ± 20% Other: ± 15% 75% (or min. 6 levels) of Cals		Every batch should have QCs, stable test samples, blanks and/or standard solutions, control charts are recommended,

Validation parameter	FDA 2018 bioanalysis [1]	EMA 2011 bioanalysis [2]	ICH 2005 [3]	Eurachem 2014 [4]
	≥ 50% of QCs per level ± 15%	**QCs:** ≥ 67% of QCs ± 15% ≥ 50% of QCs per level ± 15%		
Dilution	QCs for planned dilutions, 5 replicates per dilution factor, A & P: ± 15% of nominal conc. or RSD	Cover the dilution applied to the study samples. Spiking matrix above the ULOQ and diluting with blank matrix (≥ 5 determinations per dilution factor). **A & P**: ±15% of nominal conc. or RSD		
Incurred Sample Reanalysis (ISR)	Must reanalyze samples for control: first 1000: 10% remaining: 5% **Sample selection:** C_{max} and in the elimination phase **Acceptance Criteria**: 67% ± 20% of the mean	Must reanalyze samples for control: first 1000: 10% remaining: 5% **Sample selection:** C_{max} and in the elimination phase **Acceptance Criteria:** 67% ± 20% of the mean		
Repeat Analysis	No re-analysis of individual calibrators and QCs is permitted. Reanalysis should be based on reasons At least the same number of replicates for repeats as originally tested	Example cases: Run did not fulfil the acceptance criterias, IS response significantly differing from cal. or QCs response (if criteria pre-defined), Improper sample injection or malfunction of equipment, Obtained concentration above ULOQ or below LLOQ, Analyte levels in blanks too high, Poor chromatography		
Parallels – Samples that have been taken through the entire measurement procedure (each has had independent sample pretreatment) EMA: LLOQ, L: within three times the LLOQ (low QC), M: around 30 - 50% of the calibration curve range (medium QC), H: at least at 75% of the upper calibration curve range (high QC). FDA: LLOQ, low (L: defined as three times the, LLOQ), mid (M: defined as mid-range), high (H: defined as high-range)				

Validation parameter	FDA 2018 bioanalysis [1]	EMA 2011 bioanalysis [2]	ICH 2005 [3]	Eurachem 2014 [4]

References:
1. U.S. Department of Health and Human Services Food and Drug Administration, Bioanalytical Method Validation, Guidance for Industry, 2018: https://www.fda.gov/media/70858/download
2. Guidance on bioanalytical method validation, European Medicines Agency, 2011: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-bioanalytical-method-validation_en.pdf
3. 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: https://database.ich.org/sites/default/files/Q2_R1__Guideline.pdf
4. 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): https://www.eurachem.org/images/stories/Guides/pdf/MV_guide_2nd_ed_EN.pdf

2. Selectivity and identity confirmation

In this chapter we will rationalize the 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 mainly in matrix and consequently 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)(1) Instead of the term (analytical) method used in this course, IUPAC uses the term (analytical) procedure. (method). As the definition implies, methods can be selective to a 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.

An 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 an analyte and yield the signal indistinguishable from an analyte's signal. For example, in case of LC-UV analysis, an interferent would have the retention time very similar to that of an analyte and would absorb UV-radiation at the same wavelength as an analyte.

In the case of chromatographic methods it is very common to assess selectivity regarding chromatographic separation of compounds. The ability of the 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) – α. For the two peaks on the chromatograms in Figure 1, a separation factor is calculated as follows (follows (t_{RA} t_{RB} are retention times of, respectevely, compounds A and B; t_M is hold-up time – time for an unretained compound to reach the detector):

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

The larger 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 the peaks in chromatogram 2 are not completely resolved. This is the main drawback of a separation factor as numerical characteristic of chromatographic separation – it only characterizes peak separation in terms of retention times, but not peak resolution in terms of lack of overlap.

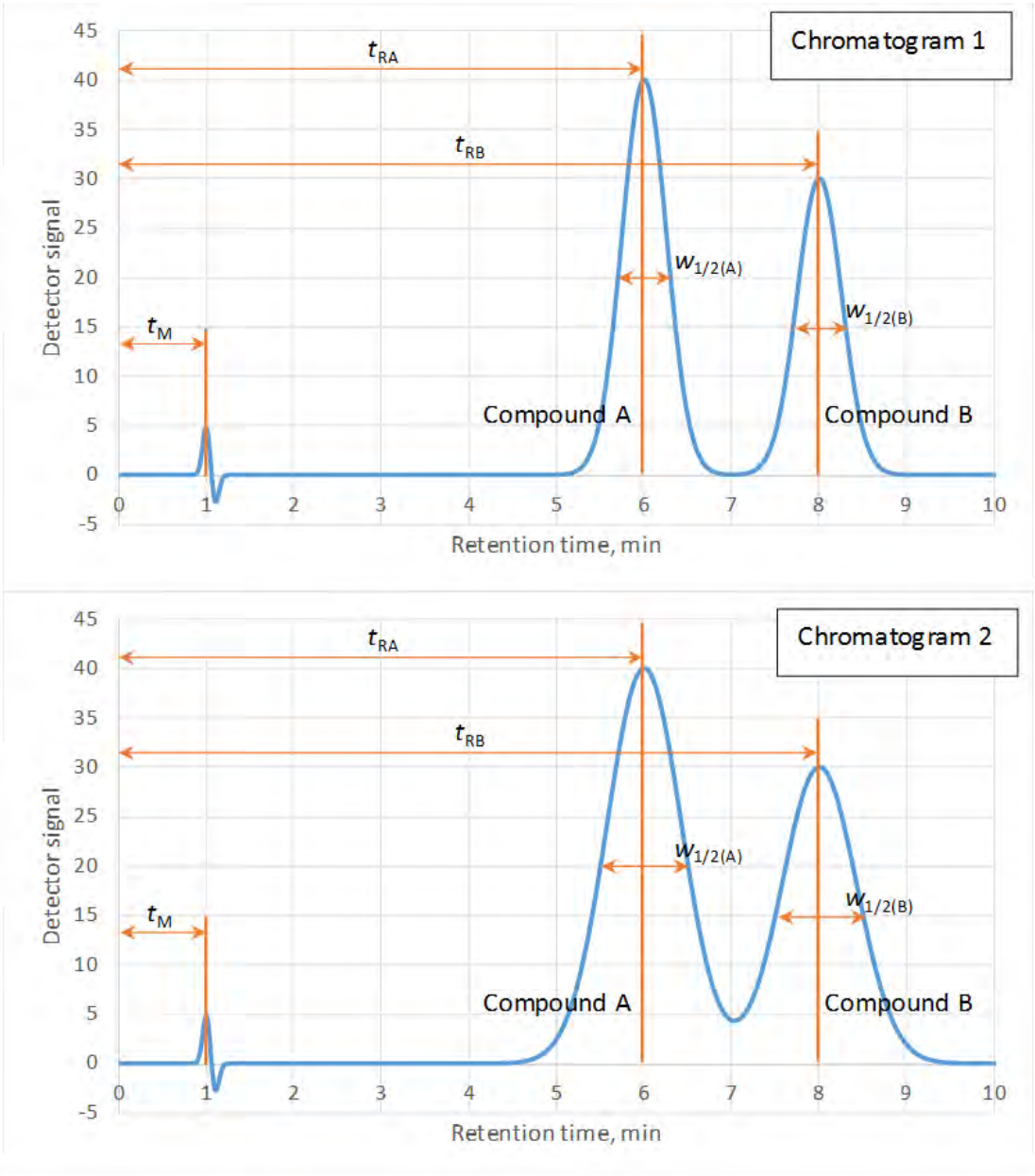


Figure 1. Chromatograms with a similar separation factor but different resolution.

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

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

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 demonstration of suitable peak resoluton, AOAC (also ICH) requires that no other compound should be detectable at the analyte retention time, when methods like IR, NMR or MS are used. To check for the presence or absence of coeluting compounds, Eurachem suggests a 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

Note that current version of FDA guideline does not directly specify resolution R_S.

2_1_intro_selectivity_identity.pdf 495 KB

Selectivity.
Confirmation of identity.

Introduction – selectivity

- Any analytical method must produce analytical signal in response to the presence of analyte in the sample.
- The signal should be unaffected by other substances in the sample.
- The extent to which this condition is met is termed **selectivity**.

Introduction – identity confirmation

- When an analytical signal is registered from the sample, it must be proved that the signal is due to analyte – **confirmation of identity**.
- Information required for identity confirmation is collected during validation.

LC-MS is one of the most selective analytical techniques, which also enables identity confirmation.

Selectivity

- IUPAC definition of selectivity:

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

- **Specificity** – 100% selectivity.

IUPAC. Compendium of Chemical Terminology, 2nd ed. (1997) Compiled by A. D. McNaught and A. Wilkinson.

Selectivity terms in validation guidelines

„Selectivity“	„Specificity“
AOAC	ICH
EMA	NordVal
Eurachem	
FDA	
IUPAC	

Selectivity

- IUPAC definition of selectivity:

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

- **Interference** …
  - substance behaves like analyte.
  - interferes with analytical signal (matrix effect).

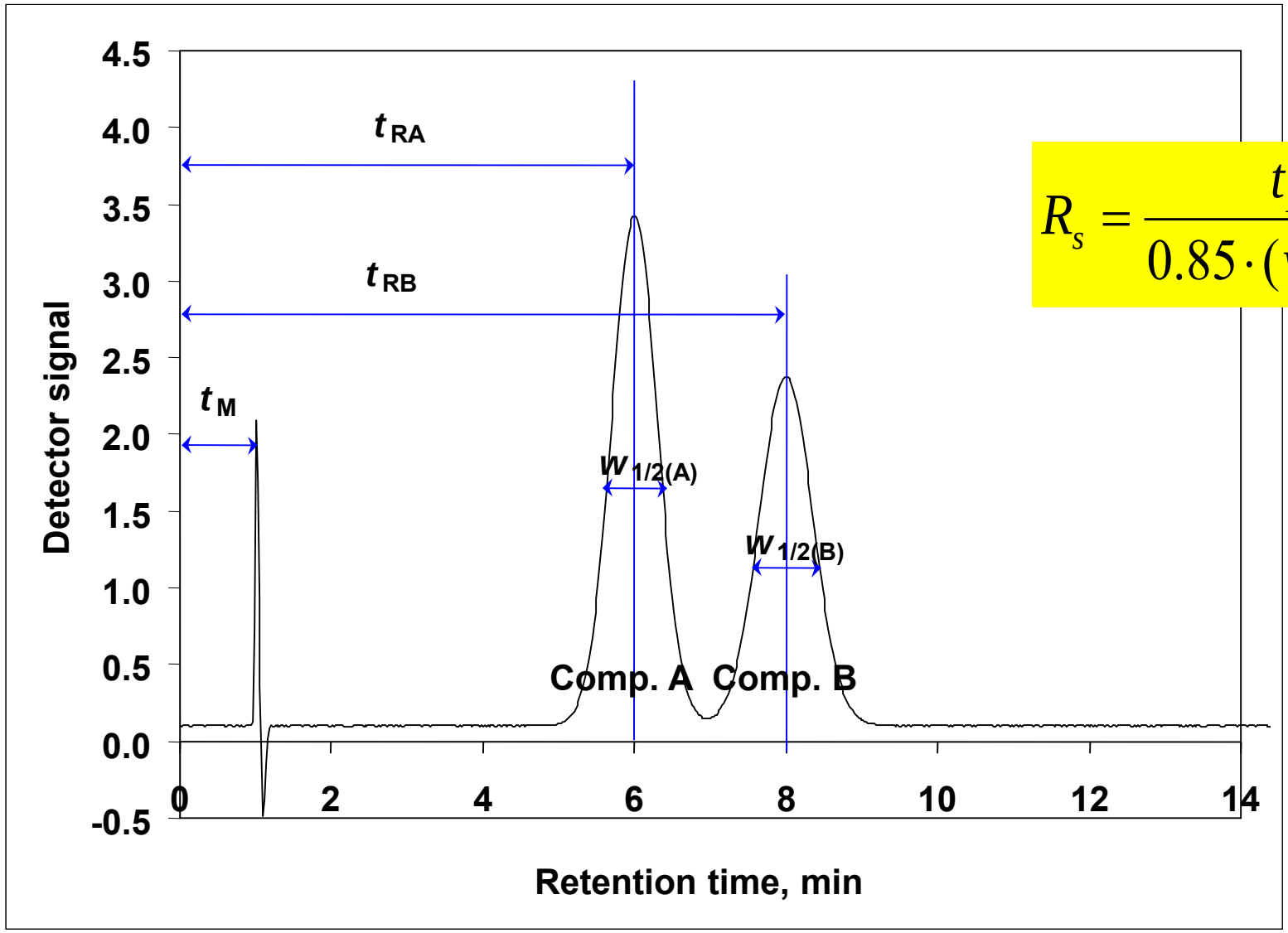
IUPAC. Compendium of Chemical Terminology, 2nd ed. (1997) Compiled by A. D. McNaught and A. Wilkinson.

Selectivity in LC-MS

- Selectivity arises from both:
  - LC separation process.
  - MS detection.

LC selectivity

Peak resolution



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

Required LC selectivity

- Selectivity in terms of peak resolution:
  - FDA: $R_S \geq 2$
  - AOAC: $R_S \geq 1.5$

- Eurachem: demonstrate separation on different column.
- AOAC and ICH: no other compound should be detectable at analyte retention time when other methods (like IR, NMR and MS) are used.

Why are the requirements so different?

- Pharmaceutical analysis.
  - In active pharmaceutical ingredient (API) all synthesis by-products, degradants and additives may cause side-effects on patients.
  - Therefore each component of the sample appears as an analyte and must be chromatographically separated.
- Pesticide residue analysis in vegetables.
  - Vegetable extract is rich in compounds, which may co-elute with analyte.
  - Analysis results can still be adequate, if **detector** is only responsive to analyte and not matrix components.

2.2. Selectivity: detector-side (MS) selectivity

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

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

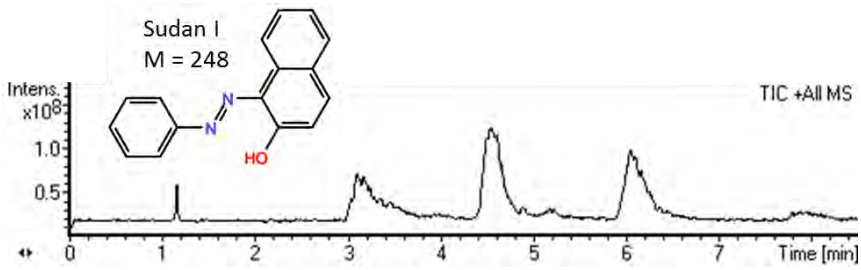


Figure 1. Total ion chromatogram (TIC) of Sudan I dye analysis.

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 the 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.

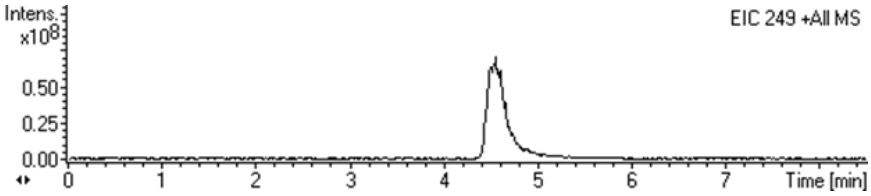


Figure 2. Extracted ion chromatogram (EIC) of Sudan I dye analysis. The *m/z* value 249 correspond to protonated Sudan I molecule $[M+H]^+$.

What if compounds of identical or very close *m/z* are analysed? In this case there are three approaches. The best (and cheapest) would be the 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 analyses, at least in complex sample matrices, are 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

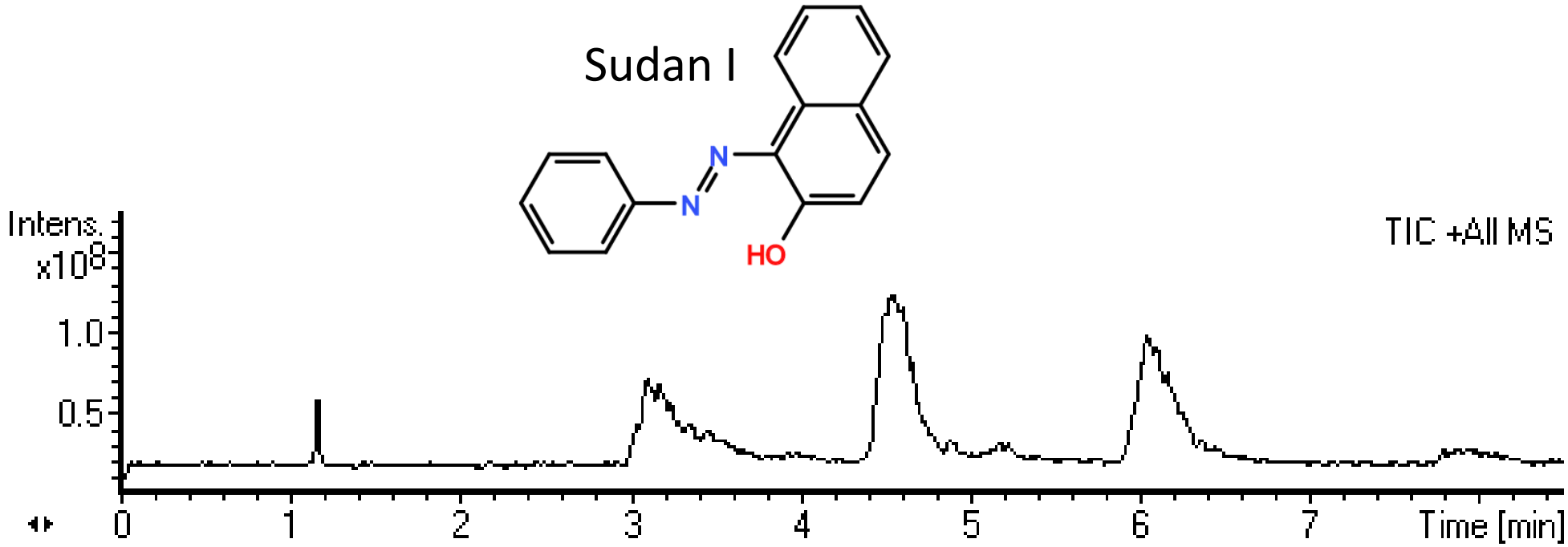
Detector-side selectivity

MS selectivity

- MS detection adds new dimension to the analytical signal.
- m/z dimension enables additional selectivity.

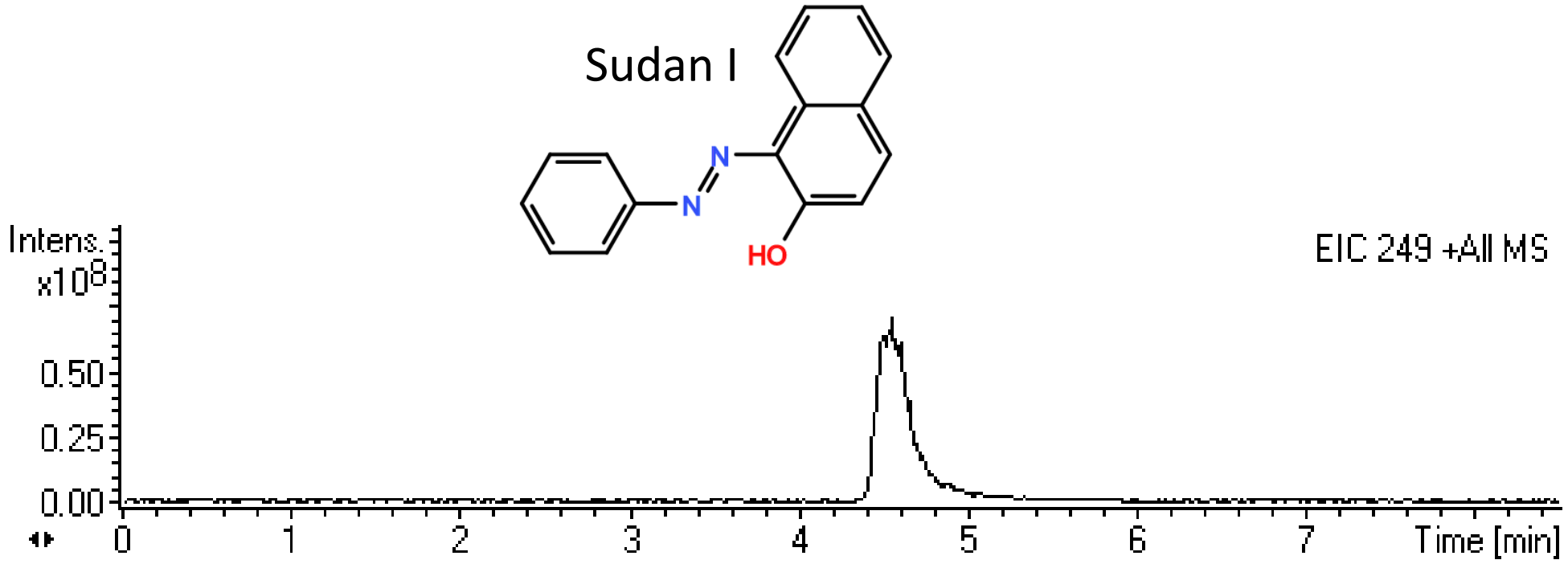
Analysis of Sudan I dye – TIC

- Total ion chromatogram (TIC) – all sample components are “visible“.



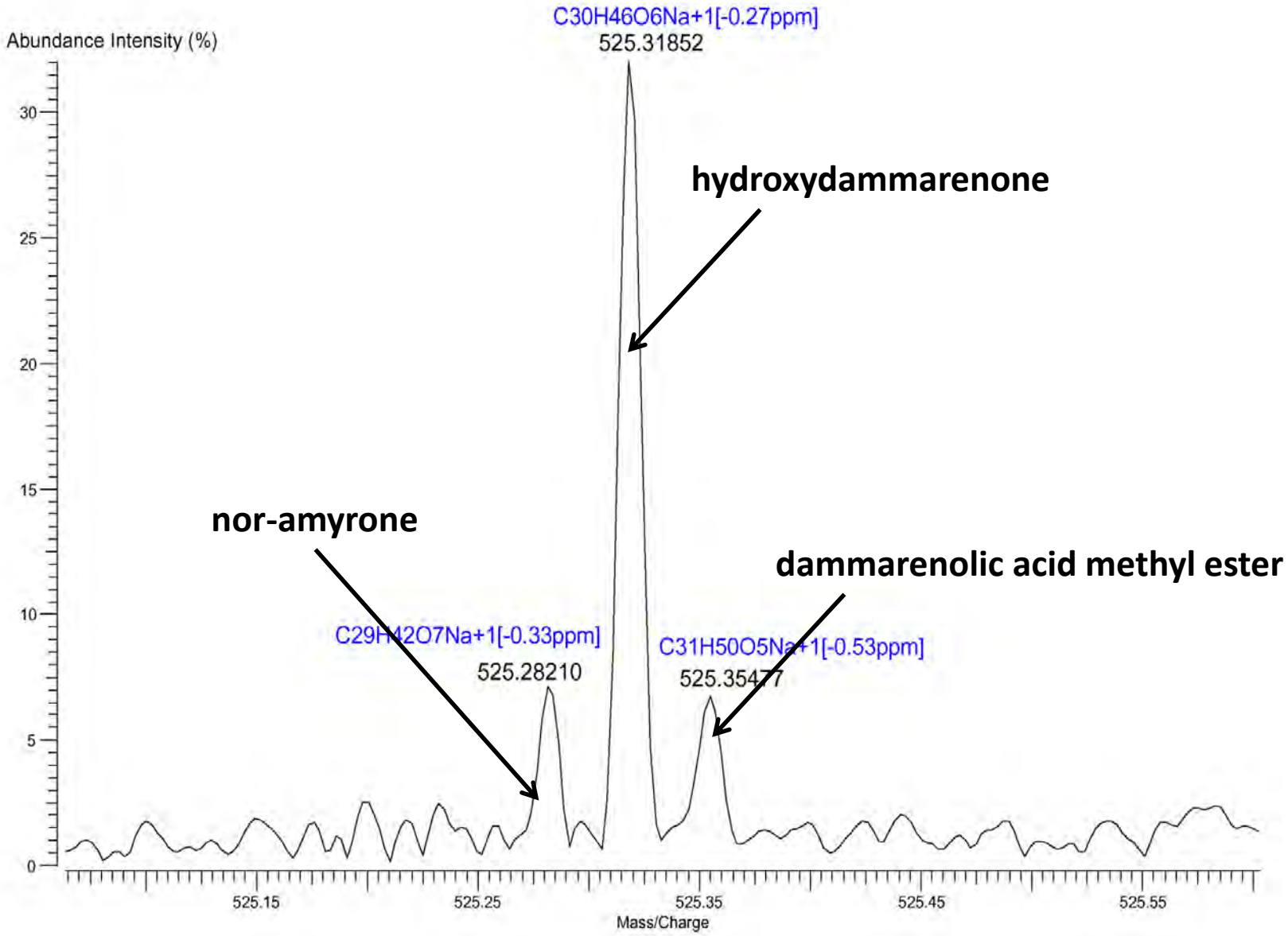
Analysis of Sudan I dye – EIC

- Extracted ion chromatogram (EIC) – only compound(s) with m/z 249 are registered.



MS selectivity

- Even higher selectivity can be obtained:
  - Tandem MS (MS/MS).
  - High resolution MS (HRMS).



Could MS be used without LC?

- In principle “yes“, BUT …
  - Mixture components may yield ions with the same m/z as the analyte.
  - Mixture components may alter analyte signal (matrix effect).

LC and MS selectivities are orthogonal to each other.
Highest selectivity is achieved if LC and MS are combined.

2.3. Selectivity examples

Example 1

Not all validation guidelines explicitly specify a required limit for chromatographic peak resolution (R_S). In LC-MS an incomplete chromatographic separation of analyte peak from neighbouring peaks may or may not cause problems. Often *m/z* of analyte differs from that of neighbouring peaks or an 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 a peak resolution of at least $R_S \geq 1.5$.

The chromatogram in Figure 1 was recorded while developing a method for analysis of compounds A, B, C and D. One MRM transition is presented for each compound and is shown in different colour. If one hadn't recorded the transition shown in green, this peak would have been missing from the chromatogram and one wouldn't have known that compound C (RT 5.33 min) – is present. But this compound could interfere with the 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 other peaks.

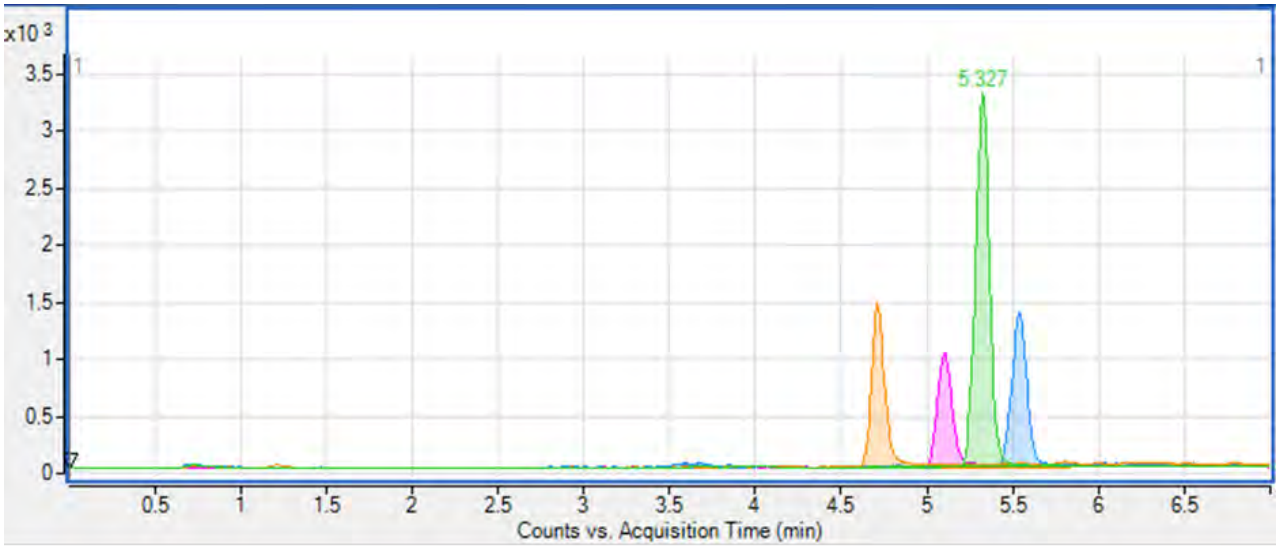


Figure 1. Initial chromatographic separation of compounds A, B, C and D.

Table 1 presents the chromatographic data for above separation. The peak resolution (R_S) from the preceding peak was calculated for the last 3 peaks according to equation 1. It appears that while the separation of the compounds B and C is at the limit (R_S= 1.5), then the 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

The chromatographic separation was further optimized (1) There are several parameters that can be varied for selectivity optimization in reversed phase LC:

- Percent of organic solvent in the mobile phase
- Organic solvent itself (methanol, acetonitrile, THF or their mixture)
- pH of the eluent (note the pH limits of your column; in case of LC-MS only volatile pH-buffer components can be used)
- Different column chemical and physical properties
- Temperature (the effect is usually small, but may help) and the chromatogram presented in Figure 2 was obtained. As a result, sufficient peak resolution (R_S > 1.5) has been achieved for all of the analytes (Table 2).

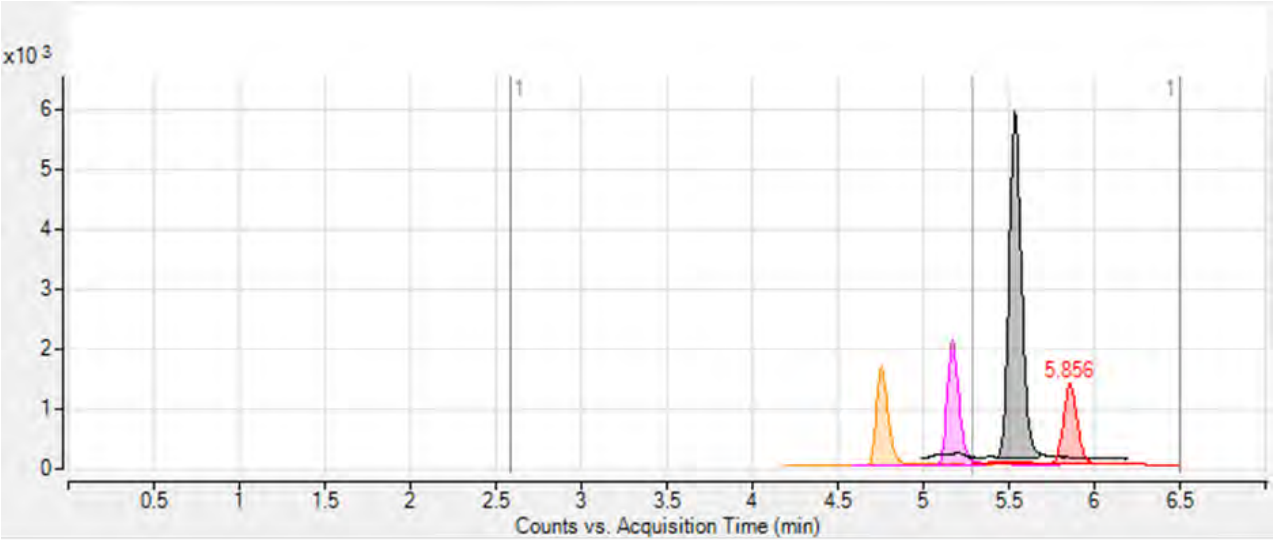


Figure 2. Chromatographic separation of compounds A, B, C and D.

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 the chromatographic separation. 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).

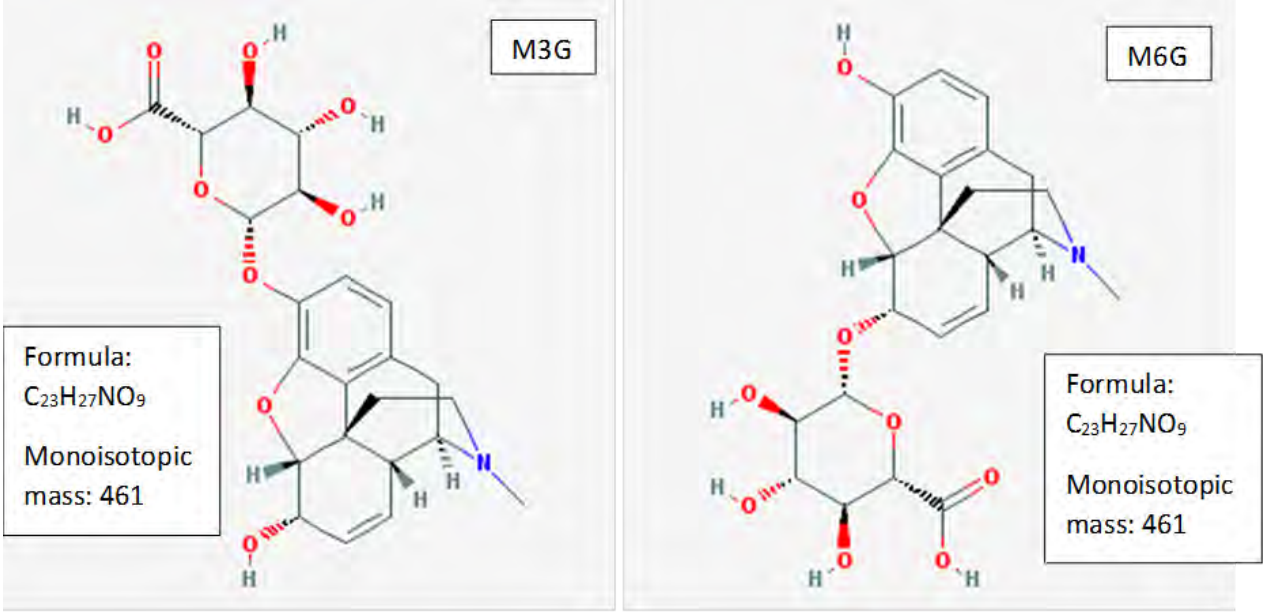


Figure 3. Morphine-3-glucuronide (M3G) and morphine-6-glucuronide (M6G).

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

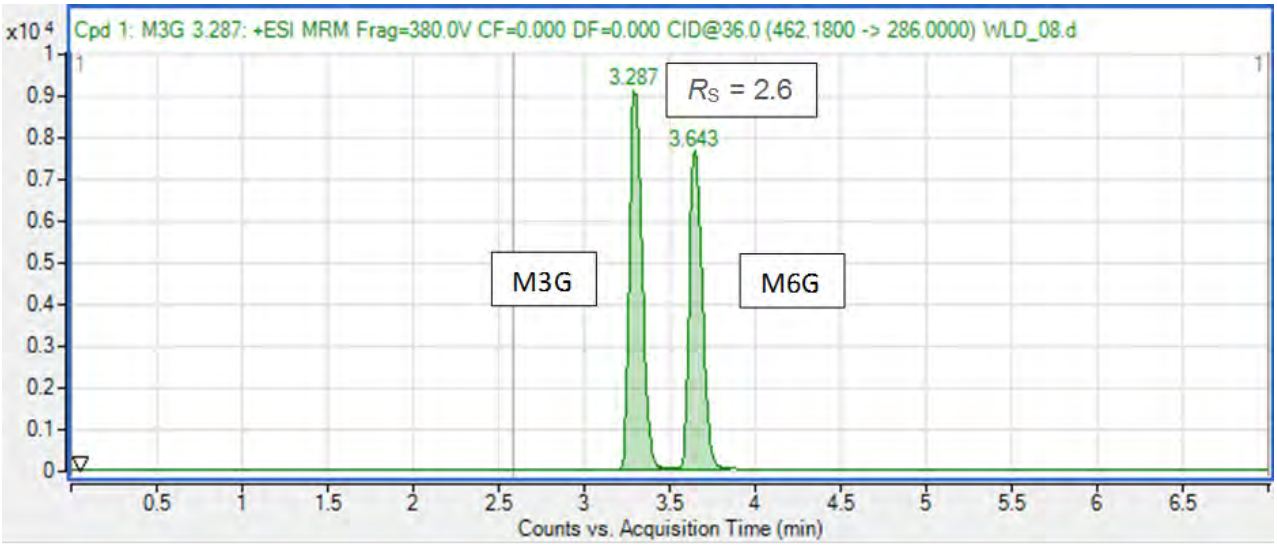


Figure 4. Chromatographic separation of morphine metabolites Morphine-3-glucuronide (M3G) and morphine-6-glucuronide (M6G).

***

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

- Percent of organic solvent in the 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 chemistry and physical properties
- Temperature (the effect is usually small, but may help)

2.4. Introduction to identity confirmation

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

The following evidence is used in chromatographic methods for confirming the identy of an analyte:

1. Chromatographic retention time. The basic requirement in chromatography is that the retention time of the compound must be the same in the standard solution and in the sample. But similar retention time alone does not confirm the identity of the analyte 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 the structural features of compounds, but UV-Vis and fluorescence spectra are usually not sufficiently characteristic to the analytes for their positive identification. Mass spectrometric techniques, especially high resolution MS and tandem MS (MS/MS), are the most powerful methods 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 and summarizes general requirements instead of considering 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 an analyte. Proper selection of potentially interfering compounds requires scientific judgement, experience and knowledge.

The following two sections look in more detail at the 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

Confirmation of identity

Introduction

- Millions of compounds exist.
- Several compounds have chromatographic behavior similar to the analyte.
- **Identity confirmation** provides evidence that the analytical signal registered from the sample is due to the analyte.
- Information required for identity confirmation is collected during validation.

Ways of confirming identity

- Comparison of retention times.
  - Analyte retention time should be equal in sample and in standard solution.
  - But several structurally different compounds can have similar retention times.

Ways of confirming identity

- Spectroscopic methods.
  - Provide evidence that chemical structure of the detected compound is the same as of the intended analyte.
  - On-line (as chromatographic detectors) or off-line.
  - Fluorescence and UV-Vis absorption spectra provide some evidence, but usually not enough for identity confirmation.
    - But enable exclusion – if spectra are different, then the compound can't be the analyte.
  - Mass spectrometry is the most powerful method for identity confirmation.

Identity confirmation by ICH guidelines

- Demonstrate that the method is able to discriminate between compounds of closely related structures.
  - Positive results from samples containing the analyte (true positive).
  - Negative results from samples, which don’t contain the analyte (true negative).
  - Negative results from compounds, which are structurally similar or closely related to the analyte (true negative).
- Consider, which compounds could potentially interfere – experience.
- Scientific judgement of properties of the sample components, analyte and analytical method – knowledge.

2.5. Identity confirmation: retention time

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

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

According to the most recent SANTE/SANCO validation guideline, retention times of an analyte in a sample and in a standard solution must not differ more than by 0.1 min (1)(1) The previous SANCO guideline established the retention time tolerance as ± 0.2 min [ref 4]. . 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 an analyte retention time relative to the standard substance retention time. The relative retention time of an analyte in a standard solution must correspond to that in a sample within tolerance of 2.5%.

The shape of the chromatographic peak may also be used to evaluate identity of an analyte. If peaks in a standard solution and in a sample differ in shape, it may indicate that the compound detected in the sample is not the analyte or that some compound coelutes with an analyte. It is possible that the analyte peak is distorted in a sample because of the sample matrix. In order to avoid 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 the peak shape of an analyte and a 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_0.pdf 367 KB

Identity confirmation by LC

LC data for identity confirmation

- LC data alone is not sufficient for analyte identity confirmation.
- Without adequate chromatography the identity confirmation can’t be completed.

- Identity confirmation data from LC:
  - Retention time.
  - Peak shape.
  - Altered chromatographic conditions.

Retention time

- The retention time of the analyte must be at least twice the system dead time, i.e. retention factor $k \geq 1$. (SANCO, 2002/657/EC)
- The retention times of the analyte in sample and in standard solution must not differ by more than ± 0.2 min. (SANCO) ±0.1 min according to new version.
- The ratio of the retention times of analyte and internal standard must correspond to that of standard solution within a tolerance of ± 2.5%. (2002/657/EC)

Peak shape

- The peak shape of the analyte in sample must correspond to the peak shape of analyte in standard solution.
- The peak shape of isotopically labelled internal standard shall match the peak shape of the analyte.

- To account for the matrix-dependent changes in the peak shape and/or retention time of the analyte use …
  - Matrix matched standard solutions.
  - Isotopically labelled internal standards.

Altered chromatographic conditions

- When chromatographic conditions are altered, the retention times and peak shapes of the analyte in sample and in standard solution should remain similar.
- Chromatographic conditions to alter:
  - Different column (ideally, different separation principle).
  - Solvents (including pH variation).
  - Gradient.
  - Temperature.

2.6. Identity confirmation by MS

For identity confirmation, the MS provides *m/z* values of ions formed from the 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 the sample analysis. Preferably, reference spectra should be recorded within the same analysis batch as the 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 which lose (eg H_2O or NH_3) are of little diagnostic value.
- The choice of diagnostic ions depends on matrix interferences.

The usage of principles outlined in this section is illustrated in section 2.7.

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

Identity confirmation by MS

MS data for identity confirmation

- MS provides the most valuable information for analyte identity confirmation.
- Identity confirmation data from MS:
  - *m*/*z*-s of quasimolecular ion, adduct ions and product ions.
    - *m*/*z*-s from high resolution (high mass accuracy) MS are particularly valuable.
  - Ion intensity ratios.

- With respect LC-MS analysis validation guides by SANCO and 2002/657/EC are most specific – give guidelines and set criteria.

Requirements for mass spectrometry

- Reference spectra for the analyte.
  - Use the same instrument and operating mode as for samples.
  - Preferably, record within the same analysis batch with samples.

- Diagnostic (characteristic) ions.
  - The quasimolecular ion should be involved in identification procedure.
  - 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.

Requirements for mass chromatograms

- Extracted ion chromatograms for the analyte.
  - Analyte peaks in sample should have similar retention time, peak shape and response ratio as in calibration standard.
  - Chromatographic peaks on extracted ion chromatograms for the same analyte must overlap.
  - For quantitation use the ion that shows the highest S/N ratio and no chromatographic interference.

Ion intensity ratios

- In addition to presence of specific ions in mass spectrum, also their relative intensity must match that of standard substance.
  - 2002/657/EC tolerance limits for relative ion intensities.

Relative intensity (% of base peak)	LC-MS, LC-MS^n (relative)
> 50%	± 20%
> 20% to 50%	± 25%
> 10% to 20%	± 30%
≤ 10%	± 50%

  - SANCO sets tolerance limit of ± 30% (relative) independent of peak intensity.

Required number of ions

- Number of ions required for analyte identification depends on used MS technique and operating mode.
  - For example, identification criteria of SANCO.

MS mode	Identification requirements
Single stage MS (low resolution)	≥ 3 diagnostic ions
Single stage high resolution MS	≥ 2 diagnostic ions; at least one fragment
MS/MS	≥ 2 product ions

  - For all diagnostic/product ions criteria for ion intensity ratios must be met.

Required number of ions

- 2002/657/EC uses system of identification points.
  - 4 points for substances having anabolic effect and unauthorized substances.
  - 3 points for veterinary drugs and contaminants.
- Number of identification points earned depends on MS technique.
  - For example:

MS technique	Identification points earned per ion
MS^n precursor ion	1.0
MS^n transition products	1.5
High resolution MS^n precursor ion	2.0
High resolution MS^n transition product	2.5

Example of identification points

- Calculate the number identification points earned for LC-MS/MS method with 2 transitions for an analyte.
  - 1 precursor ion → 1 point.
  - 2 transition products, 1.5 points each → 3 points.
  - Total identification points earned: 4.

2.7. Identity confirmation examples

Example 1

It is the basic assumption of liquid chromatography, that the retention time of an analyte is the same on the chromatograms of standard solutions and sample solutions. Two validation guidelines have set a clear criteria for the retention time tolerance: 2002/657/EC [ref 5] and SANTE/SANCO [ref 4]. 2002/657/EC limits the tolerance of the 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 the sewage sludge compost was carried out. The retention times of the three analytes in the calibration solutions are presented in Table 1 along with the 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

The average retention times in calibration solutions are used as reference points to assess the deviations of retention times of analytes in sample solutions. Table 2 presents experimental retention times of analytes in sewage sludge compost extracts. The difference of a retention time from the respective reference value is calculated. Absolute values 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 the situation is different with metformin – only two results comply. A 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, the sample matrix affects the retention. Matrix matched calibration could then solve the problem. It is also possible, that the contamination gradually builds up in the column and alters the stationary phase properties and consequently the retention – more extensive clean-up of the 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.

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

Table 3. Absolute (*RT*) and relative (*RRT*) experimental retention times obtained from calibration solutions. Metformin is used as an 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 all analyte peaks in samples (Table 4). To compare *RRT* values in sample to a reference value, a relative difference (as percent) is calculated as follows:

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

where *RRT* is a relative retention time of an analyte in a sample injection and RRT_{Ref} is an 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 compliant. Only retention times of analytes in the last sample were found to be non-compliant. A reason for this is the retention time of the IS, which is smaller than expected.

Compared to tolerance limits set by SANTE, these of 2002/657/EC appear to be more lenient. IS 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 the monitored ions in mass spectrometry are presented in Table 3 of SANTE/SANCO validation guidelines. For example, the 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 3 in SANTE/SANCO guidelines [ref 4].

The mass spectrometric identification rules used by 2002/657/EC [ref 5] are more elaborate and based on the identification points. The required number of identification points depends on the type of the analyte (see [ref 5] for full details): 4 points are required for substances having anabolic effect and unauthorized substances; 3 points are required 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 the "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 an ion trap MS (unit resolution) recording 2 product ions – 5.5 points (1.0+1.5+2*1.5=5.5).
- 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%

The ion ratio is calculated as an intensity (or peak area) ratio of a less intense ion to that of a more intense ion. The reference ion ratio value is calculated as an average of ion ratios of calibration solutions. Table 7 illustrates the process of calculating such a 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
Cal. solution 1	555	259	0.47
Cal. solution 2	1176	499	0.42
Cal. solution 3	1707	803	0.47
Cal. solution 4	2404	991	0.41
Cal. solution 5	3031	1312	0.43
		Average:	0.44

In a similar manner, ion ratios are calculated for the 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 solution 1	821	281	0.34
Sample solution 2	2221	1251	0.56

To assess the 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 the identity of propamocarb in both samples is confirmed by the rules of 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 a proper selection of the diagnostic ions. In general, fragmentation due to the 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 the loss of NH_3, which leads to a noisier background.

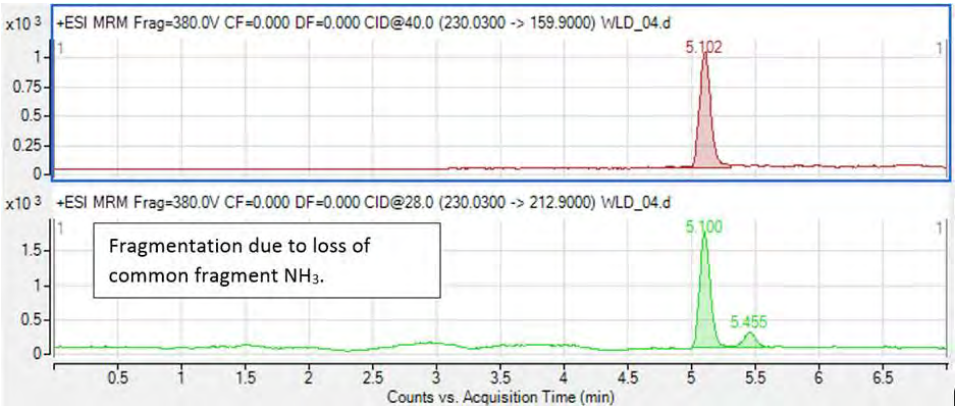


Figure 1. Influence of fragmentation pathways on chromatographic background.

Example 5

How well do all the limitations work?

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

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

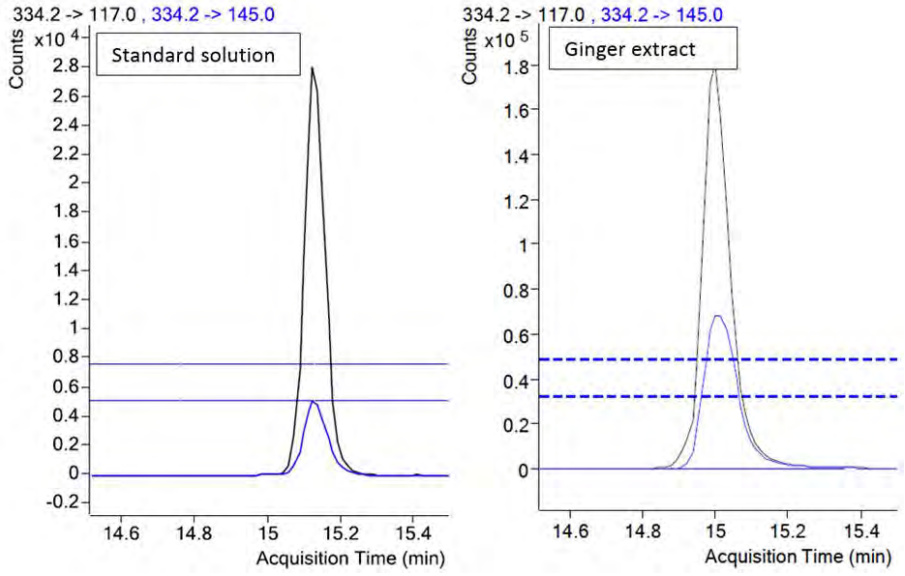


Figure 2. Chromatograms of two transitions of tebufenpyrad in a standard solution and in a sample. (Data from http://www.agilent.com/cs/library/eseminars/public/Triggered%20MRM%20for%20the%20Analysis%20of%20Pesticides.pdf)

Retention times of peaks are similar: 15.1 min in a standard and 15.0 min in an extract and would pass the requirements for the retention time match by SANTE (0.1 min). The ion ratio of the two transitions for a tebufenpyrad in a calibration standard is 0.21 and for peaks in a ginger extract 0.469. The relative difference of ion ratios is 123%, which exceeds the tolerance limits set by SANTE/SANCO and 2002/657/EC. Therefore, the presence of a tebufenpyrad in a 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 the 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

Note: Gradient of calibration graph is more commonly termed as slope of calibration graph

All section 3 slides are downloadable in one file:

3_linearity.pdf 6.7 MB

Linearity, linear range, sensitivity

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.

Sensitivity

The change in instrument response which corresponds to a change in the measured quantity.

The **gradient** of the response curve.

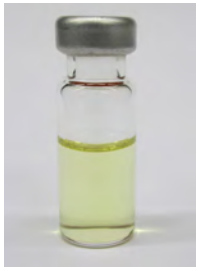
Experiment planning for evaluation of linearity

- Type of calibration samples
- Concentration range
- Number of concentration levels
- Measurement protocol

Type of calibration samples

- Standard solutions
  - matrix-free

- Matrix-containing samples
  - Matrix-matched calibration
  - Blank matrix extract (preferably the same type as the sample)



Concentration range

- Appropriate for the method
  - Expected working range ± 10%…20%
  - 70%…130% of expected analyte concentration
  - Blank sample

Number of concentration levels

- Minimum 6 concentration levels are acceptable
  - Suggested 10
  - Evenly placed
  - Prepared from independent dilutions

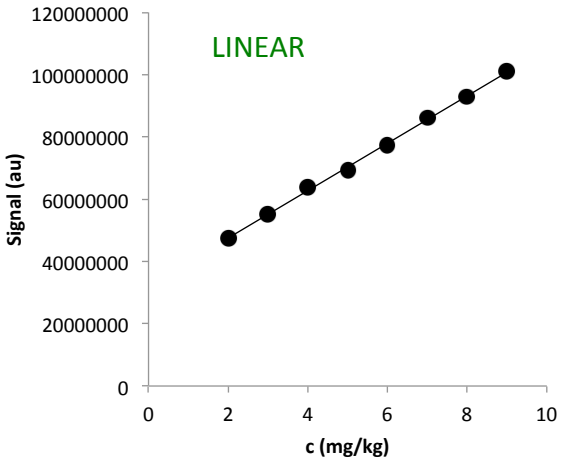
Measurement protocol

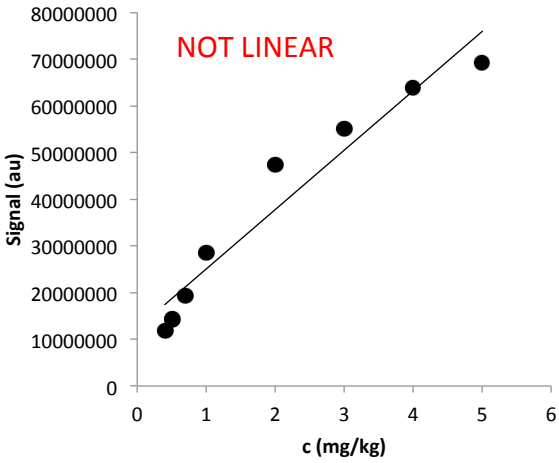
- As similar as possible to the real life situation
  - Random order
  - Between the samples
- Analysed at least twice

Evaluation of linearity

- Visual evaluation
- Residuals

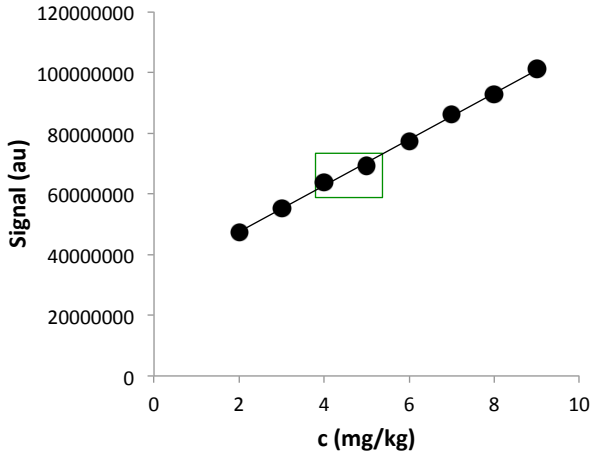
Visual evaluation

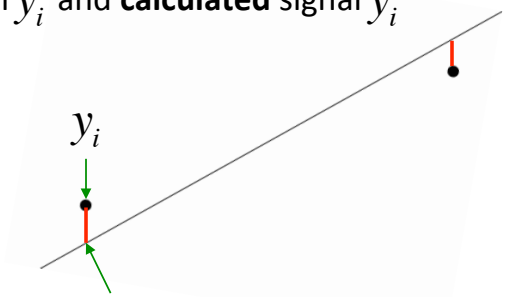




Residual

- Difference between **experimental** signal y_i and **calculated** signal $\hat{y}_i$

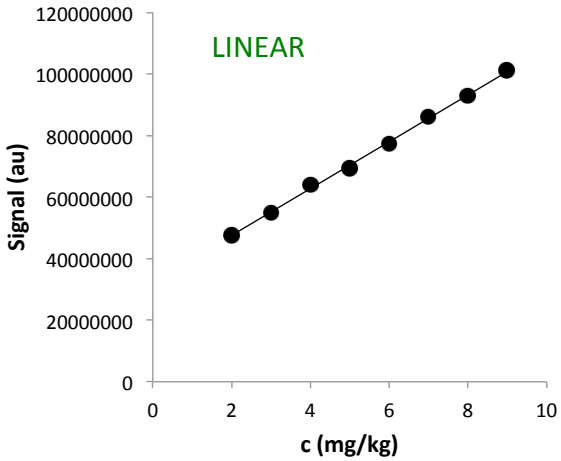


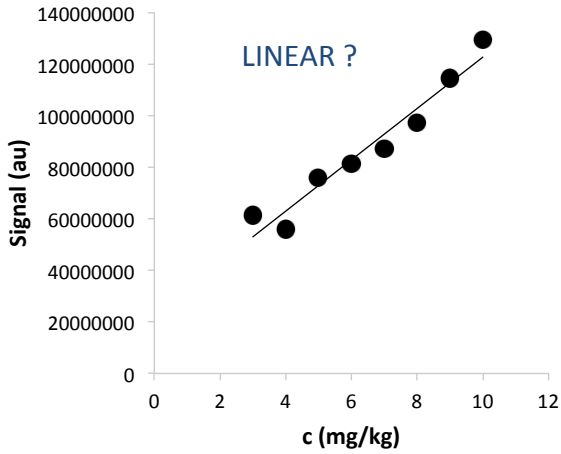


$$\hat{y}_i = slope \cdot x_i + intercept$$

$$\varepsilon_i = y_i - \hat{y}_i$$

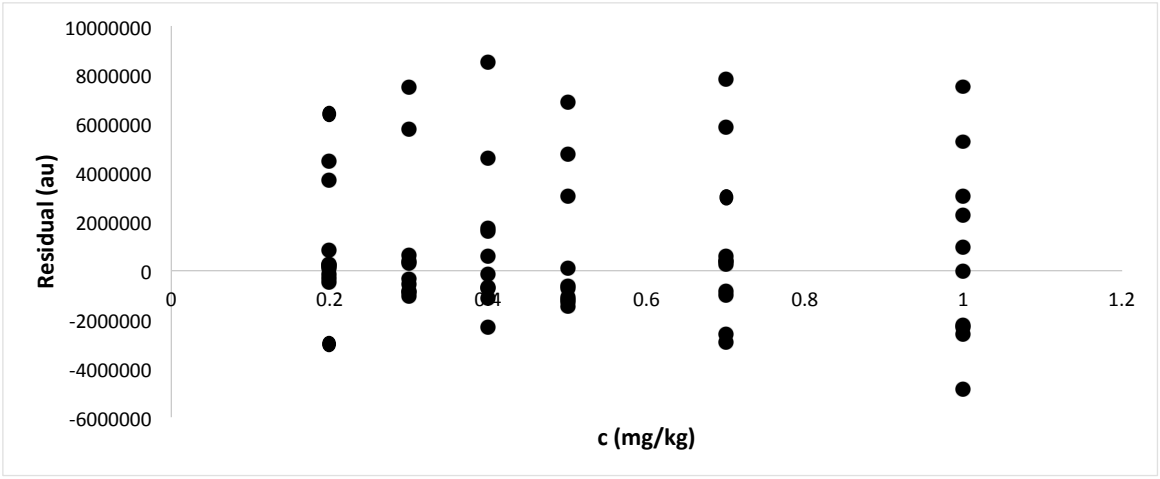
Visual evaluation



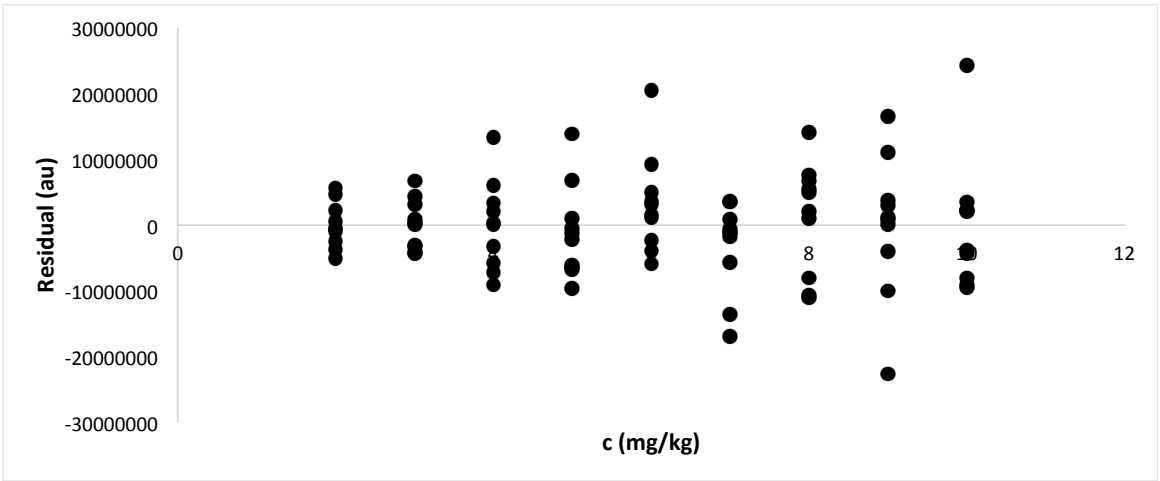


Plot of residuals

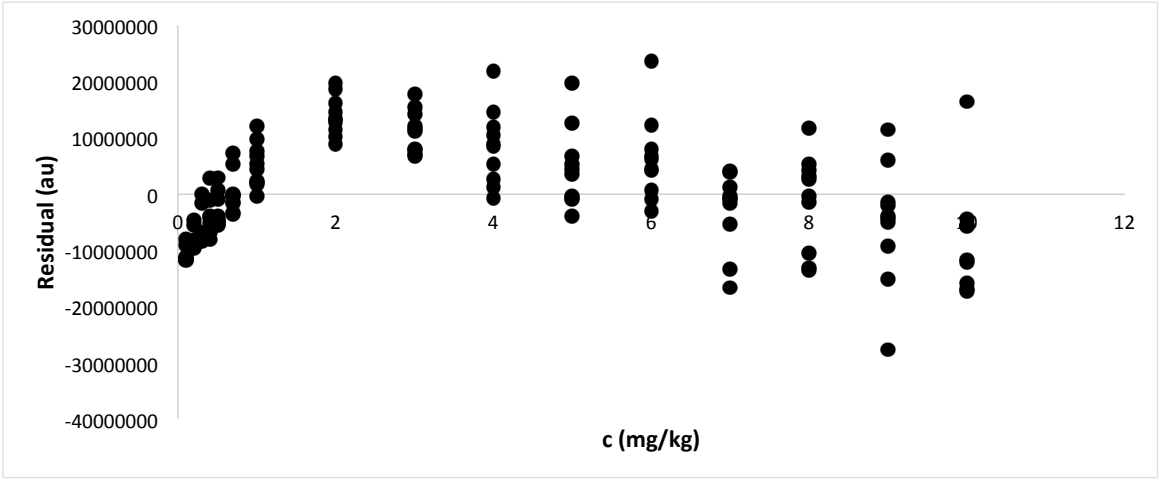
Linear, homoscedastic



Linear, heteroscedastic



Not linear, heteroscedastic



Relative residuals

- ±20% is acceptable

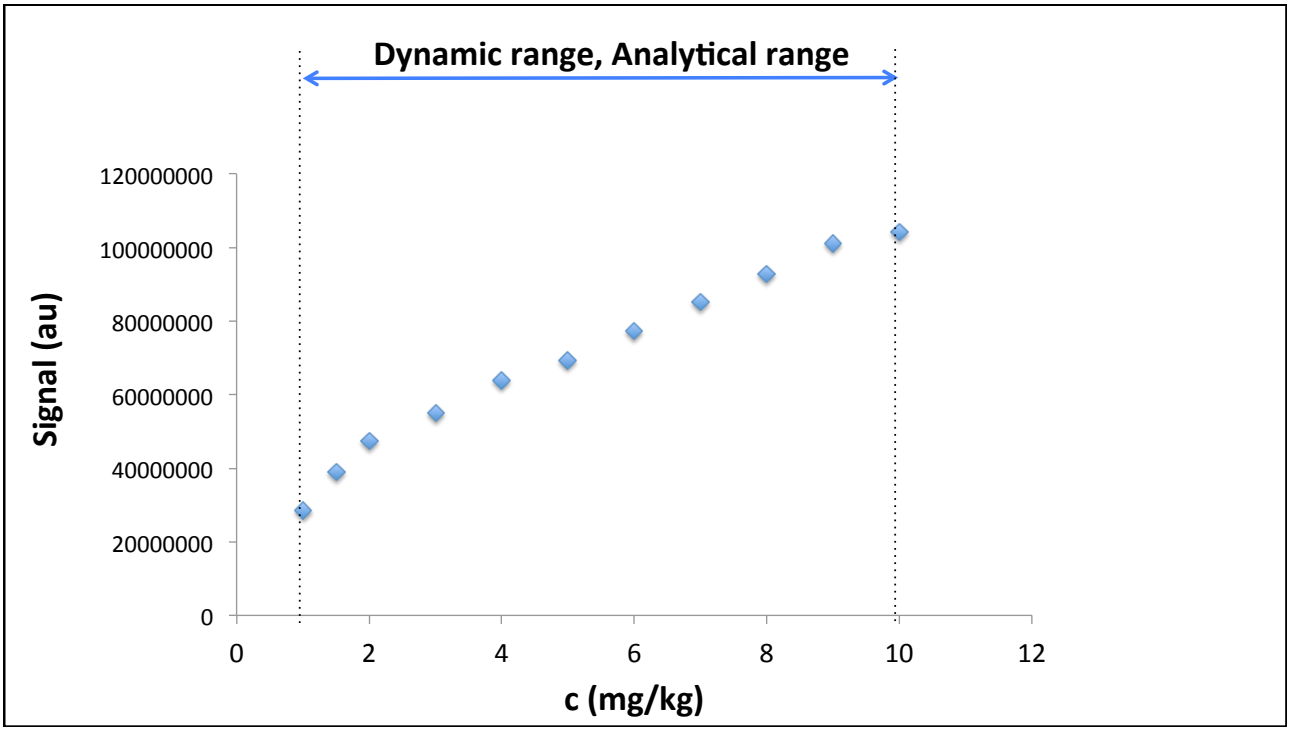
$$Y_i = \frac{y_i - \hat{y}_i}{\hat{y}_i}$$

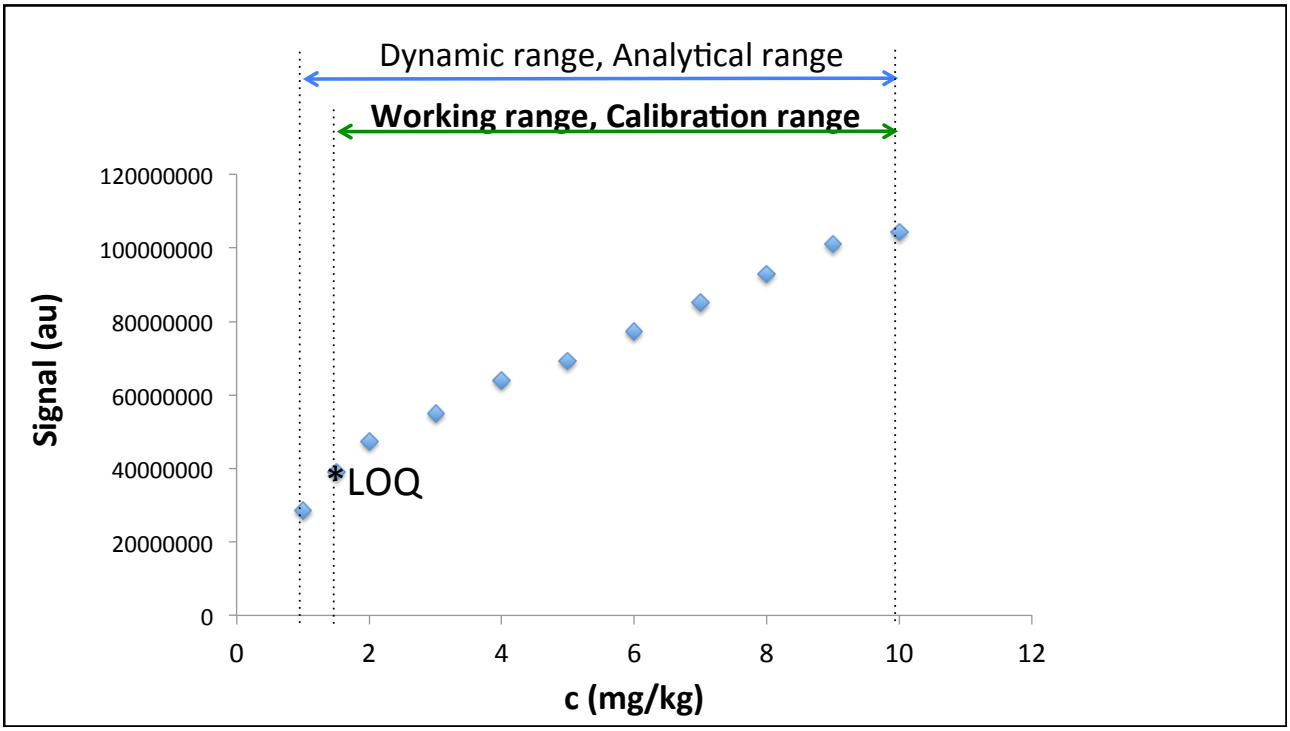
Correlation coefficient *r*
Coefficient of determination r^2

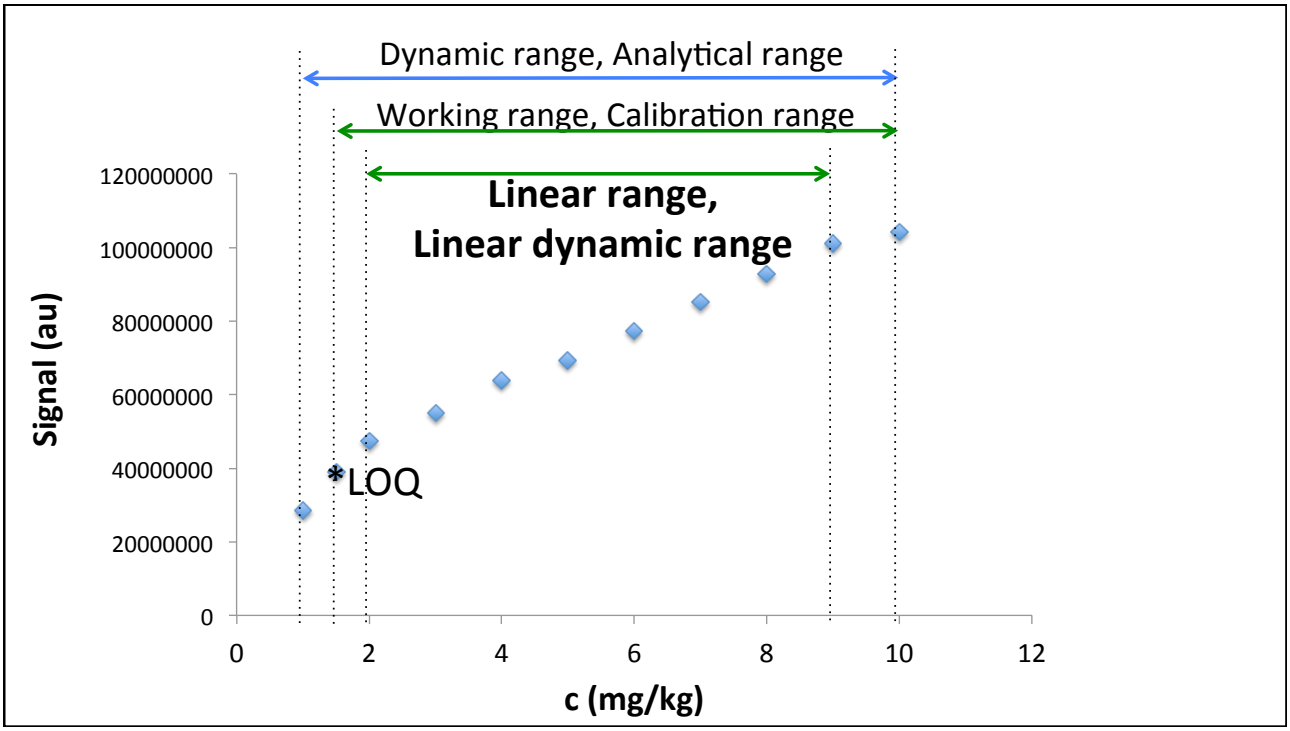
- Total variability in the response that is accounted for by the model
- **Not a measure of linearity!**

Linear range

“range”


Dynamic range, Analytical range
120000000
100000000
80000000
60000000
40000000
20000000
0
Signal (au)
0
2
4
6
8
10
12
c (mg/kg)

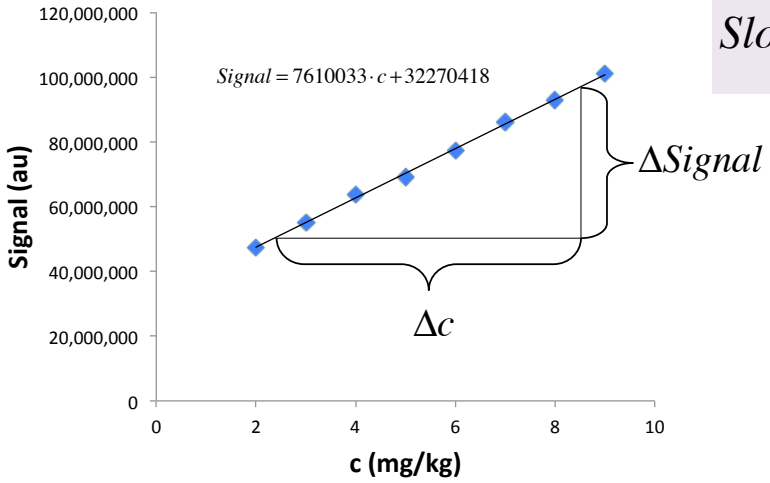

Dynamic range, Analytical range
Working range, Calibration range
120000000
100000000
80000000
60000000
40000000
20000000
0
Signal (au)
*LOQ
0
2
4
6
8
10
12
c (mg/kg)


Dynamic range, Analytical range
Working range, Calibration range
Linear range,
Linear dynamic range
120000000
100000000
80000000
60000000
40000000
20000000
0
Signal (au)
*LOQ
0
2
4
6
8
10
12
c (mg/kg)

Sensitivity

- gradient of the calibration graph

Sensitivity



$$Slope = \frac{\Delta Signal}{\Delta c}$$

Sensitivity

- Method optimization
- Quality assurance
- Routine monitoring of the instrument

Connected to other parameters

- Calculating matrix effect calibration graph method

$$\%ME = \frac{Slope_{matrix-matched}}{Slope_{solvent}} \cdot 100\%$$

- Calculating LoD and LoQ

$$LoD = 3.3 \times \frac{S_{y.x}}{Slope}$$

$$LoQ = 10 \times \frac{S_{y.x}}{Slope}$$

3.1. Linearity

Linearity is the method's ability to obtain test results, which are directly proportional to the concentration of the analyte in the sample. (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 the LC–MS signal and concentration of the analyte.

The term linearity of signal, in the context of LC–MS, has 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 also used in this course. The reason being that, if the analyte signal in samples is linear, then it is almost certain that it is also linear in calibration solutions, while the opposite might not be true. The most common cause for this is the phenomenon called matrix effect, which is addressed in the chapter 5 of this course. Linearity of the calibration graph is closely related to choosing the 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 the decrease or loss of linearity. Therefore, it is very 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)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 the clusters can also cause losses of ions.
- Thirdly, the linearity of the 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 the mass analyzer design is the deciding factor here. Mass analyzers are characterized by **the transmission efficiency** (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. , which is the ratio of number of ions that are finally detected and number of ions that entered the mass analyzer. In order to demonstrate the linear behaviour, the transmission must be independent from the 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 the LC–MS signal and concentration of the analyte.

(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 the 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 influence 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 calibration solutions in solvent are used, a comparison of calibration graphs in matrix and in solvent should be carried out. (1)(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 a 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 by using an F-test.

Here is an example where the slopes of the two graphs do not differ statistically and the use of a standard calibration graph is justified: 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. 92 (2009) 612–621.

Here is an example where the matrix interferes with the analysis and a matrix matched calibration is used: 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 3 (2011) 606–613.

2) Concentrations of calibration samples

The highest and lowest concentrations of the calibration samples should be appropriate for the method, keeping in mind the predicted variation of the analyte concentration 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 the validation, some concentrations might fall out of the linear range. Therefore, using 10 concentration levels encompassing the expected linear range is recommended. Moreover, concentrations should be approximately evenly spaced over the chosen concentration range, to ensure that the different parts of a 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. Consequently, the analysis order of calibration samples should be random.
It is useful to analyze calibration solutions in a manner as similar as possible to the unknown samples, i.e. calibration samples should be in random order and placed between unknown samples in the analytical run. Calibration samples should be analyzed at least twice (and average values should be 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 the concentrations are on the x-axis and the signals are on the y-axis. (2)(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.”

From this plotted graph we can have the first evaluation of the linearity by using the visual evaluation approach. In addition, calculation of the residuals can also be useful, due to the fact that the residuals show a more clear picture if the applied linear model actually fits the data.

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 the experimental (y_i) and calculated ($\hat{y}_i$) signal values: $y_i - \hat{y}_i$. In addition, relative residuals can be used (Eq 1).

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

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

Statistical approaches for evaluation of linearity *Note 1

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

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

**Note 1: Please see comments in the Mandel's test section below on the simplification in the video. The Lack-of-fit equation in this video should be understood the same way as (Eq 2) in the Lack-of-fit section below.*

Different expressions of signal values are used for statistical approaches:

y_{ij} is the experimental signal value at the concentration level i for replicate measurement j,

$\bar{y}_i$ is the average value of the experimental signals from p replicate measurements at the concentration level i,

$\hat{y}_i$ is the signal value at the 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 each 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(α;DoF1;DoF2),
where:
α is the probability of the F distribution (on 95% confidence interval α=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)).

As a generalisation, the number of degrees of freedom is equal to the number of data points minus the number of parameters calculated from the data. Example: In simple linear regression of the type $y = b_0 + b_1 \cdot x$ DoF is typically $n - 2$, where n is the number of data points. The „2" means that two parameters are found: slope and intercept. If the intercept is forced to zero, i.e. if the regression has the form $y = b_1 \cdot x$ then DoF = $n - 1$, because only one parameter is found.

*The 0.05 is found from 95% as follows: (100% – 95%) / 100% = 0.05.

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 the random scatter of the points is estimated from the 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 the mismatch of the calibration model (mean sum of squares due to lack of fit MSS_{LOF}).

$$F_{calculated} = \frac{MSS_{LoF}}{MSS_{error}} = \frac{\sum_{i=1}^{n} p(\bar{y}_i - \hat{y}_i)^2/(n-2)}{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \bar{y}_i)^2/(n(p-1))}$$ **(Eq 2)**

If the $F_{calculated}$ is higher than the $F_{tabulated}$, the model cannot be considered fit for the data, because the unexplained variance in the model is too big.

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 ($MSS_{residuals}$).

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

If the $F_{calculated}$ is higher than the $F_{tabulated}$, 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 (S_{y1}) with the fit of a nonlinear model (S_{y2}). In this case a three-parameter model, parabola, is used. Three-parameter model is suitable for almost all cases. For both of the models, the residual standard deviation is found using this equation:

$$S_{y_1} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n \cdot p - 2}} \qquad S_{y_2} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n \cdot p - 3}}$$ **(Eq 4)**

, where n is the number of calibration points (assuming one point per calibration level, i.e. $p = 1$). In case $p > 1$, it is assumed that all concentration levels have the same number of replicates. In many materials $n \cdot p$ is denoted by N.

The $F_{calculated}$ is found:

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

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

Please note that in different materials the Mandel's fitting test is presented at different level of rigor. In the video, a simplified version is presented, assuming that p is always 1 and disregarding the difference of degrees of freedom between linear and nonlinear model, using n - 2 in both cases. A more rigorous version, presented here. In practical application, the differences between the two are not big.

Intercept

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

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

$$A = c \cdot l \cdot e$$ **(Eq 6)**

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

The statistical way to evaluate the importance of an intercept would be via a *t*-test. In order to carry out a *t*-test, the linear regression is run with an 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

$$\text{Intercept} < 2 \cdot \text{Stdev\_intercept}$$

then it can be assumed with 95% confidence that the intercept is insignificant and can be disregarded in the calibration model. The following form of the 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 the intercept to zero.

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

- a linear model is fitted to data that are in fact nonlinear (e.g. saturation of signal at higher concentrations);
- blank samples produce signal because of a 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=32138

https://youtu.be/l07_KulyYoc

***

(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 a 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 by using an *F*-test.

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

An example where the matrix interferes with the analysis and a 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 25 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 a response is linear, it can be specified as a dynamic linear range. (1)(1) http://goldbook.iupac.org/D01874.html

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

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

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

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

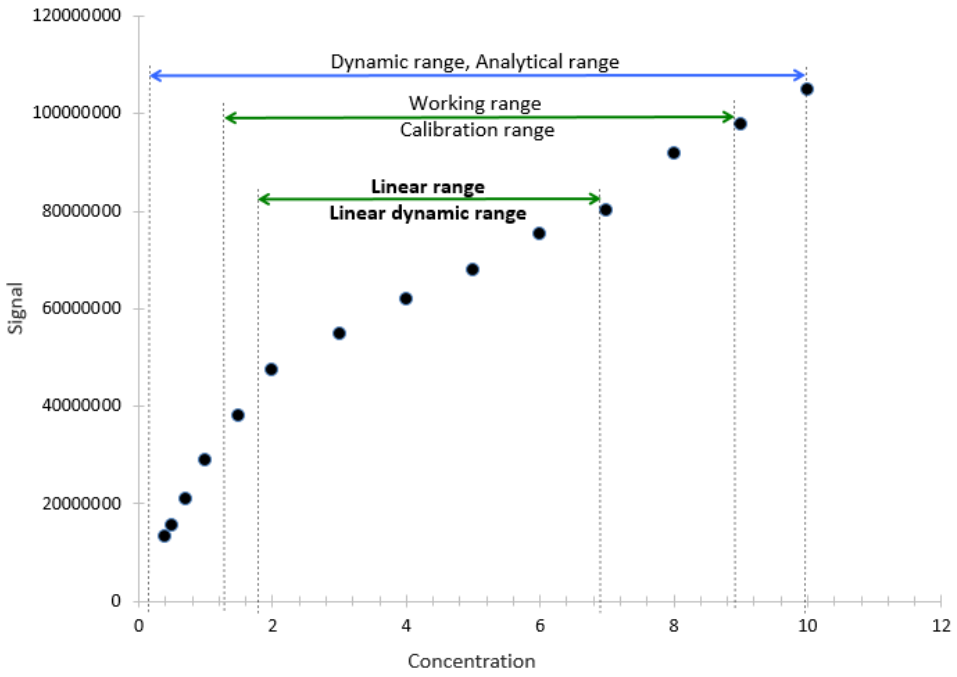


Figure 1. Different ranges

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

While a linear range for the LC–MS instruments is usually fairly narrow (and depends on the compound), several possibilities have been used in order to widen a linear range. Using the isotopically labeled internal standard (ILIS) is one of them. While the signal-concentration dependence of the compound and an 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 a linear range of the signals.

One way of increasing a linear range is by working at the lower concentrations (and diluting samples), if the analyte signal is intense enough. Another way to increase a linear range, specifically in the case of a LC-ESI-MS, is to decrease the charge competition by lowering the flow rate in the ESI source, e.g. by using a 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 the 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)(1). Sensitivity is often interpreted as related to the detection/determination ability. For example, in the recent FDA's Bioanalytical Method Validation guidance document, a 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. The numerical value (and unit) of a sensitivity is arbitrary and depends on the instrument used and its settings. Sensitivity is not an essential parameter during a method validation. However, it is very important in the 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 the 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 the ionization suppression – in fact the essence of an ionization suppression is a decrease of sensitivity due to the co-eluting compounds (see chapter 5).

The main use of a sensitivity as a parameter is threefold:

1. an optimization (for maximizing sensitivity) of the method parameters during a method development,
2. a 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, a 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 the 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 a 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 a 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 an accuracy of the results by assessing both the **systematic** and **random** effects (errors) on single results. These errors are caused by a range of reasons, such as the 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 the 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 an accuracy is a **measurement uncertainty**. (Figure 1)

Introduction to precision, trueness and accuracy

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

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

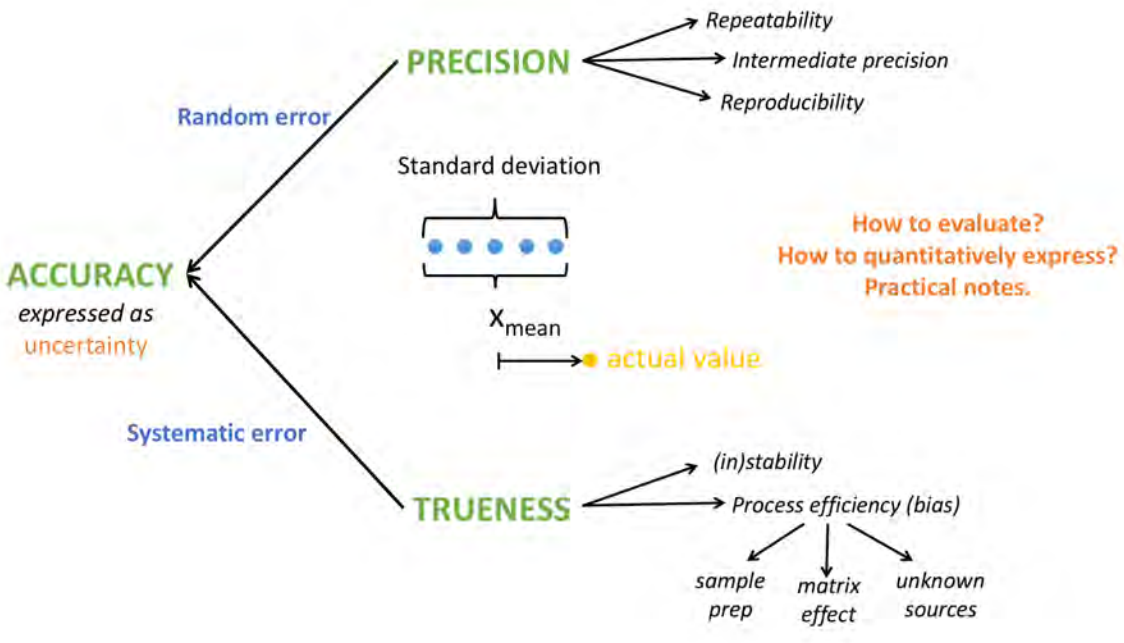


Figure 1. Summarized graphical presentation of the sections 4 to 7.

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 the 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 the measurement results obtained at *different laboratories*. Sometimes a mistake is made and a term reproducibility is used for a 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

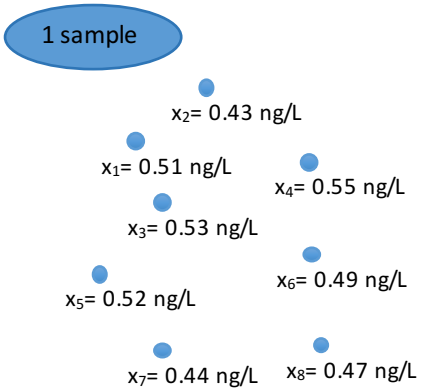
https://www.youtube.com/watch?v=ur9gRxpFpPw

4.1_precision.pdf 83 KB

Precision

Repeatability, intermediate precision, reproducibility

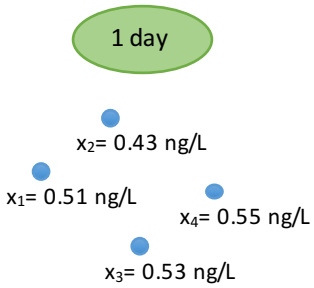
Glyphosate in ground water:



Precision characterizes the closeness of agreement between the measured values.

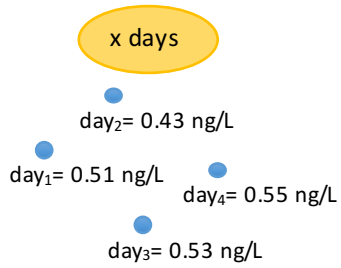
Precision relates to the random error of a measurement system.

Glyphosate in ground water:



Repeatability

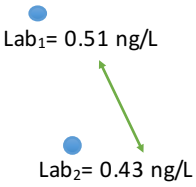
one operator, one experimental setup, one set of reagents



Intermediate precision

different analysts, calibrants, batches of reagents, columns, spray needles

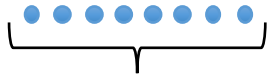
Reproducibility



Precision between measurement results obtained at different laboratories

*Not always needed for single-lab validation **but** important if the method is going to be used in several laboratories*

Glyphosate in ground water:



Standard deviation

Expressed as absolute standard deviation:

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

also as:

Relative standard deviation (RSD)

Variance (s^2)

Coefficient of variance (CV)

$$RSD = \frac{s}{x_{mean}}$$

As a general rule…

$$s_{repeatability} \leq s_{intermediate\ precision}$$

If not:

- Too few replicates
- Too short time interval

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 the 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 varies in different validation guidelines, but is typically between 6 and 15 for each of the material used in the study.

Precision is usually expressed as "imprecision": 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 the 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) in order 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 the time constraints (e.g. stability) or because the available amount of sample is insufficient for replicates. 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 the 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 the 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 a 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

The specification of the 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 the 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 the CV should be smaller than 20%.

It is important to keep in mind that ideally the determination of precision should be carried out for all matrices that are within the scope of the application of the method. This general principle becomes even more important for the LC–MS methods, where the 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 a strong evidence that repeatability and intermediate precision in LC-MS is concentration-dependent.

4.4. Examples

Example for calculating repeatability
https://www.uttv.ee/naita?id=32142

https://youtu.be/4eKvvgNOH3c

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 focuses on the different aspects of trueness and its quantitative estimate, bias. In section 5.1, the general information and definitions are given. This section, in particular, explains the bias constituents (i.e. causes of bias). Section 5.2 addresses the evaluation of (overall) bias in practice. Sections 5.3 and 5.4 look at more specific aspects related to trueness (bias), specifically in case of 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 an absolute bias (Eq. 1 below), i.e. simply the difference, or as a relative bias (Eq 2 below), i.e. as a difference divided by the reference value. The main causes of bias in LC-MS results can be termed as bias constituents and are the following:

1. Bias caused by the analyte loss during sample preparation, expressed quantitatively by recovery (R);
2. Bias due to the limited stability of the analyte (see in Stability section) in the 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)(1) Xlab: average of results obtained by laboratory; Xref: reference value	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)(5) Bias can also be determined as the short-term i.e. within-day bias, but the long-term bias is more useful, e.g. for measurement uncertainty estimation, as is explained in section 7. occurring at any stage of the analytical process. Bias can be expressed as an absolute or relative bias.
	Relative bias (2)(2) Can be expressed as a simple ratio or as a percentage (in the latter case the ratio is multiplied by 100). $bias(\%) = \frac{X_{lab} - X_{ref}}{X_{ref}} \cdot 100\%$ **(Eq 2) (3)(3) This bias is expressed as a percentage. Absence of bias corresponds to 0%. Negative bias values indicate negatiive and positiive bias values positiive bias.** $bias = \frac{X_{lab}}{X_{ref}} = ME_{ionization} \cdot R \cdot B_{stab} \cdot B_{other}$ **(Eq 3) (4)(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 Xlab/Xref 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.**	Absolute bias is useful when it is either constant over the used concentration range or if it is evaluated separately at different concentrations. 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)(2) Can be expressed as a simple ratio or as a	$PE = \frac{X_{lab}}{X_{ref}}$ **(Eq 4) (6)(6) This equation holds, if Bstab and Bother are insignificantly different from 1.**	Process efficiency (PE) refers to the joint effect of possible losses during sample preparation and ionization

percentage (in the latter case the ratio is multiplied by 100). , *PE*	$PE = \frac{m_{\text{analyte detected}}}{m_{\text{analyte in sample}}} = ME_{\text{ionization}} \cdot R$ **(Eq 5)**	suppression/enhancement in the ion source. *PE* is a useful parameter for characterizing the analysis method when it is either required for the characterization of the method or when it is intended to carry out correction with *PE* (more generally: bias correction).
Recovery (2)(2) Can be expressed as a simple ratio or as a percentage (in the latter case the ratio is multiplied by 100). , *R*	$R = \frac{m_{\text{analyte extracted}}}{m_{\text{analyte in sample}}}$ **(Eq 6)**	Recovery (7)(7) In the case of most other analytical techniques, recovery would also include the possible matrix effects, so that it would be effectively equal to PE as defined above. expresses the efficiency of the sample preparation step: the proportion of an analyte obtained from the sample during sample preparation (see also (4)(4) This way of expressing a 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 Xlab/Xref is interpreted as a recovery, i.e. the four bias components in this equation would combine into recovery. However, in LC-MS it is useful to make a distinction between recovery – relating specifically to sample preparation – and other bias components.).
Ionization suppression/enhancement (matrix effect), $ME_{\text{ionization}}$ (2)(2) Can be expressed as a simple ratio or as a percentage (in the latter case the ratio is multiplied by 100). , (8)(8) There are different ways for expressing the matrix effect. We use the way which is similar to expressing a recovery and a process efficiency.	$ME_{ionization} = \frac{m_{\text{analyte detected}}}{m_{\text{analyte extracted}}}$ **(Eq 7)**	In LC-MS the term matrix effect refers to the suppression (usually) or enhancement (rarely) (9)(9) In the case of alleged ionization enhancement, it may under closer examination turn out that the reason for the enhanced signal is not the analyte signal enhancement but the 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]. 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 the analyte decomposition. Depending at which stage of the sample preparation the 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 the laboratory; X_{ref}: reference value.

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

(3) This bias is expressed as a 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. The absence of bias in this case corresponds to 1. Bias values below 1 indicate a negative bias and values above 1 indicate a 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 a distinction between recovery – relating specifically to sample preparation – and other bias components.

(5) Bias can also be determined as the short-term i.e. within-day bias (see below), but the 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 also include 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 the enhanced signal is not the analyte signal enhancement but the 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 the 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 a 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, a process efficiency (*PE*) is often used as a LC-MS specific term for the overall trueness (if B_{other} and B_{stab} are insignificant).

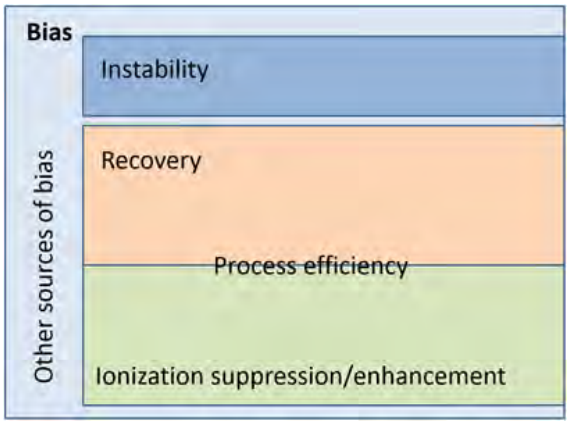


Figure 1. Relations between the bias constituents.

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, the within-day bias is larger than the long-term bias – the longer time we examine, the more effects will switch their status from systematic to random. Consequently, a within-day precision (repeatability, s_r) is smaller than a 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 an intermediate precision and a long-term bias estimate is more useful and reliable than a repeatability and a short-term bias estimate.

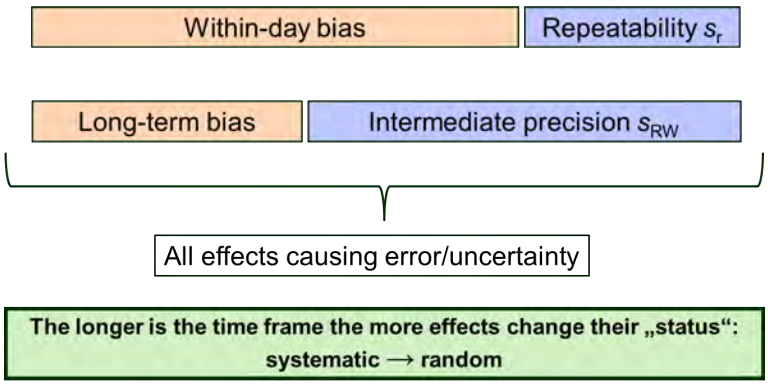
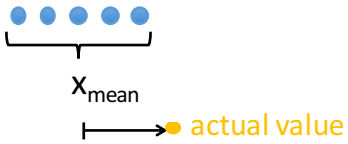


Figure 2. Dependence of bias on the examined timeframe.

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

Recovery, ionization suppression and process efficiency



Trueness relates to the systematic error of a measurement system.

$$\text{bias} = X_{lab} - X_{ref}$$

Agreement between the mean value of replicate measurements and the true value of the measured quantity

Important! Bias takes into account the effects that are systematic over a long term, occurring at any stage of the analytical process

Bias in LC-MS

$$\text{bias} =$$

Bias in LC-MS

$$\text{bias} = B_{stab} \times$$

(in)stability refers to analyte stability or instability during analysis

Bias in LC-MS

recovery

$$\text{bias} = B_{stab} \times R \times$$

(in)stability

refers to analyte loss during sample preparation step

$$R = \frac{m_{\text{analyte extracted}}}{m_{\text{analyte in sample}}}$$

Bias in LC-MS

recovery

$$\text{bias} = B_{stab} \times R \times ME_{ionization} \times$$

(in)stability

matrix effect

refers to the suppression or enhancement of analyte ionization by co-eluting compounds originating from the sample matrix.

$$ME = \frac{m_{\text{analyte detected}}}{m_{\text{analyte extracted}}}$$

Bias in LC-MS

recovery

$$\text{bias} = B_{stab} \times R \times ME_{ionization} \times B_{other}$$

(in)stability — matrix effect — eg. purity of the standard substance, calibration bias of volumetric ware.

Bias in LC-MS

recovery

$$\text{bias} = B_{stab} \times R \times ME_{ionization} \times B_{other}$$

(in)stability — matrix effect

Process efficency, PE

$$PE = \frac{m_{\text{analyte detected}}}{m_{\text{analyte in sample}}}$$

Bias in LC-MS

recovery

$$\text{bias} = B_{stab} \times R \times ME_{ionization} \times B_{other} - 1$$

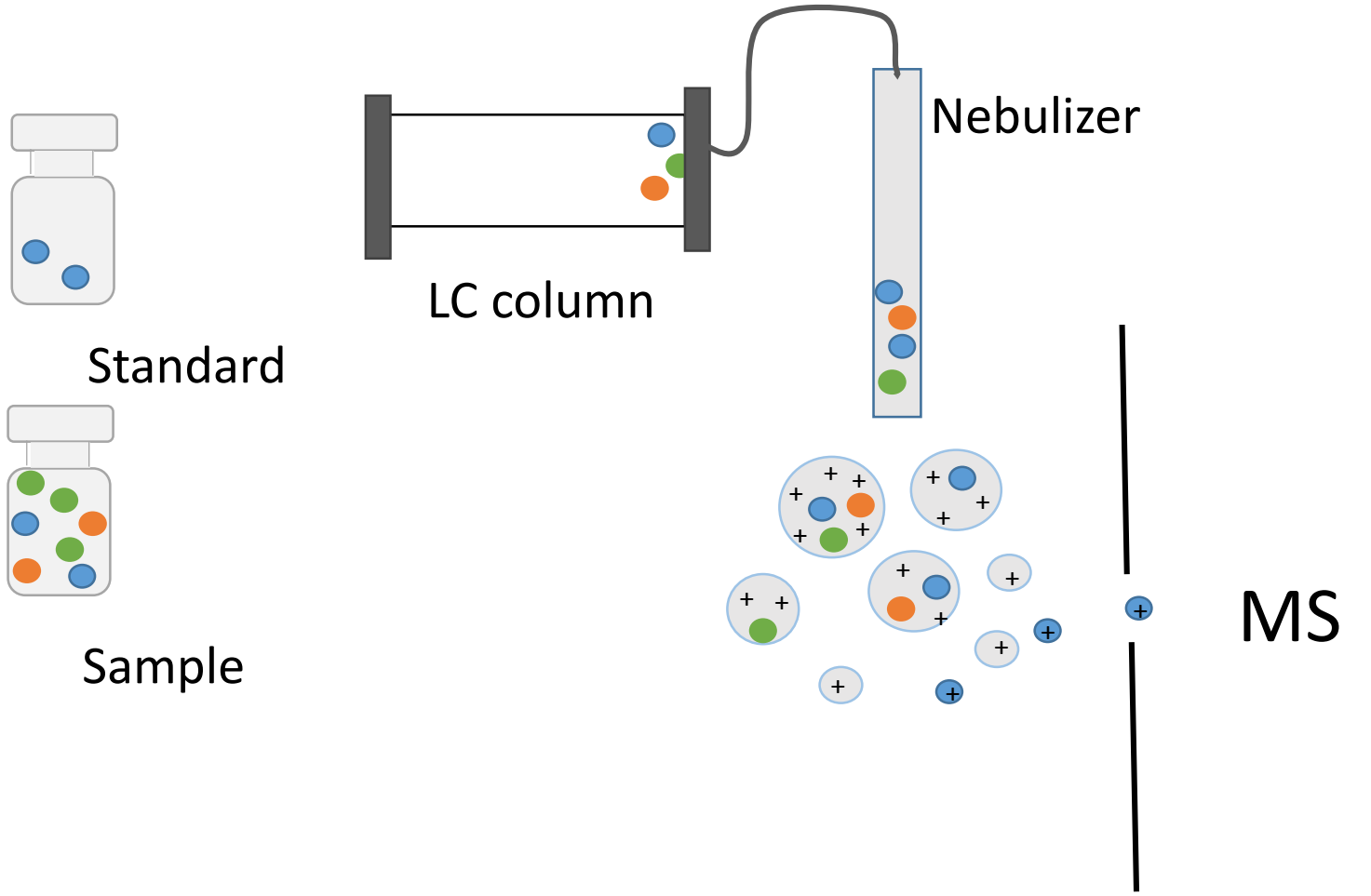
(in)stability

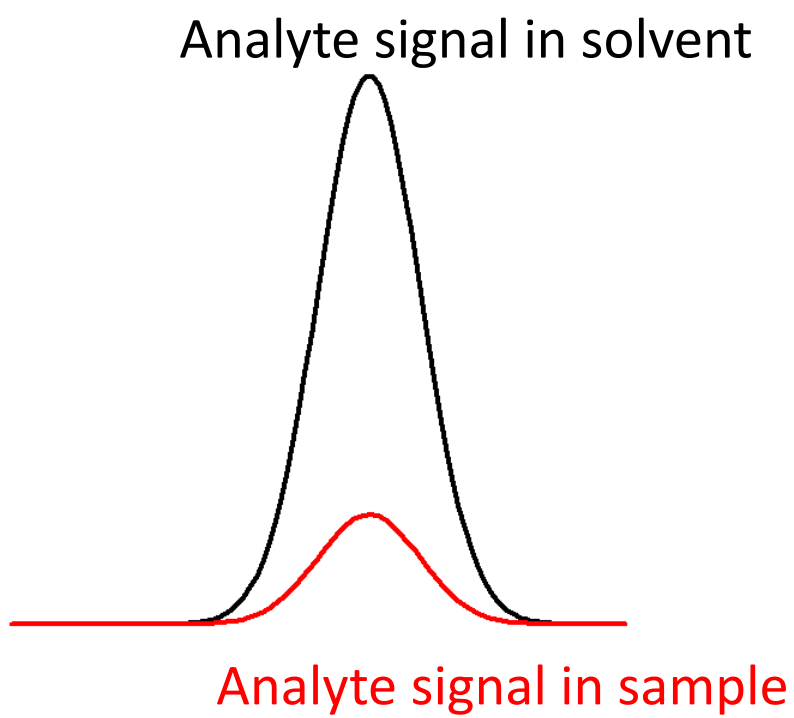
matrix effect

Process efficency, PE

$$PE = \frac{m_{\text{analyte detected}}}{m_{\text{analyte in sample}}}$$

Matrix effect





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 values are essential for determining bias. There are essentially 4 approaches for 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 testings or interlaboratory comparisons.

(c) Reference values obtained from spiking studies. Spiking means adding a known amount of the 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 the matrix are accurately measured and if the analyte is carefully dispersed into the matrix, then the *increase* of the analyte content in the matrix will be accurately known and can be used as a reference value.

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

Reference values provided by these four approaches have different qualities. Reference values from CRMs are generally considered the most reliable, but an important prerequisite is a matrix match between the CRM and the samples actually analyzed. Reference values of the reference materials without a certificate are usually less reliable and a matrix match is required as well as in the case of CRMs. Spiking can give reliable reference values if done carefully and correctly. 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, but very difficult with soil or minerals). If spiking is feasible, then the ideal matrix match can be achieved, because then the routine samples can be spiked. Lastly, 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 a 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 a 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 the matrix effect is based on detecting the presence of ionization suppression (1)(1) The following text is relevant for both ionization suppression and enhancement. However, since ionization suppression is significantly more common, we use only the term "suppression" in the following text. by recording the matrix effect profile with post-column infusion [ref 25]. For this the blank sample extract – which does not contain 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. The MS monitors the analyte signal.

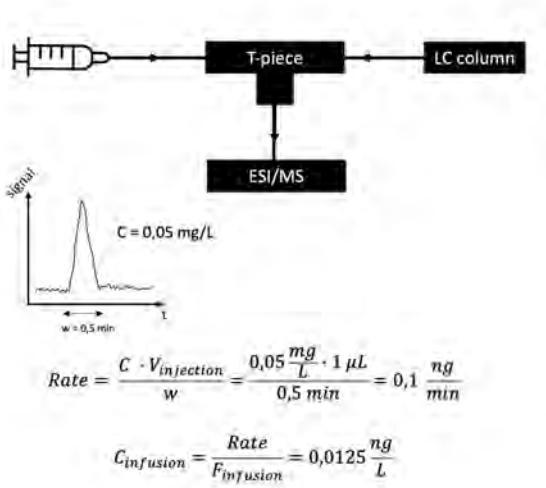


Figure 1. The experimental setup for recording the matrix effect profile.

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 the sample and solvent injection within the precision of the MS signal. For the retention time region where the ionization suppression occurs, the MS signal in the sample injection decreases. In order to have a method not affected by an ionization suppression, the analyte peak should elute away from the suppression region. This approach does not enable quantitative evaluation of the ionization suppression/enhancement but only enables confirmation of its presence/absence and this is very useful for tuning the chromatographic separation of the analyte and possible compounds causing ionization suppression.

The most recent FDA validation guideline proposes an evaluation of ionization suppression by assessing the parallelism of dilution plots [ref 8].

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

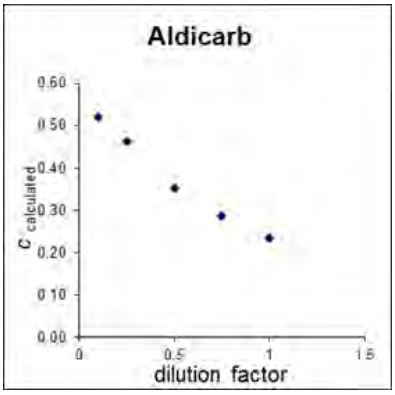


Figure 2. Example dilution plot of Aldicarb determination in garlic.

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

decreases 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 the ionization suppression is possible with post-extraction addition methods as is explained in the following videos. The first video explains the principles of evaluating a 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=J1O1bD7gEy4

The second video explains evaluating a 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 a known concentration is prepared in the 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 ia thereafter analyzed giving the peak area (signal) S_{sample}. The ionization suppression/enhancement effect can be calculated:

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

$ME_{ionization}$ value of 100% indicates no effect, less than 100% indicates an ionization suppression and $ME_{ionization}$ over 100% indicates an ionization enhancement due to the coeluting sample compounds. From this definition, though most often used in the LC-MS literature, some possible misunderstandings can 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% indicate an ionization enhancement and below 0% a suppression. The corresponding equation is:

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

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

A similar approach is described in the most recent FDA validation guideline. In this approach, two calibration graphs are constructed, one in the solvent and the other one in the post-extraction spiked samples (i.e. sample extracts obtained from sample preparation). 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 an approach based on the slope, linearity of the method 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 the statistical significance of the obtained matrix effect values.

(c) All of the calculations described above can be done either in the signal scale or concentration scale. The obtained results are fairly similar if the samples used for the 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 the concentration-based calculations.

If signal- or concentration-based calculations are used (not slope-based), the number of samples and replicates used for the 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 effects depend 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 the method optimization/validation and then 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 the ionization suppression to the level of being insignificant, it should be taken into account when calculating the results. Several approaches have been suggested and tested for reducing the ionization suppression effect, mainly focusing on the ESI ionization source. In broad terms the approaches can be categorized as based on (a) the sample preparation, (b) the instrumental modifications and (c) the 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 of the 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 the sample preparation.

Different sample preparation techniques have been compared and for example found that for phenacetin and caffeine determination in endogenous plasma, protein precipitation is the least favourable technique for LC-ESI-MS analyses while LLE was the most favourable [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 the matrix compounds similar to the analyte (i.e. potentially co-eluting from HPLC with the analyte) [ref 30]. The reason probably being 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 a 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 the matrix effects [ref 31, ref 32, ref 33, ref 34]. Also, 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 the use of APCI nor switching to 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 the 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 the 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

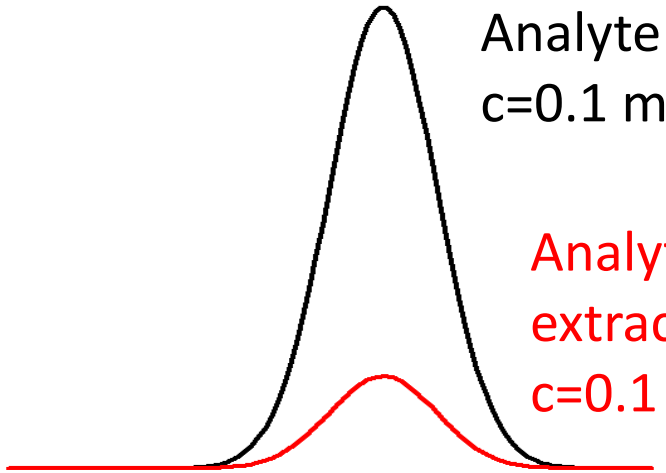
https://www.youtube.com/watch?v=5lonQCdmcis

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

Evaluating matrix effect

Signal-based method



Analyte signal in solvent
c=0.1 mg/kg

Analyte signal in post-extraction spiked matrix
c=0.1 mg/kg

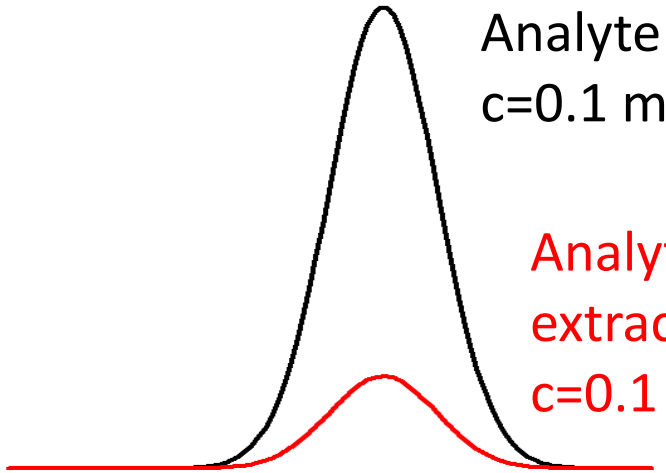
$$\%ME = \frac{\text{Analyte signal}_{\text{post-extraction spiked matrix}}}{\text{Analyte signal}_{\text{solvent}}} \cdot 100\%$$

100% no matrix effect

<100% ionization suppression

>100% ionization enhancement

Signal-based method



Analyte signal in solvent
c=0.1 mg/kg

Analyte signal in post-extraction spiked matrix
c=0.1 mg/kg

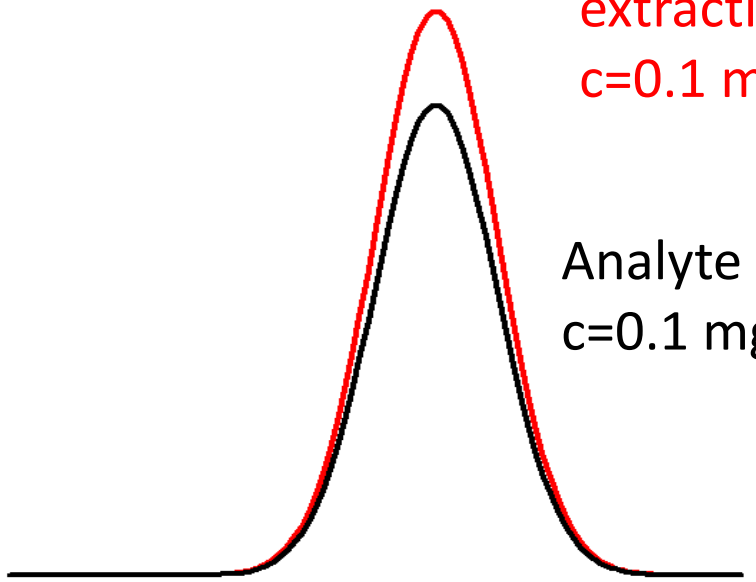
$$\%ME = \frac{\text{Analyte signal}_{\text{post-extraction spiked matrix}}}{\text{Analyte signal}_{\text{solvent}}} \cdot 100\%$$

100% no matrix effect

<100% ionization suppression

>100% ionization enhancement

Signal-based method



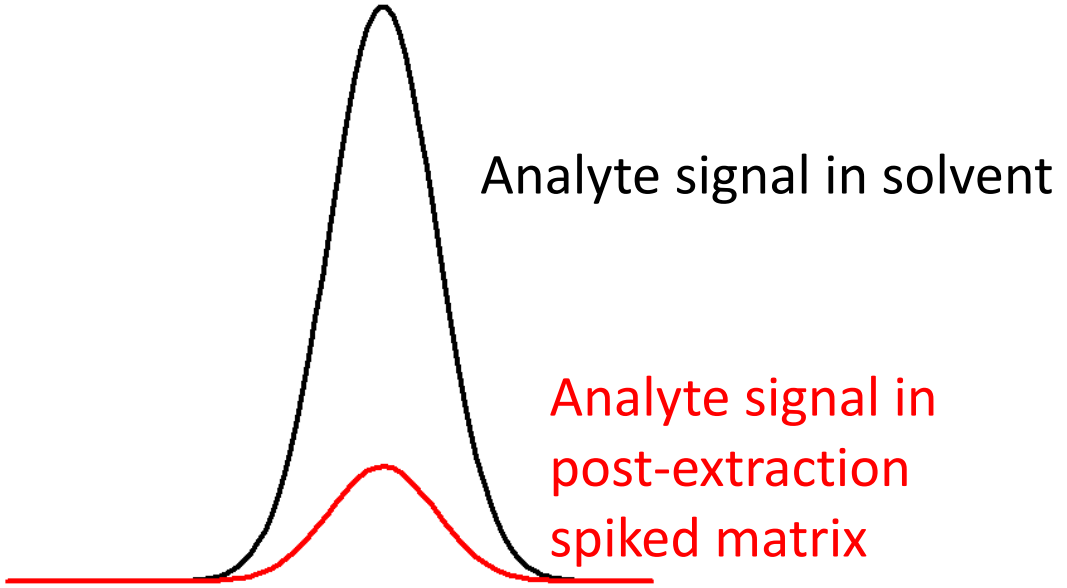
$$\%ME = \frac{\text{Analyte signal}_{\text{post-extraction spiked matrix}}}{\text{Analyte signal}_{\text{solvent}}} \cdot 100\%$$

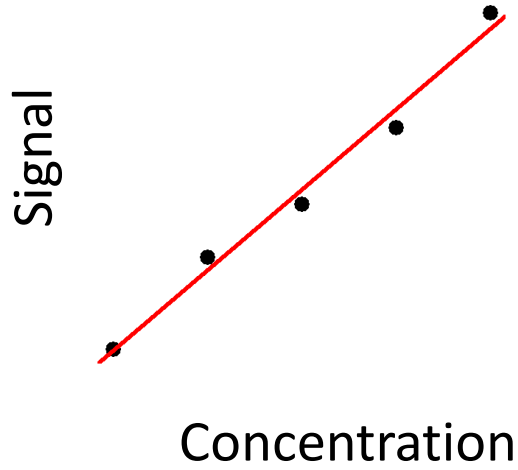
100% no matrix effect

<100% ionization suppression

>100% ionization enhancement

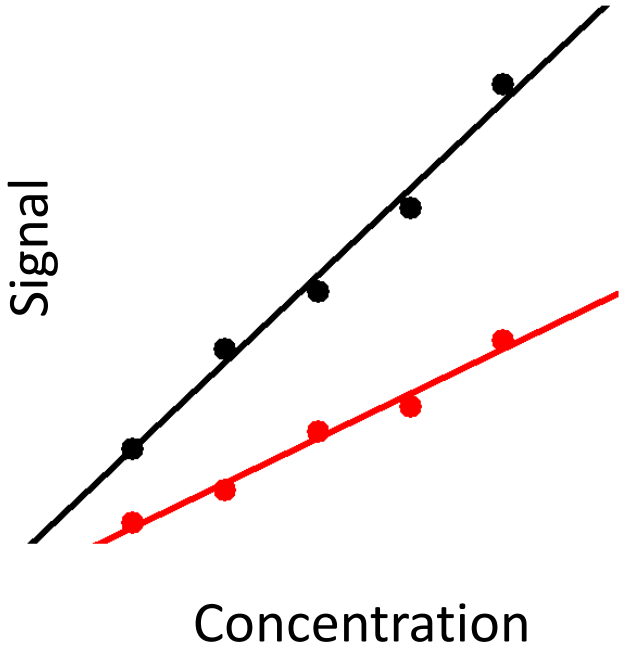
Concentration-based method





$$\%ME = \frac{\text{Analyte concentration found}_{\text{post-extraction spiked matrix}}}{\text{Spiked concentration}_{\text{solvent}}} \cdot 100\%$$

Calibration graph method



Calibration graph in solvent

Matrix-matched calibration graph

$$\%ME = \frac{\text{Slope}_{\text{matrix-matched}}}{\text{Slope}_{\text{solvent}}} \cdot 100\%$$

%ME does not depend on analyte concentration
Suitable for samples already containing analyte

Combining recovery and matrix effect

$$\%PE = \frac{\%R \cdot \%ME}{100\%}$$

Process efficiency:
Common term for expressing trueness in LC-MS

Recovery:
Loss of analyte in sample pre-treatment

Matrix effect:
Matrix influence on analytes ionization

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 the 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 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

6_practical_aspects.pdf 82 KB

Precision, trueness, accuracy

guidelines

Validation guidelines

- ICH
- AOAC
- Eurachem
- IUPAC
- EMA
- FDA
- SANCO
- Nordtest Guide for measurement uncertainty

Validation guidelines

- ICH
- AOAC – AOAC International
- Eurachem
- IUPAC – Union of Pure and Applied Chemists
- EMA
- FDA
- SANCO
- Nordtest Guide for measurement uncertainty

Single laboratory method validation guidelines

Validation guidelines

- ICH
- AOAC
- Eurachem
- IUPAC
- EMA – European Medicines Agency
- FDA – The United States Food and Drug Administration
- SANCO
- Nordtest Guide for measurement uncertainty

Bioanalytical methods

Validation guidelines

- ICH - International Conference on Harmonization
- AOAC
- Eurachem
- IUPAC
- EMA
- FDA
- SANCO
- Nordtest Guide for measurement uncertainty

Registration of pharmaceuticals

Validation guidelines

- ICH
- AOAC
- Eurachem
- IUPAC
- EMA
- FDA
- SANCO
- Nordtest Guide for measurement uncertainty

Validation guideline for pesticide residue analysis in food and feed

Validation guidelines

- ICH
- AOAC
- Eurachem – "The Fitness for Purpose of Analytical Methods"
- IUPAC
- EMA
- FDA
- SANCO
- Nordtest Guide for measurement uncertainty

Validation guidelines

- ICH
- AOAC
- Eurachem
- IUPAC
- EMA
- FDA
- SANCO
- Nordtest Guide for measurement uncertainty

Terminology - precision

- **Repeatability** = within-run precision (FDA, EMA)

- **Intermediate precision** = between-run precision (FDA, EMA)
  = intra-lab reproducibility (Eurachem)
  = within-laboratory reproducibility (Nordtest)

Terminology

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

Intermediate precision

- Time interval only specified in:
  - EMA – two different days
  - Nordtest* – at least several months

Recommendation

Ideal – several months

In practice – estimate with one month

How many concentration levels?

- 2 – IUPAC
- 3 – FDA, NordVal
- 4 – EMA, AOAC
- Not specified in Eurachem
- ICH
  - 1 level if concentration doesn't vary
  - Can choose levels as long as 9 measurements in total

Recommendation

Choose accordingly to concentration levels you are expecting!

How many replicates?

- 5 replicates – NordVal, EMA, FDA, AOAC
- 3 replicates – ICH
- 6 replicates – ICH (1 concentration level)
- 6-15 replicates – Eurachem

Recommendation

5 – 10 replicate measurements

Can reduce number of replicates if a long period intermediate precision.

And for trueness?

- Most times done with precision measurements.

BUT

- If carried out separately from precision, a reduced design may be used (e.g., 2 replicates in five days).

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

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

- The sample matrices should represent samples which are routinely analyzed in the laboratory. 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 the precision.
- A sufficient amount of sample must be available so that the required number of replicate measurements with the subsamples can be carried out.
- A sample should be homogeneous, so that subsamples would have as similar matrix and analyte concentrations as possible.
- In the case of determining the 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 a reference value for bias calculation

Three types of samples can be used for obtaining a 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 (RM) 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 like 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 the 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. Then the sample is split into two aliquots – one is analyzed in its original state and the other one 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), however, is certainly an indication of poor trueness. Strictly speaking, trueness studies of this type only assess bias due to the 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 for both at the same time.

2. How good is the matrix match between a CRM and a routine sample and at which point is it 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 the existence of the matrix effect as a 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 a 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 contains more 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 the matrix effects may vary from one citrus fruit to another. It may be a good idea to determine the 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 plasma from an ill patient	Blood plasma from a healthy volunteer	Although blood plasma 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].	Again, it is logical that if a blood plasma RM or CRM is available then it is used for a 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)(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 the bias constituents. is frequently done or at least attempted. In very broad terms, a distinction can be made between avoiding/minimizing bias and correcting for bias is applicable.

Minimizing (in favourable 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 the analysis, an accurately measured amount of the ILIS is added to the sample and instead of just using the analyte signal $S_{Analyte}$ for the quantitation, the ratio of the analyte and the ILIS signals, $S_{Analyte}/S_{ILIS}$, is used. If some of the heavy atoms (i.e. not hydrogen atoms) in the ILIS are labeled (which is preferred as opposed to replacing hydrogen by deuterium) then the extraction, adsorption, etc properties of the analyte and the ILIS are almost indistinguishable. As a consequence, extraction losses, retention time, co-eluting interferents and consequently matrix effects for the 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, the use of ILIS is one of the most powerful approaches of assuring quality of LC-MS results. The three main drawbacks of the ILIS are (a) non-availability of ILIS-s for many analytes, (b) difficulties in dispersing the ILIS into the sample in such a way that it will be in the same molecular environment as the native analyte and (c) in most cases they are very expensive when a mixture of different analytes is considered and more than one ILIS is needed.

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. In this course, most of the discussion related to bias 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 correctthe results, therefore a 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 the 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 the possible existence of bias has to be taken into account in the uncertainty estimate of the result!

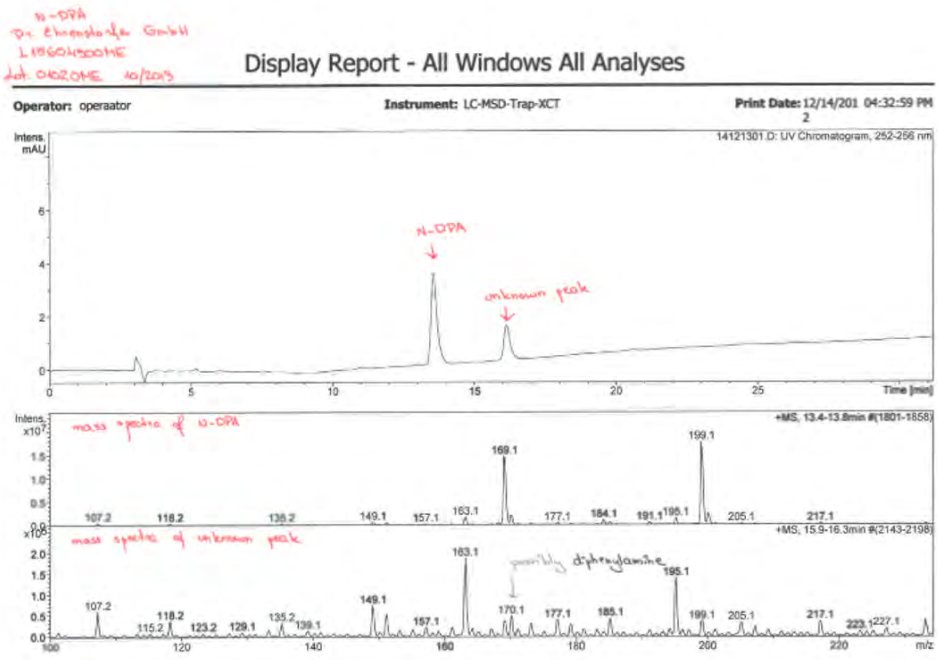
***

(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 the bias constituents.

6.4. Real life examples

Example of an impure standard

LC-UV method was used for quantification and MS was used for confirmation of the result. A calibration curve was prepared with a separate pure standard of 2-Nitrodiphenylamine (N-DPA) and this was used to determine the concentration of a certified standard solution by Dr. Ehrenstorfer. However, instead of a certified 10 ng/mL, a concentration of 7.7 ng/mL was observed for the solution. A closer look at the UV chromatogram showed two peaks for the Dr. Ehrenstorfer solution while the calibration curve solutions had just one peak. The MS spectra shows that the unknown additional peak might possibly be a diphenylamine.



Bigger picture

Figure 1. UV chromatogram of Dr. Ehrenstorfer solution and mass spectra of the N-DPA and unknown peaks.

PFOS standards by different producers

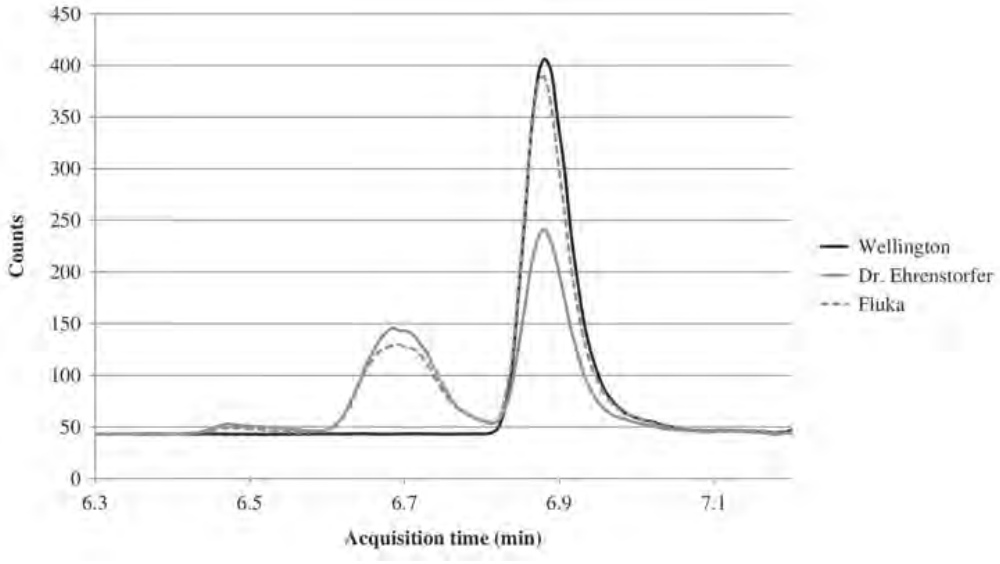


Figure 2. Chromatograms of PFOS solutions of Dr. Ehrenstorfer, Fluka and Wellington Laboratories.

PFOS is an anthropogenic fluorosurfactant and global pollutant. It is known that there can be different isomers for PFOS. During method development, various reference solutions were analysed and the chromatographic analysis revealed that both PFOS commercial standard solutions [Dr. Ehrenstorfer (Lot no 5011ME) and Fluka (Lot no SZB008XV)] contained not only linear PFOS isomer but also a branched version (Figure 2). **No indication that the standard is a mixture of isomers is mentioned by these manufacturers in the documentation.** A linear PFOS solution from Wellington Laboratories Inc. was used to quantify the amount of linear PFOS in Dr. Ehrenstorfer and Fluka PFOS solutions. [ref 59]

Problems with isotope labelled internal standards

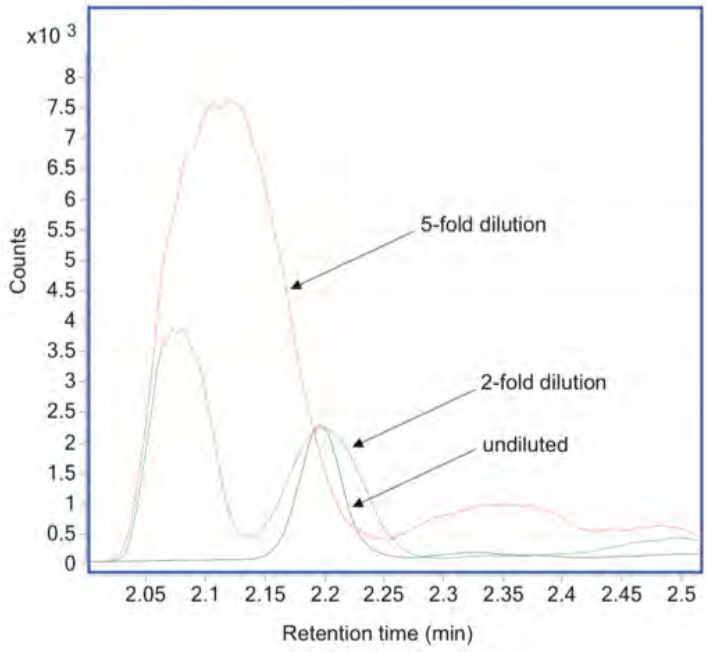


Figure 3. Normetanephrine-D3 (C = 500 pg/mL) extracted ion chromatograms (169->137) of protein precipitated (with acetonitrile) sample extracts analyzed undiluted and with 2- and 5-fold dilutions.

During method validation, an unusual interference for a D3-labeled internal standard of normetanephrine was discovered – a signal of the interfering compound increased while the matrix effects were reduced by dilution, e.g. **dilution eliminates the matrix suppression on the interfering compound**. The results stress the need to monitor interfering compounds and evaluate matrix effects at every step of the method development. **Matrix effects and interferences can be different for analytes and their corresponding isotopically labeled internal standards**. This means that the use of isotopically labeled internal standards cannot guarantee accuracy of the obtained results. [ref 60]

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 characterize 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, the accuracy of a method (i.e. accuracy of the results delivered by the method) is affected by a systematic (bias) as well as random (precision) error components and is, therefore, studied as two components: trueness and precision [ref 2].

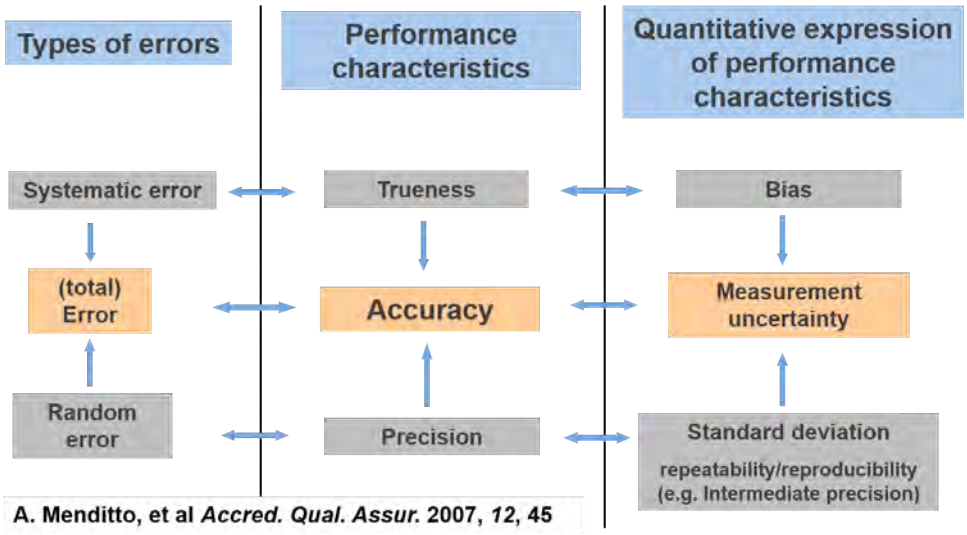


Figure 1. Interrelations between the different error types, the performance characteristics used to estimate them and the ways of expressing the estimates quantitatively. This type of scheme was originally published in ref 56. (See Note 1)

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 as a qualitative term [ref 6]. However, in practice it is useful to consider that accuracy is quantitatively expressed as a measurement uncertainty. There are different approaches for the measurement uncertainty estimation, but in practice the approach based on the 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

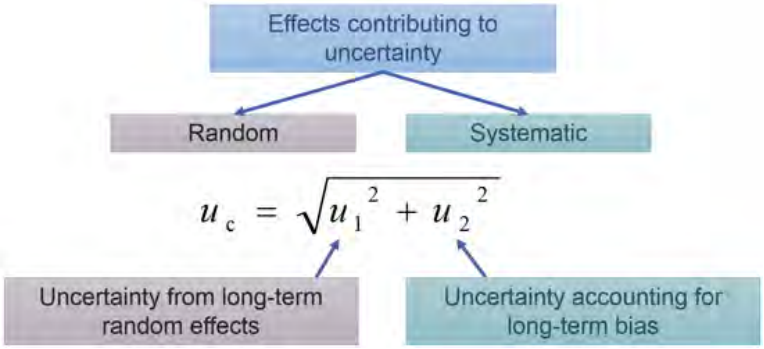


Figure 2. The main idea of estimating measurement uncertainty using validation and quality control data.

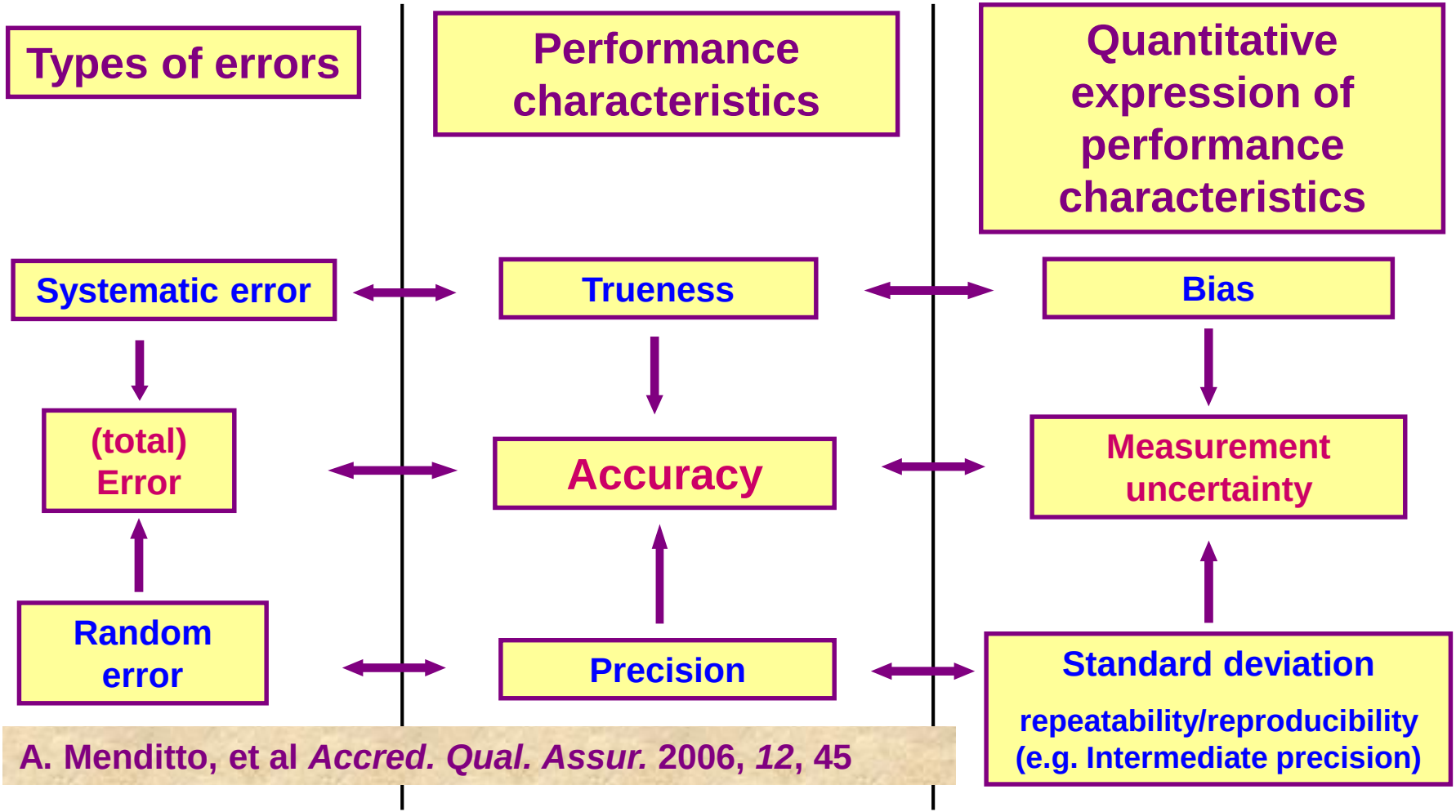
***

Note 1: Literature reference in the video has incorrect number of year. The correct reference is as stated here: A. Menditto, et al *Accred. Qual. Assur*. **2007**, 12, 45

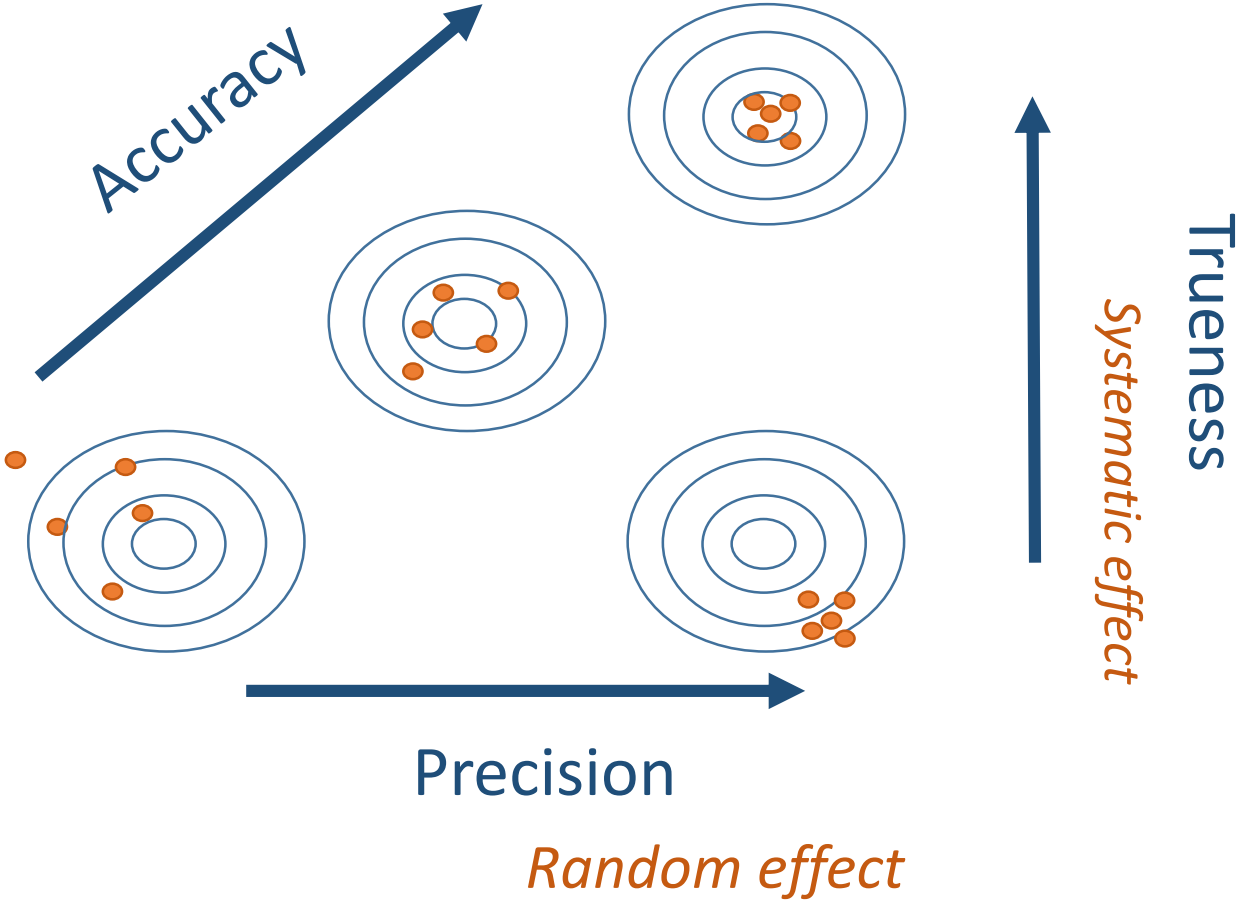
7_accuracy_and_measurement_uncertainty.pdf 614 KB

Accuracy and measurement uncertainty

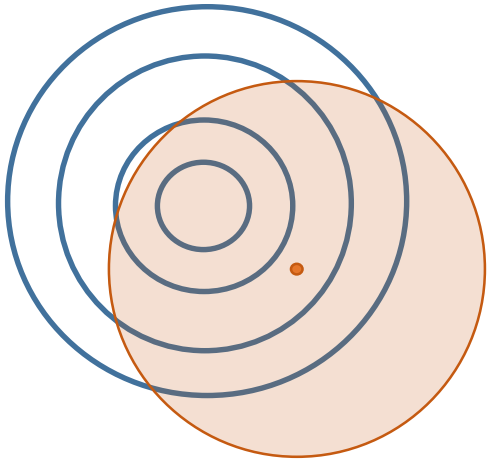
Section 7



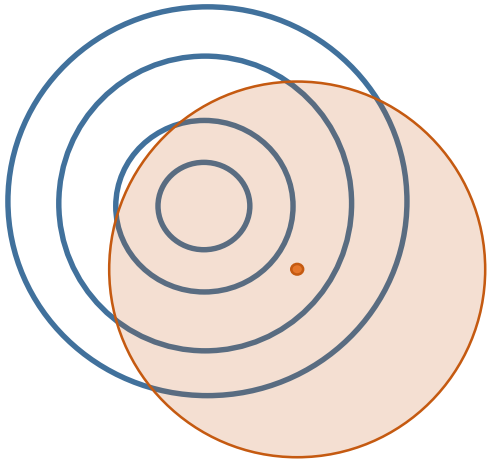
A. Menditto, et al *Accred. Qual. Assur.* 2006, *12*, 45


Accuracy
Trueness
Systematic effect
Precision
Random effect

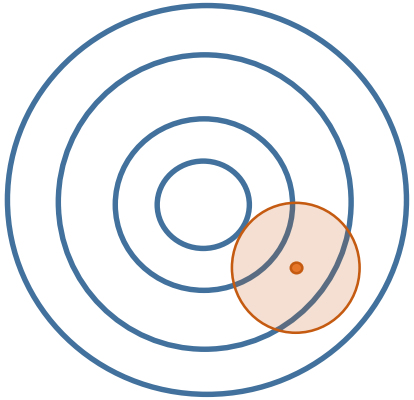
How does measurement uncertainty fit to this picture?



Which method is more accurate?

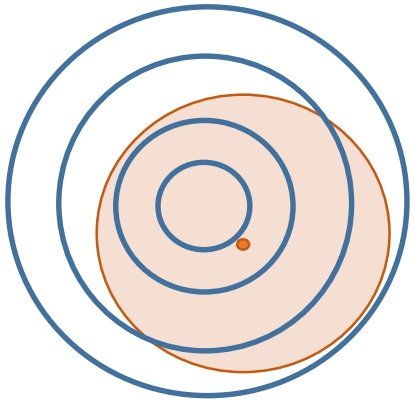


If uncertainy is acceptable to the customer then method is OK.

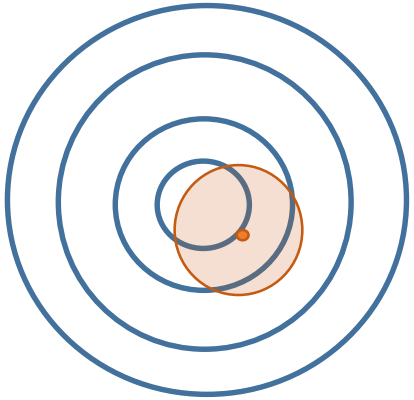


Method is not accurate enough and the uncertainty estimation should be reviewed.

Which method is more accurate?



Less accurate



More accurate

Uncertainty estimation approaches

Based on modelling

Uncertainty data of many parameters are used

Rigorous but work-intensive and needs competence

Based on validation data

Intermediate precision and long-term bias data are used

Less rigorous but easy to apply in a routine lab

Using validation data for uncertainty

Effects contributing to uncertainty

Random

Systematic

$$u_c = \sqrt{u_1^{\,2} + u_2^{\,2}}$$

Uncertainty from long-term random effects

Uncertainty accounting for long-term bias

There is a Dedicated MOOC for measurement uncertainty:

Estimation of measurement uncertainty in chemical analysis

https://sisu.ut.ee/measurement/

8. Stability

In this part of the course the issues related to the analyte's (insufficient) stability are introduced and explained. In addition, different 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 universally not included in the validation guidelines as a validation parameter (it is included only in SANTE/SANCO, EMA, FDA and AOAC). The reason for this 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 considering 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 the conditions in our method to improve analytes' stability and, finally, how can we still get acceptable results if an analyte is unstable.

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

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

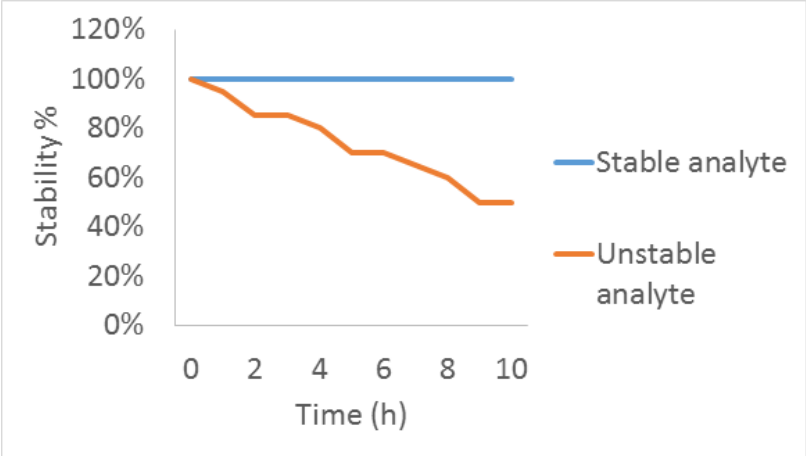


Figure 1. Analyte stability over the period of time.

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

Decomposition usually leads to a decrease in analyte content. However, in the case of analysis of decomposition products, degradation can lead to an increase in the analyte content.

If an 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 the 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, the 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

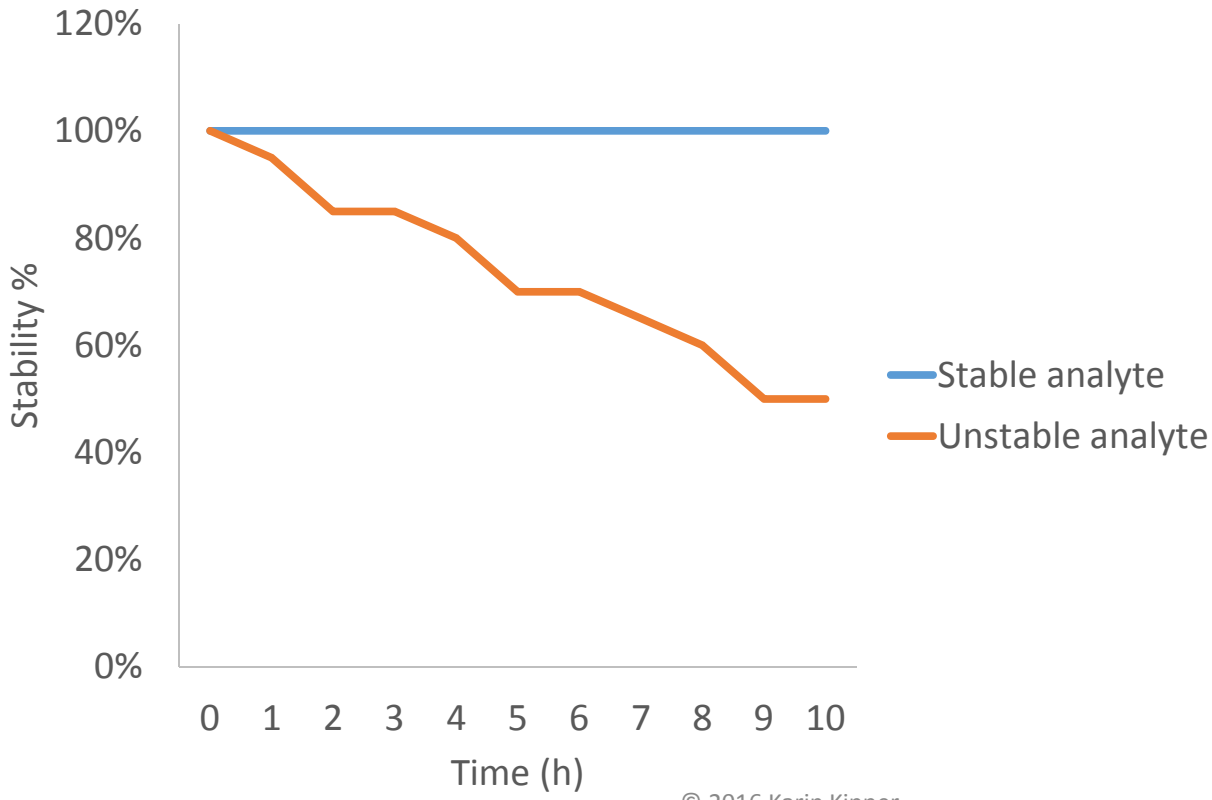
Stability introduction

Stability

Analyte stability must be ensured during **sample collection, processing, storage, extraction and duration of the analysis** to generate reliable (bio)analytical data

Stability

Decomposition of analyte during the period of time



Stability

- Instability
  - Lowering the analyte content in the sample
  - Influences trueness and precision

- The rate depends on the experimental details (matrix, temperature, light etc)

- Not only the samples, but also the standards can degrade

- Stability **can not** be proven by literature data and experiments have to be carried out in the laboratory

8.1 Different types of stability

SANTE/SANCO [ref 4] specifies briefly that an 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 an 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;
- Auto-sampler Stability.

Stability should be evaluated using both low- and high-level QC samples with at least in three replicates.

This guideline distinguishes between an analyte stability in the calibration and stock solutions and stability in the sample matrix and stresses the influence of storage conditions, matrix and container system on the stability, in addition to the intrinsic properties of the analyte itself. For example, the analyte stability evaluation in whole blood could be valuable in case an analyte is unstable in whole blood or adsorbs to cellular components during the collection procedure.

According to the EMA [ref 7] guide, a 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;
- Processed sample stability;
- On-instrument/auto-sampler stability;
- Stock and working solution 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] and FDA [ref 8] guides are 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.

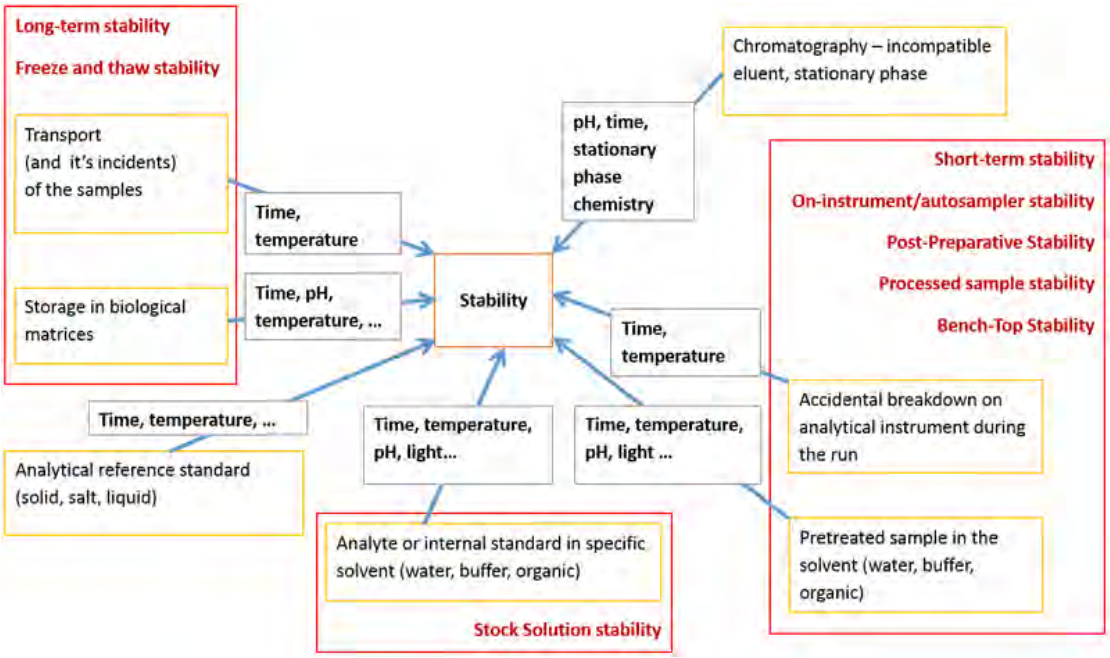


Figure 2. Different types of stability in LC–MS analysis and the chemical/physical parameters influencing stability.

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

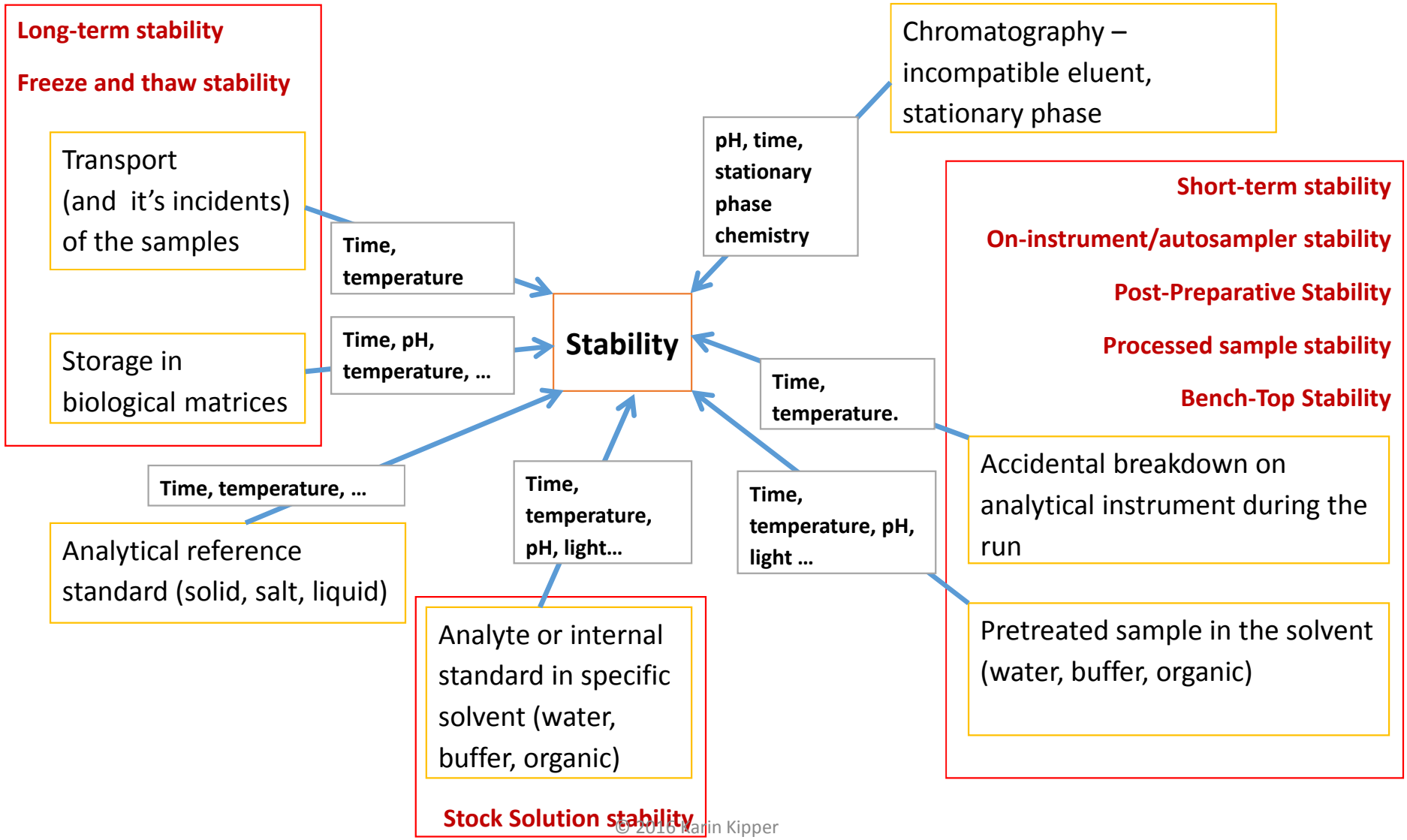
8.1_different_types_of_stability.pdf 46 KB

Different types of stability

Different types of stability according different guidelines

Types of stability according FDA	Types of stability according EMA
Freeze and Thaw Stability	Freeze-thaw Stability
Bench-Top Stability	Short-term stability at room temperature or sample processing temperature
Long-Term Stability	Long-term Freezer Stability
Stock Solution Stability	
Processed Sample Stability	

Different types of stability in LC-MS analysis and the chemical/physical parameters influencing the stability



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], dobutamine [ref 62])
Cytostatic nucleosides	Enzymatic deamination [ref 45]
Enantiomers	Racemization [ref 46]
Cis/trans isomers	Cis-trans interconversion [ref 47]
Beta-lactams [ref 10, ref 11, ref 63] (all penicillin derivatives)	Hydrolysis, enzymatic degradation, oxidation, presence of metal ions

In one example, the bench-top/short term stability of the 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 stored at -20 °C until loading into the autosampler, the stability (*ST*%) (see section 8.3) remained around 95% after 3 h and 92 % after 12 h storage in an autosampler (+4 °C) [ref 10]. Thus, storage of samples and solutions at low temperature is a powerful means of increasing stability.

In another example [ref 11], the stability of Penicillin G was examined over a 16 h time-period and also compared in different pH conditions since Penicillin G is well known for its instability in solutions and the stability of penicillins is also highly dependent on the pH of the solution. The bench-top/short term stability of the Penicillin G in the processed blood plasma samples in an autosampler thermostated at +4 °C showed rapid degradation: *ST*% was 20% after 16 h for samples with a pH 6.2. After the 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 an analyte stability.

Moreover, microbial contamination can lead to the degradation of the analyte as it was observed for two biomarkers thymidine and 2'-deoxyuridine in urine samples. No analytes were detected in urine after storage at ambient temperature for 24 h and at 4 °C for 14 days [ref 61]. The stability of samples was improved with the usage of perchloric acid.

Oxidation of dobutamine was observed when the stability of plasma samples were tested. After the addition of the ascorbic acid as the stabilizer to the plasma samples, a significantly increased sample stability has been onserved [ref 62]. Interestingly, the internal standard of dobutamine (dobutamine-D4) indicated a decrease of peak area over 115 injection run in samples where ascorbic acid was not added as a stabilizer. A stabilizer helped clearly to maintain the concentration of an analyte and an internal standard's concentration in the processed sample. However, this also means that suitable internal standards could potentially help to overcome analyte stability issues in case degradation of internal standard mimics the degradation of an analyte and additional analyte degradation would not take place before the addition of internal standard.

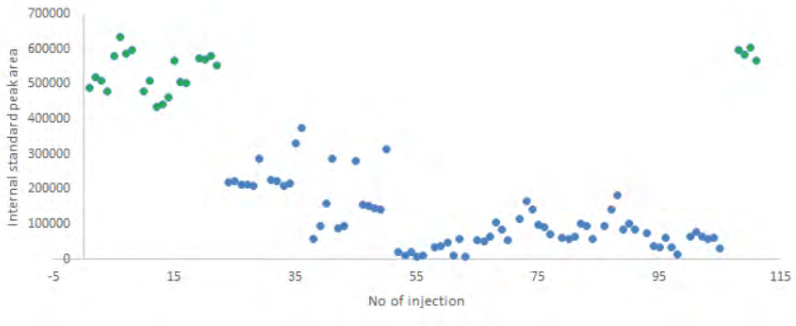


Figure 3. Dobutamine internal standard's (dobutamine-D_4) response in stabilised (green) and unstabilised (blue) plasma samples during the analytical batch run [ref 62].

8.2_examples_of_unstable_analytes.pdf 62 KB

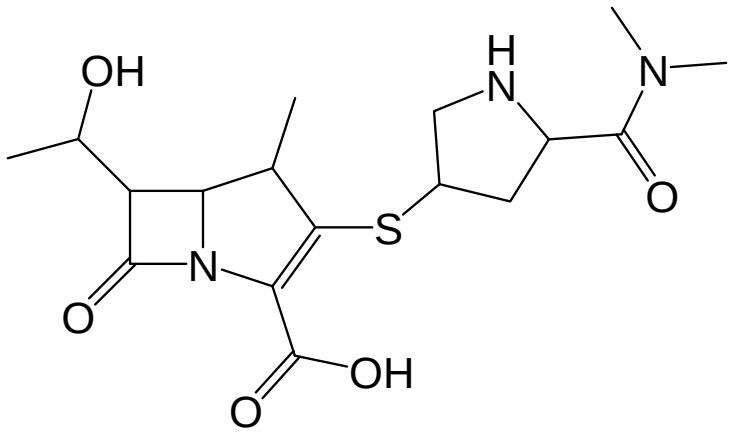
Examples about unstable analytes

Examples about unstable analytes

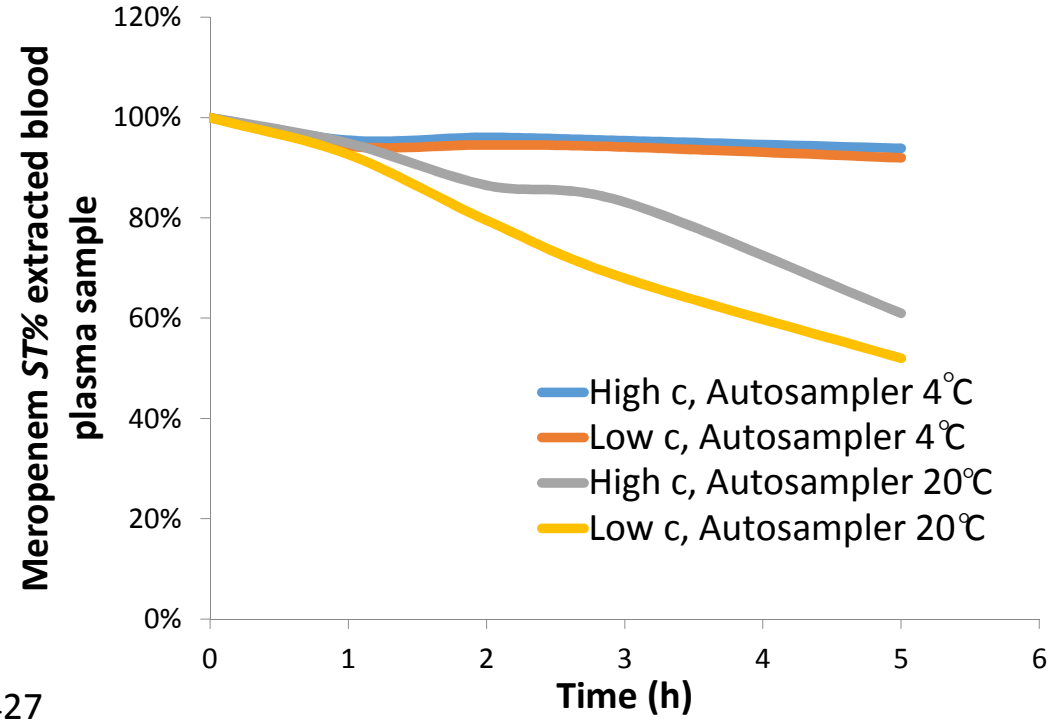
Compound	Issue
Acylclucoronide metabolites	Hydrolysis to parent drug, acyl migration
Lactones	Reversible hydrolysis to hydroxyl acid
Prodrugs	Conversion into active drug by ester bond cleavage
Amides	Hydrolysis to acid
Oxidizable compounds	Oxidative degradation
Cytostatic nucleosides	Enzymatic deamination
Thiols	Dimerization
Enantiomers	Racemization
Cis/trans isomers	Interconversion
Beta-lactams (all penicillin derivatives)	Hydrolysis, enzymatic degradation, oxidation, presence of metal ions [*]

*Kipper et al, Chromatographia, 70 (2009) 1423 - 1427

The stability of beta-lactam antibiotics – the effect of the temperature

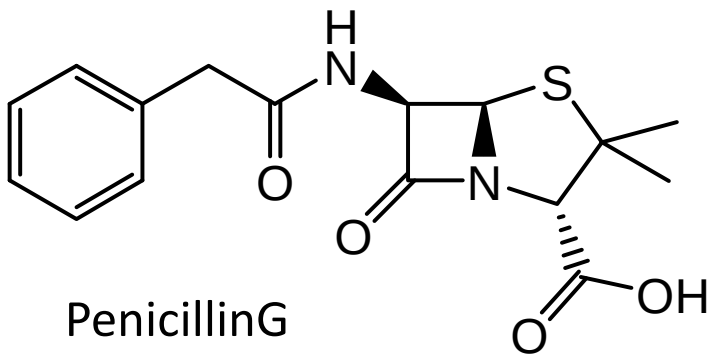


Meropenem

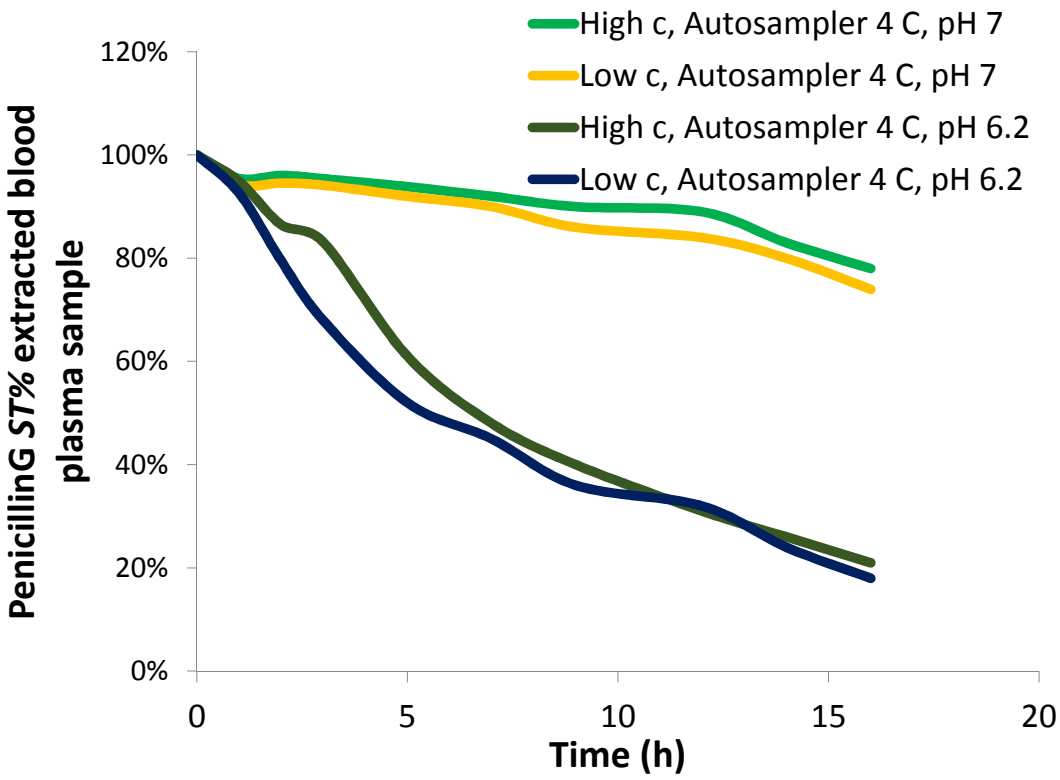


Kipper et al, Chromatographia, 70 (2009) 1423 - 1427

The stability of beta-lactam antibiotics – the effect of the pH



PenicillinG



Kipper et al, MSACL conference (2015)

8.3 Evaluation of stability

The time during which the stability is tested, is important, but it is usually not specified in the guidelines and only brief recommendations are given. The reason being 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 the 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 processed samples should be carried out covering the actual duration of the sample storage during the proposed research study. Analyte concentration should be evaluated over a specific storage period and in conditions that would improve the stability over time, should be identified.

The experimental design of the 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 the analyte or internal standard.**

Stability should be studied at least at two concentration levels – low and high concentration level and in a 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 the analyte stability under sample preparation conditions. Freeze-thaw stability is usually evaluated during three thawing cycles to predict the influence of possible delays and glitches on the sample handling.

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

The analyte or internal standard stability in the 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.

Moreover, stability can also be evaluated as %nominal of QC samples [ref 7, ref 8]. Analytes are considered to be stable stability remains within +/-15% [ref 7, ref 8] of analyte's nominal concentration.

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

8.3_evaluation_of_stability_and_overcoming_stability_issues.pdf 52 KB

Evaluation of stability

Experimental design of stability

- Time during which stability is tested, is important and can vary
  - Usually the autosampler stability injections are carried out overnight
  - Short time stability should be carried out at least over a stability analysis period of 48 h

Evaluation of stability

- At least at two concentration levels – low and high; matrix matched samples
- Bench-top stability/short-term stability – at room temperature or sample processing temperature
- Freeze-thaw stability – during three thawing cycles

- Stability should be evaluated at several different time points over the stability analysis period
- Samples should be analyzed in six replicates

Numerical expression of stability

- Evaluation via chromatographic peak areas

$$ST\% = \frac{S_t}{S_0} \times 100\%$$

- Evaluation via concentrations

$$ST\% = \frac{C_t}{C_0} \times 100\%$$

Numerical expression of stability

- The average percentage of analyte found in the sample under the specific conditions
  - The freshly prepared calibration standards are considered as containing 100% of the initial analyte content

Overcoming stability issues

Aspects to be considered while analytes are unstable

Measure	Challenge
pH control	Extract pH is crucial; Esterification of alcohols
Addition of stabilizers	Some additives itself are not stable Exact pipetting of stabilizers is required
Light protection	Working in dark or under yellow light
Reducing of sample processing time	Time critical process to be managed
Cold storage and handling	Every other step in the processing has to be cooled as well
Cold storage long-term (e.g. -70 ◦C; -80 ◦C)	Availability of -70 ◦C or -80 ◦C freezers
Derivatization (e.g. at sample collection)	Time critical and time consuming Completeness of reaction Matrix effects Reproducibility issues, IS or ILIS required

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	Sample extract pH may be crucial and therefore has to be monitored carefully. A possible problem may be esterification of alcohols. If a low pH is used for sample preservation, then the 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 themselves are not stable and therefore their concentration and thus also stabilizing efficiency changes in time. The precise quantity of stabilizers in the sample is required to enable the optimal stabilizing effect [ref 62].
Protecting samples and standards from the light	Keeping samples and standards on the bench for minimum time possible. Working in the dark or under the yellow/red light.
Reducing sample processing time	Time-critical processes have to be managed – e.g. optimization of the 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:

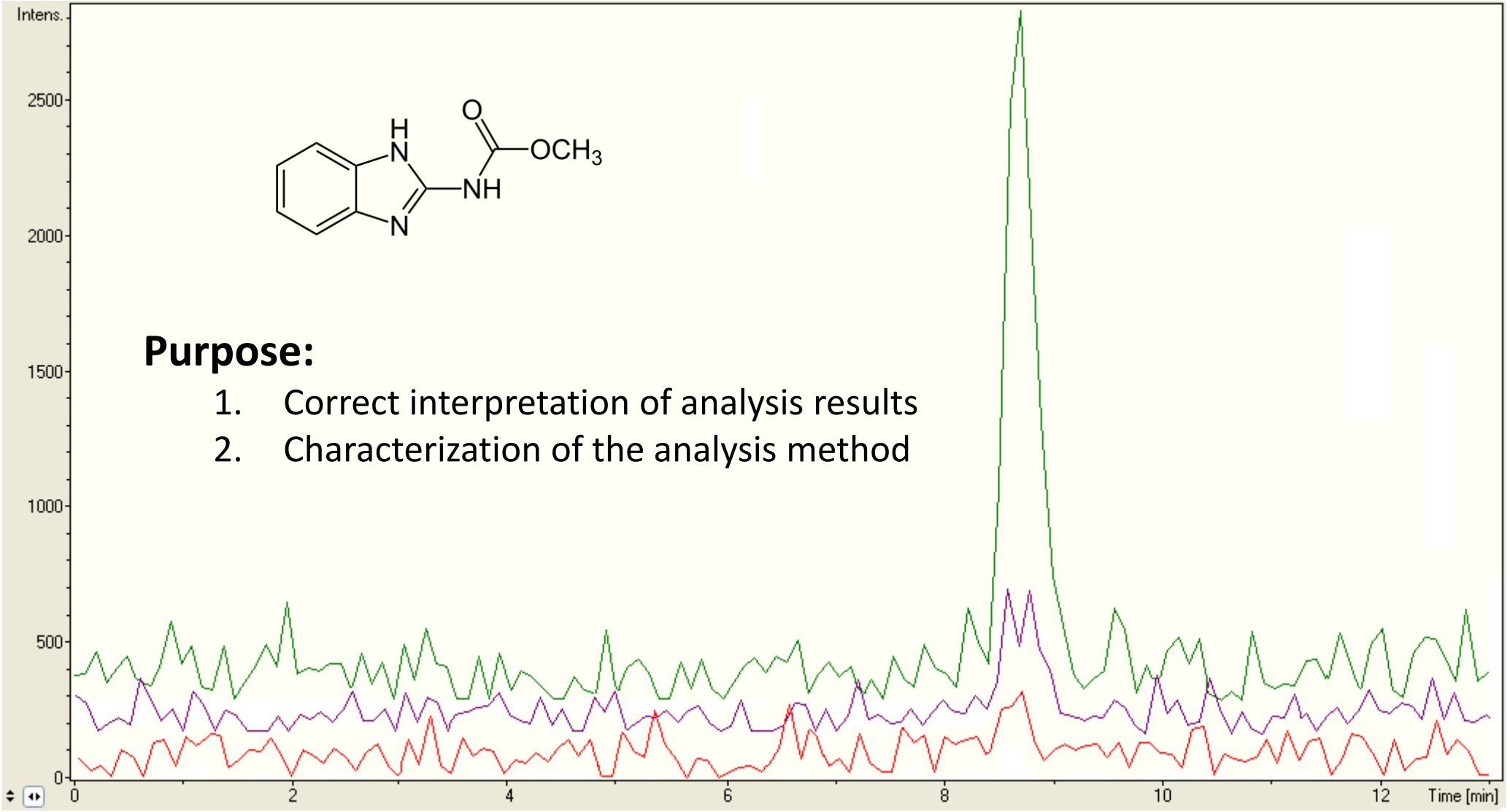
- Limit of detection (LoD) (also called detection limit) – the smallest amount or concentration of the analyte in the test sample that can be reliably distinguished from zero [ref 12].
- Decision limit (CC_α) – the concentration level, as determined by the method, at which there is a probability α (usually 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 the sample at which there is a 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 the 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

Limit of Detection

LC-MS/MS chromatogram of carbendazim (m/z 192 -> 160)



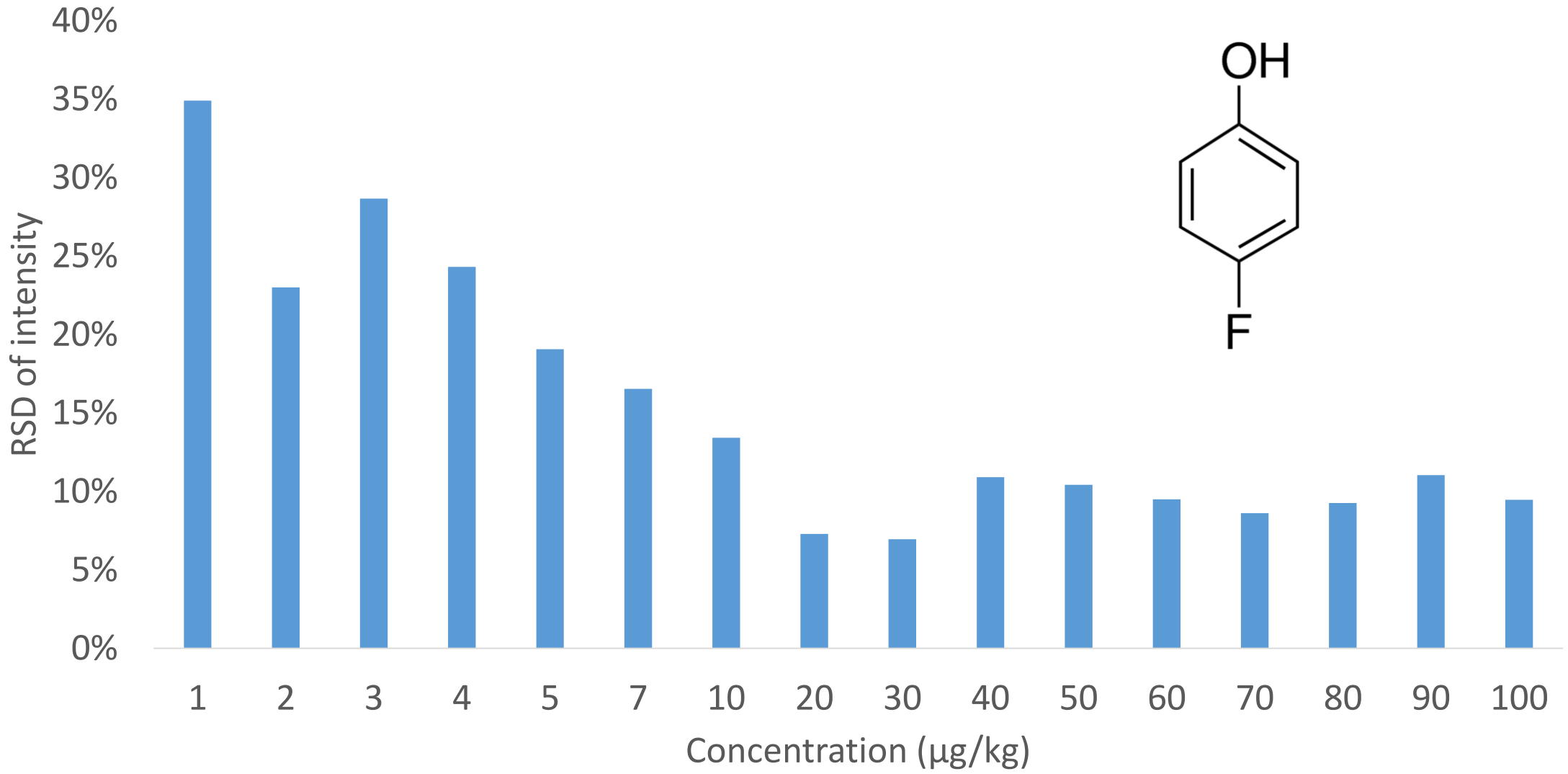
Purpose:

1. Correct interpretation of analysis results
2. Characterization of the analysis method

IUPAC definition

In broad terms, the detection limit (limit of detection) is the smallest amount or concentration of analyte in the test sample that can be **reliably** distinguished from zero.

Limit of Quantitation

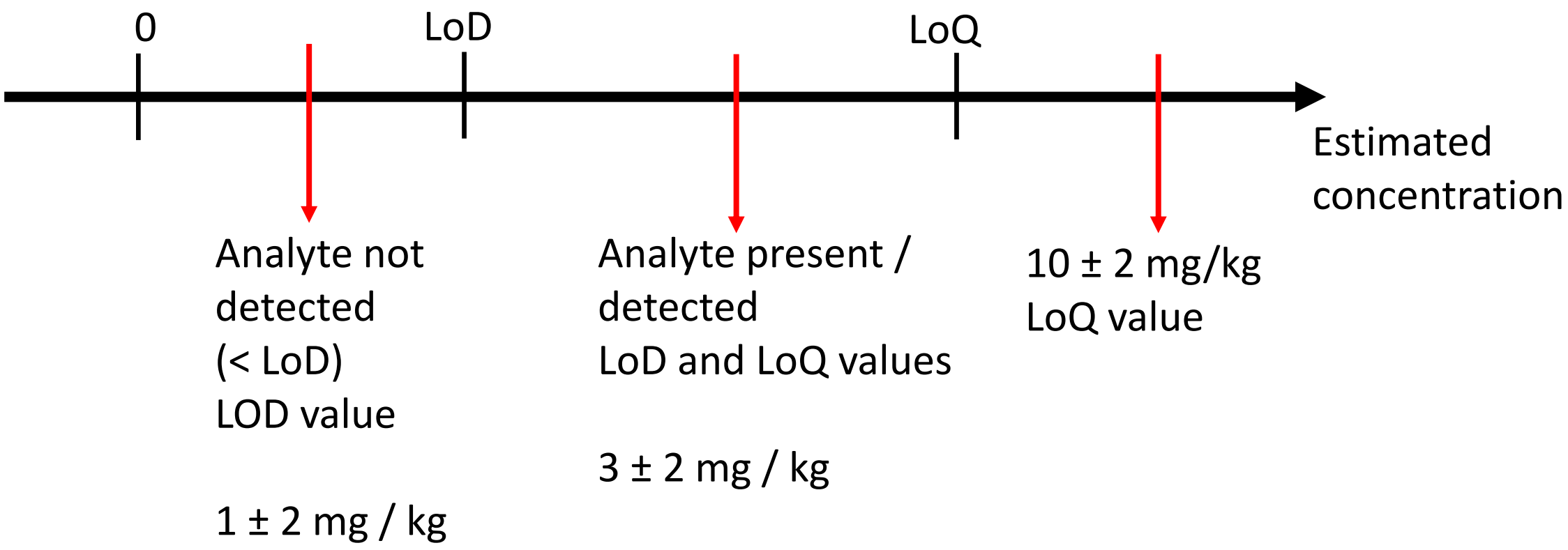

40%
35%
30%
25%
20%
15%
10%
5%
0%
RSD of intensity
1
2
3
4
5
7
10
20
30
40
50
60
70
80
90
100
Concentration (µg/kg)
OH
F

Limit of Quantitation (LoQ)

- Definition by Eurachem:
  - The lowest concentration of analyte that can be determined with an acceptable repeatability and trueness
- Repeatability and trueness limits for LoQ can be set by relevant guidelines or standards
  - For example SANCO demands ≤ 20% repeatability and trueness between 70-120%
- Quantitation below LoQ is possible
  - In range of LoD uncertainty becomes large, comparable to the result

Interpretation of analysis results with LoD and LoQ

Interpretation of analysis results with LoD and LoQ

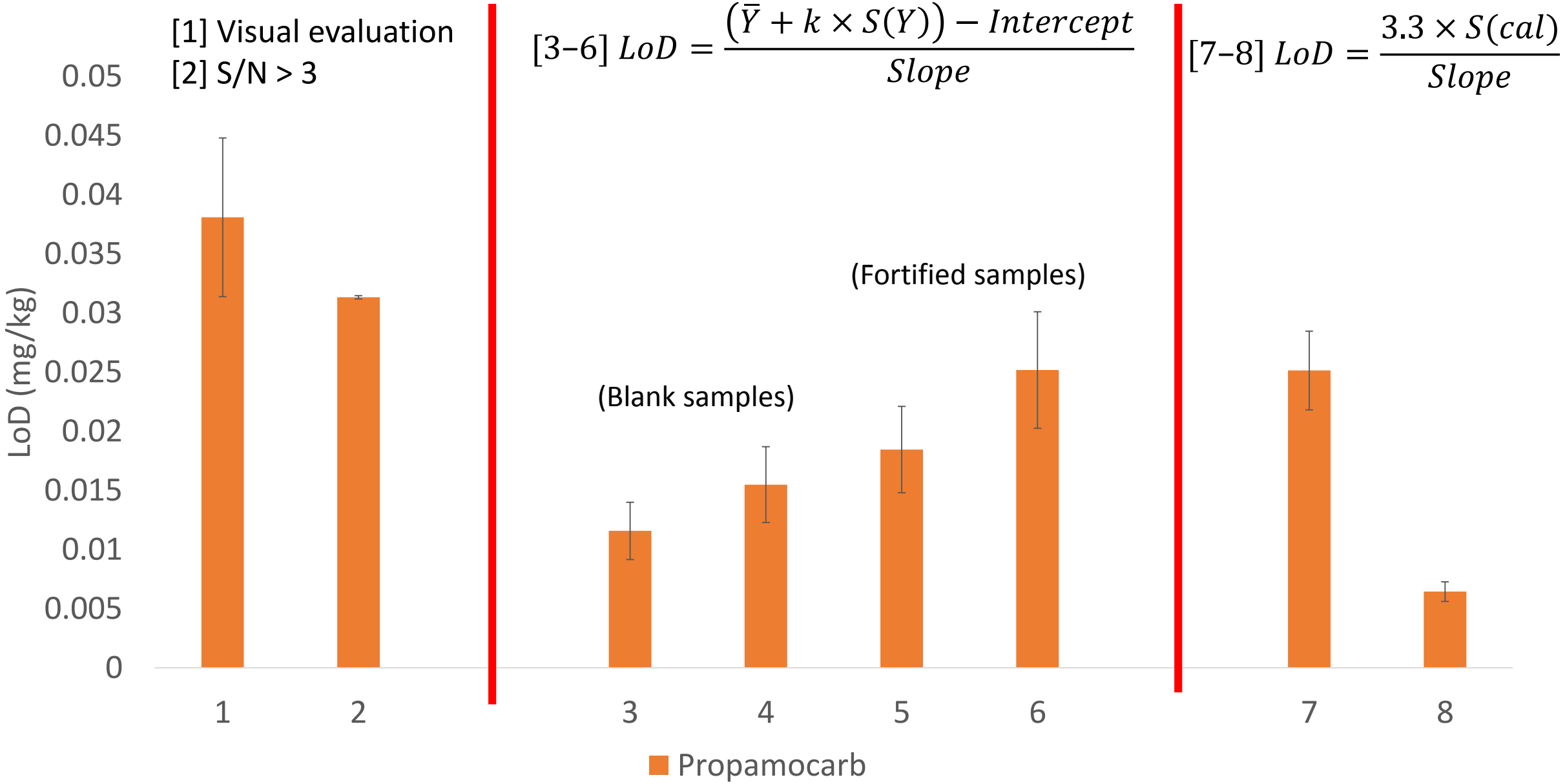


Different approaches to estimate LoD

Important aspects

- Guidelines give different approaches
  - E.g. FDA, IUPAC, Eurachem, NordVal, US EPA, etc.
  - Not all approaches are fitting for all analytical methods

- Different approaches make different assumptions


[1] Visual evaluation
[2] S/N > 3
[3–6] $LoD = \frac{(\bar{Y} + k \times S(Y)) - Intercept}{Slope}$
[7–8] $LoD = \frac{3.3 \times S(cal)}{Slope}$
(Blank samples)
(Fortified samples)
LoD (mg/kg)
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
1
2
3
4
5
6
7
8
Propamocarb

Conclusions

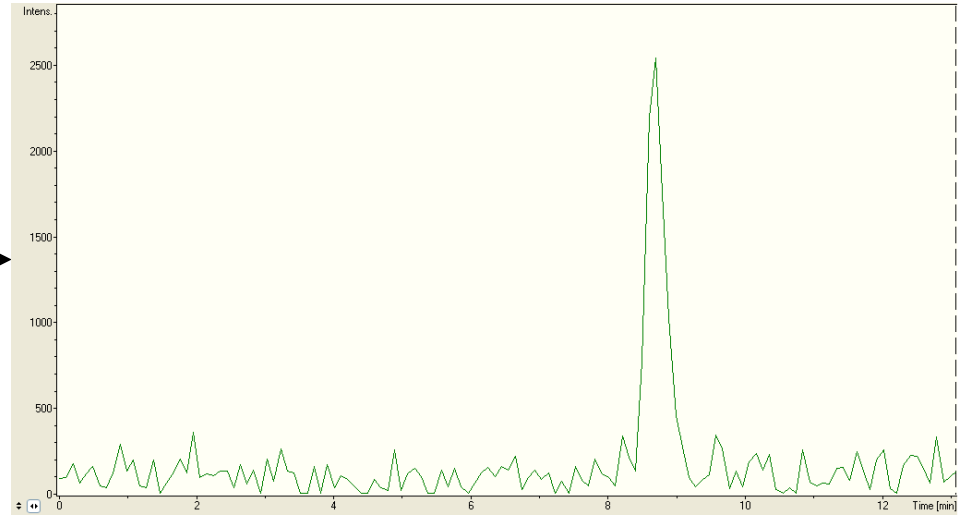
- Results of different approaches are not comparable

- LoD depends on
  - Variance
  - Slope and intercept
  - Only an estimate of LoD can be found

Instrumental LoD and method LoD

Instrumental LoD:

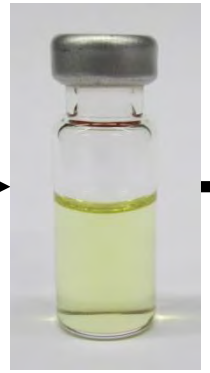
MEASUREMENT



Method LoD:

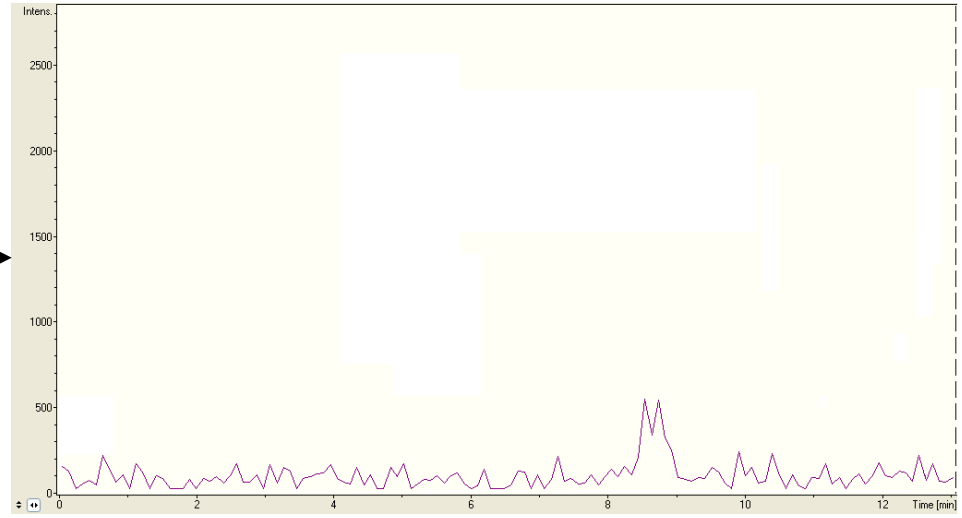
SAMPLE PREPARATION

Recovery
(Loss of analyte)



MEASUREMENT

Matrix effects



Lead to increased
variance and higher LoD

Conclusion

- For a whole analysis method instrumental LoD is not suitable
  - Blank matrix matched samples must be used
  - All samples must go through the whole method
- Similar conclusions for LoQ

Decision limit (CC_α) and detection capability (CC_β)

CC_α and CC_β

- The general definition of LoD is ambiguous
  - False positive and false negative results

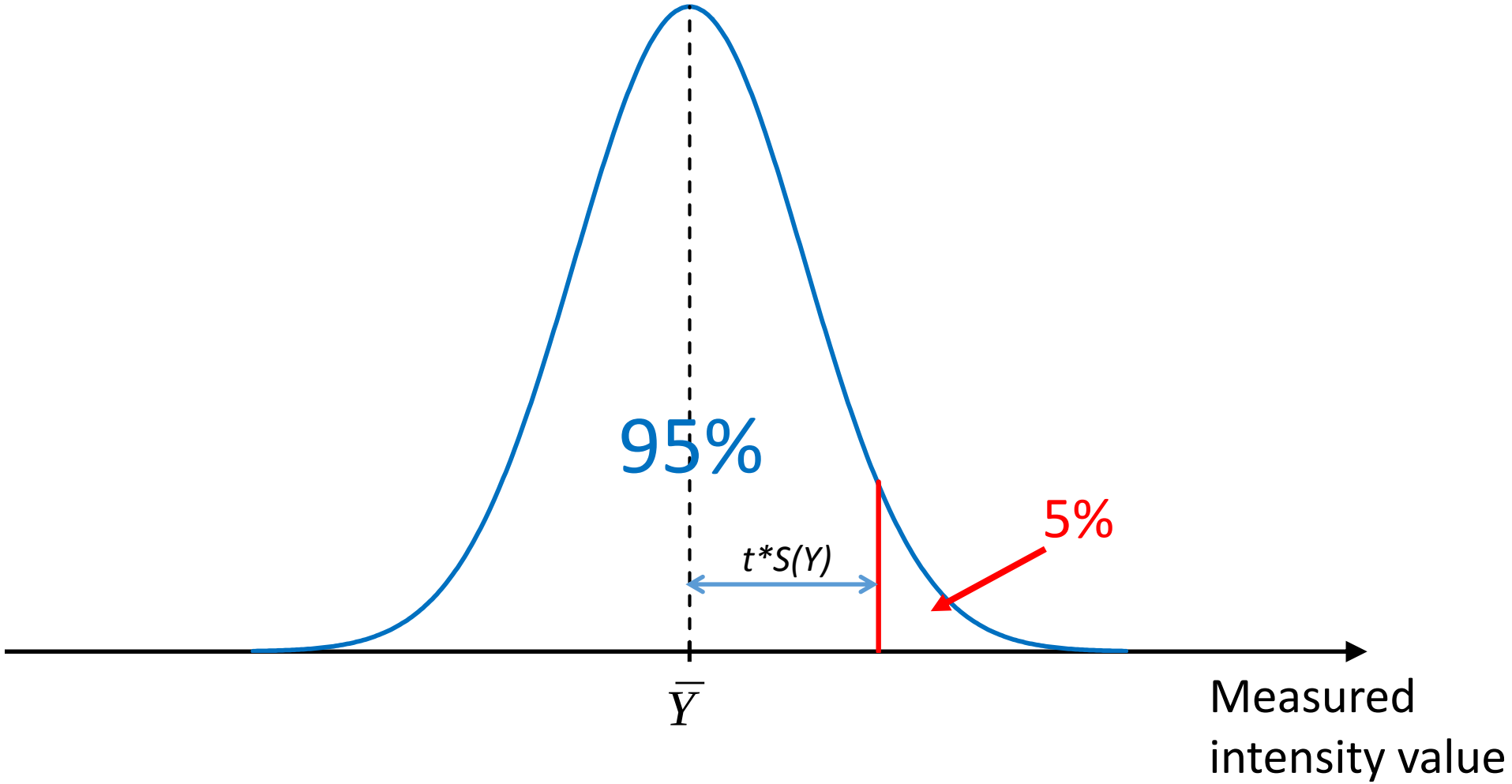
	A positive result is received	A negative result is received
Sample is truly positive	✓	False negative result
Sample is truly negative	False positive result	✓

Definitions

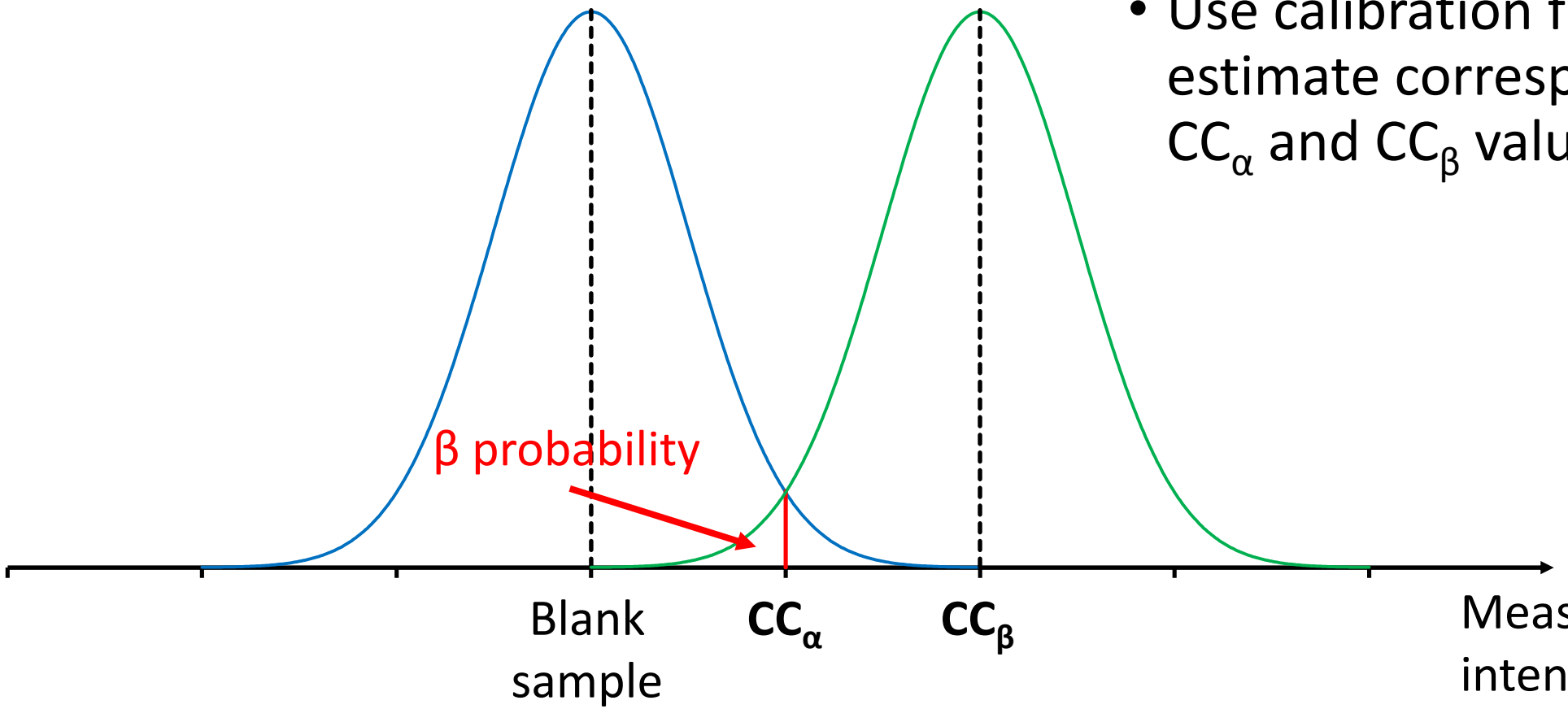
- Decision limit (CC_α) – analyte concentration level above which we can state that the signal is caused solely by the noise with the probability below α
  - α = 5% or 1%

- Detection capability (CC_β) – analyte concentration level in a sample above which there is less than β probability that the result will be randomly below CC_α (and therefore interpreted as a negative result)
  - β = 5% or 1%

Normal distribution of measurement results



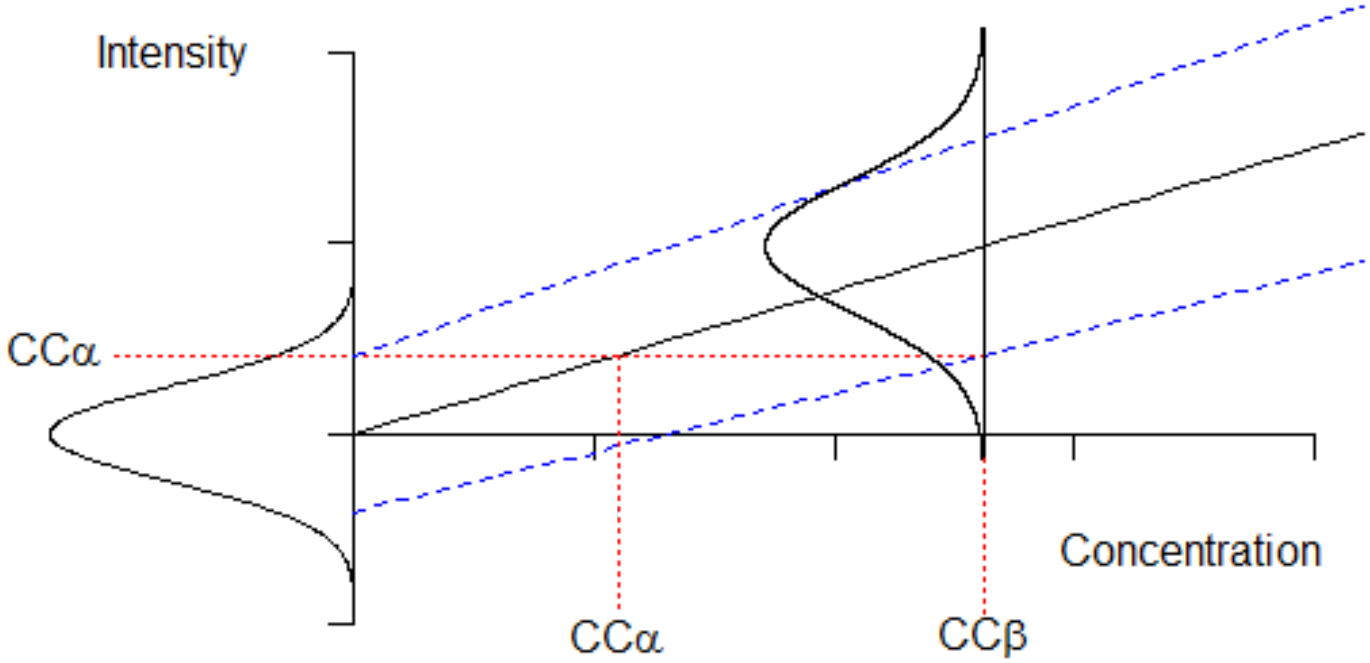
CC_α and CC_β



- Use calibration function to estimate corresponding CC_α and CC_β value

Calculating CC_{α} and CC_{β}

Calculating CC_α and CC_β

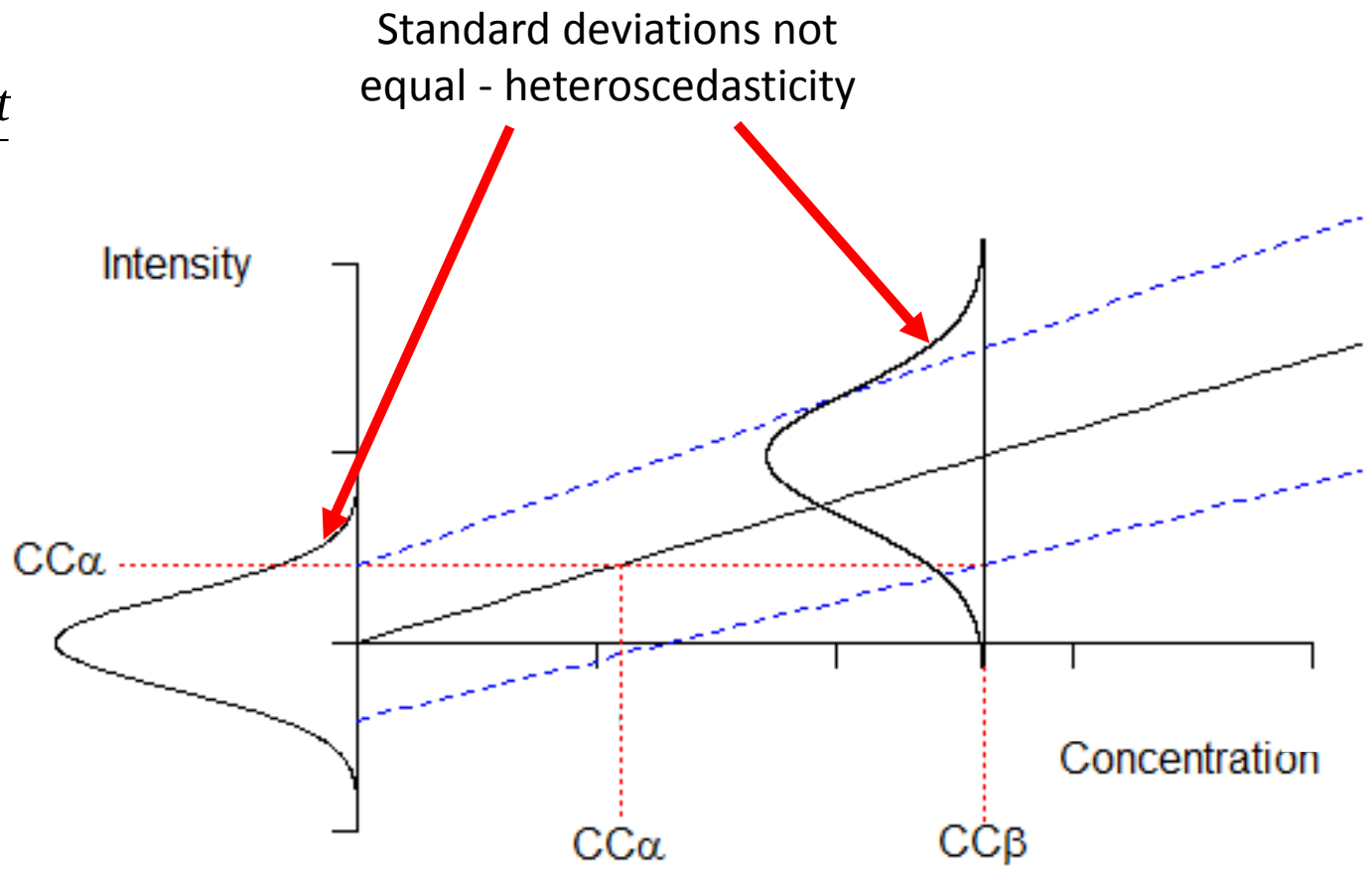


$$CC_{\alpha} = \frac{(\bar{Y}(blank) + t \times S(blank)) - Intercept}{Slope}$$

$$CC_{\beta} = \frac{(\bar{Y}_{CC\alpha} + t \times S(Y_{CC\beta})) - Intercept}{Slope}$$

$$LoD = \frac{(\bar{Y} + k \times S(Y)) - Intercept}{Slope}$$

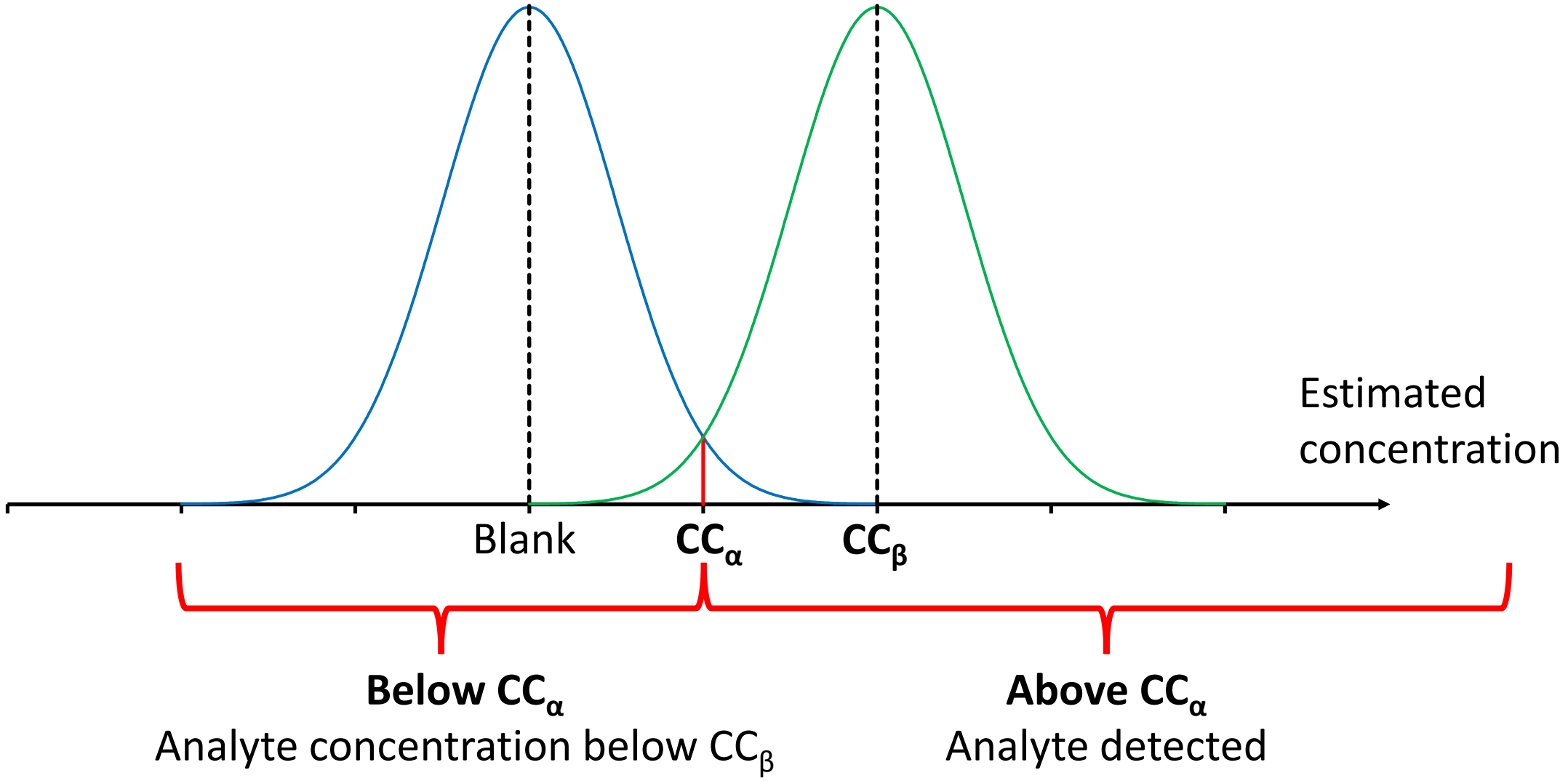
$$LoD = \frac{3.3 \times S(cal)}{Slope}$$



- Complex approaches with less assumptions and simplifications exist
  - ISO 11843-2

Interpreting results with CC_{α} and CC_{β}

Interpreting results with CC_α and CC_β



Conclusion

- CC_{α} – for making the decision (analyte detected or not)

- CC_{β} – for characterizing the approach
  - When analyte is not detected
  - For comparing different analytical methods
  - For comparing a method with a set limit

- With decision also give
  - CC_{α} and CC_{β} values
  - Result with uncertainty if necessary

Important aspects of estimating LoD and CC_{α}, CC_{β}

Choosing between LoD estimation approaches

1. Is LoD necessary?

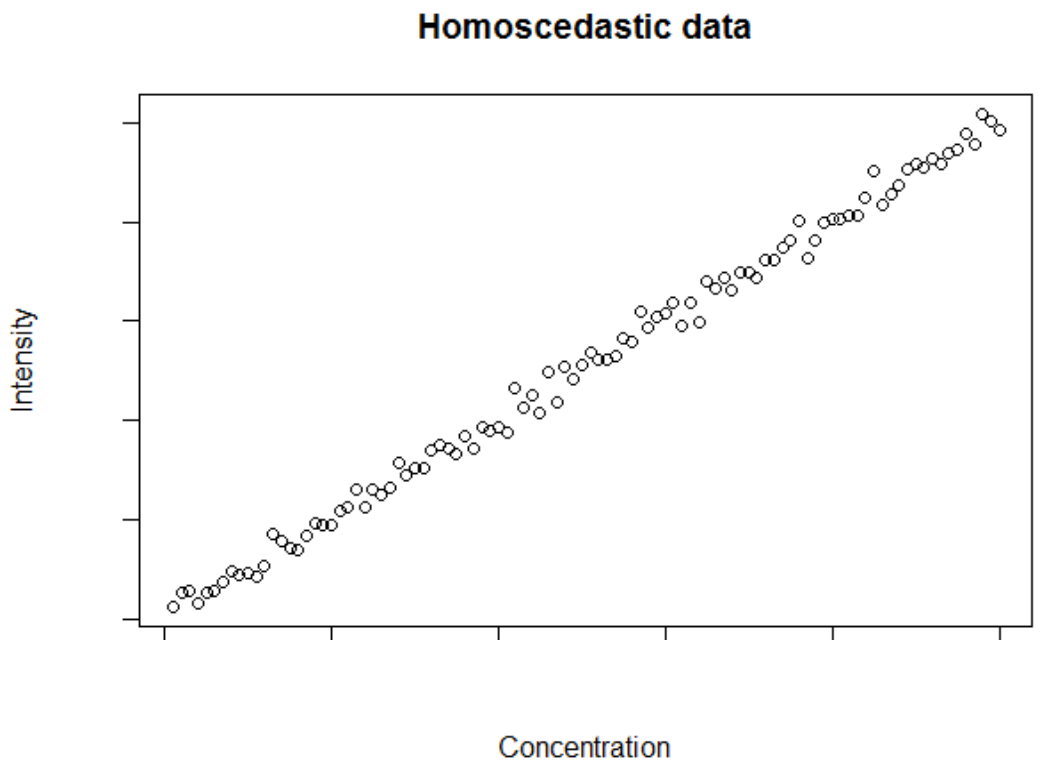
2. Should a standard (or guideline) be followed?

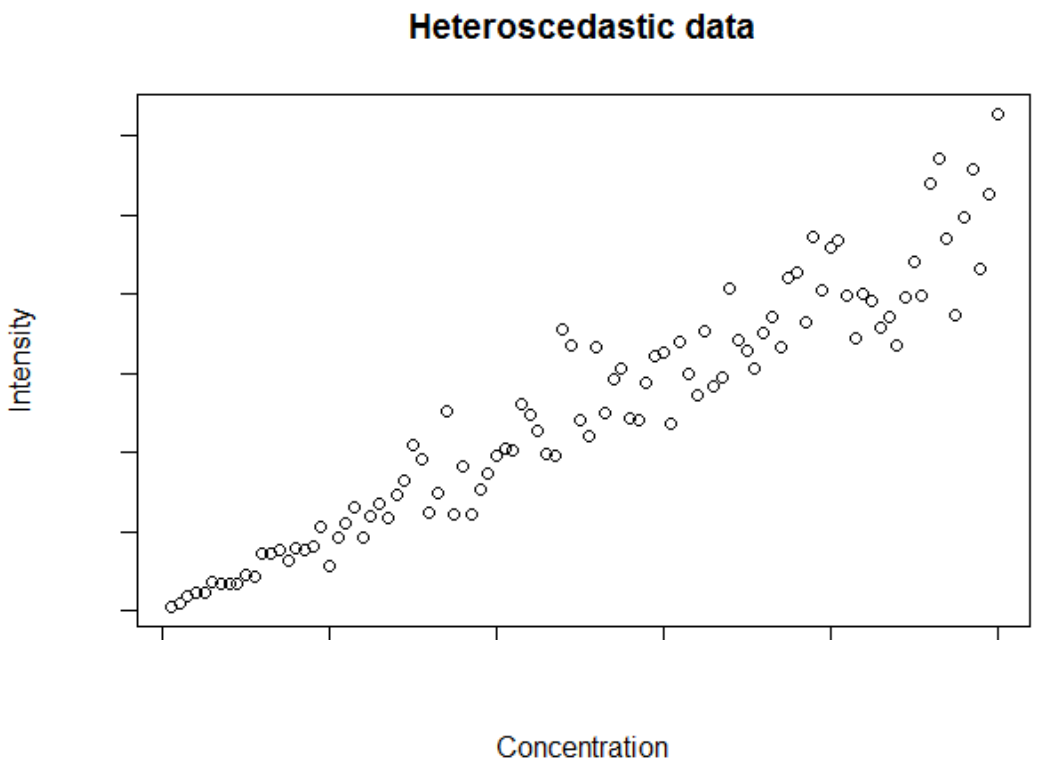
3. Are critical decisions based on LoD?
   - If “Yes” then estimate CC_{α} and CC_{β} (procedure given in ISO 11843-2)
   - If “No” using the following equation is suggested:

$$LoD = \frac{3.3 \times S(residuals)}{Slope}$$

Linearity and scedasticity

- Data should be in range of LoD and linear
- Data should be homoscedastic
  - Use narrow concentration range





Other important aspects to consider when estimating LoD (CC_α and CC_β)

- LoD varies between measurement series and days
  - Regular reevaluation is recommended

- Not all approaches are appropriate for all analytical methods
  - Integration of blank samples with LC-MS/MS

- Matrix matching of used samples

Different approaches to estimate LoQ

LoQ estimation approaches

1. Trueness and precision at each concentration

- Preferable, but labor-intensive
- Different relevant guidelines or standards set precision and trueness limits
- Can be based on uncertainty at different concentration levels

2. Approaches related to LoD estimation

- Same data for LoQ, LoD
- *k* values different in guidelines
- Trueness and precision are not estimated
- S/N and visual evaluation

$$LoQ = \bar{Y} + k \times S(Y) \qquad LoQ = \frac{10 \times S(cal)}{Slope}$$

LoQ estimation

- Choice depends on importance of LoQ parameter
  - When necessary specific guidelines must be followed

- Our recommendation:
  - If LoQ is critical use precision and trueness estimation
  - Otherwise use ICH suggested approach

$$LoQ = \frac{10 \times S_{y.x}}{Slope}$$

Important aspects of LoQ estimation

- LoQ is used for:
  1. Is the sample concentration high enough for “fit for purpose” quantitation?
  2. To characterize the analytical method
     - The used LoQ estimation approach must be stated

- When estimating LoQ:
  - Use data in range of LoQ
  - Use matrix matching samples
  - LoQ changes between measurement series and days

- For approaches that use calibration function:
  - Linearity and scedasticity

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 the limit of detection (LoD) (also called detection limit) as the smallest amount or concentration of the 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 a 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 a decision limit (CC_α) and a detection capability ($CC\beta$).

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.

It can be seen, that 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 the analytical methods (how low analyte levels is the method able to detect).

In Table 1 (of Section 9.3) different approaches of how to estimate LoD with detailed information are presented. Only widespread approaches, taken from the prominent guidelines and articles, are presented, thus it is not an exhaustive list of all possible approaches. Approaches can be based on the 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. How to choose 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 used. Section 9.3 gives guidance relevant for LC-MS methods.

In many cases regulations set the maximum allowed limits for the 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) from which analyte content in the sample, as *determined by the method*, 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 the LoD value. 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 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 the 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

A distinction has to be made between the instrumental LoD and the method LoD. Instrumental LoD is found from the analysis of the analyte in pure solvent. Therefore, instrumental LoD only shows the capability of the instrument to detect the analyte and can only be used 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 a 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)(1) This is in most cases when analysis is performed to establish whether the analyte level in the sample is above or below the 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. – 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, then it can be said that the analyte is present at trace levels, 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, then simply stating that the analyte is present in the sample and not providing any quantitative information, only gives part of the information available from the analysis. However, it is often still suggested that the measured value and its uncertainty should be reported even if the results are below LoQ (or even LoD), so that the quantitative data can be used for further calculations if necessary. For example, in case of pharmacokinetic studies it has been suggested that results below LoQ can be given out so that better pharmacokinetic models could be built. Further discussion on the topic can be found at the following reference [ref 58].

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 the 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 may or may 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.

An example is given in Table 1 to better explain the concept of false positive and false negative results on the basis of a widely banned pesticide DDT (dichlorodiphenyltrichloroethane). The concentration of DDT is compared to the capability of the analytical method.

Table 1. Explanation of false positive and false negative results.

	DDT is found to be present	DDT is not found to be present
Sample has enough DDT in it for the analysis method to detect it	Everything is OK.	**False negative results**: DDT is present at high enough concentration for the analysis method to detect it, but the analysis has given a result that DDT concentration is below the detection capability of the analytical method.
Sample does not have enough DDT in it for the analysis method to detect it	**False positive result**: DDT is not present in the sample or at least is below detectable level, but the analysis has given results that DDT has been detected.	Everything is OK.

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). For example, when validating the analysis method for measuring DDT, it was found that CC_α was at 0.005 µg/L. Therefore, if a higher result than 0.005 µg/L is received from the analysis there is ≤ 5% probability that the result is just noise (noise can be defined as signal received when no analyte is present or the analyte concenrtation is so low that the analysis method cannot detect it). If we are relatively certain that the results are not simply noise, then we can say that we have detected DDT. In other words, CC_α can be considered as the concentration at which we can decide that we are not measuring noise but we are receiving a signal from the analyte.

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 a 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 CC_β definitions).

For example, in the example of DDT analysis method, CC_β was found to be Y. This means that if a result above Y is received then the probability that the measurement is just noise is ≤ 5%. Because of random variability of results, even if DDT is present in the sample at a concentration above CC_α, there is the possibility that we receive a result that is below CC_α and therefore interpret the result as negative (see the red curve in Figure 1). Therefore, we need to find CC_β where the probability for this kind of false negative probability is small, in our example ≤ 5%.

By these definitions, if the analyte is present at concentration of CC_β (or above) it can be said with good certainty that the measurements are not simply noise (which otherwise would lead to false positive results) and that the analysis method can measure this concentration with only small probability of giving results below CC_α (i.e. false negative results).

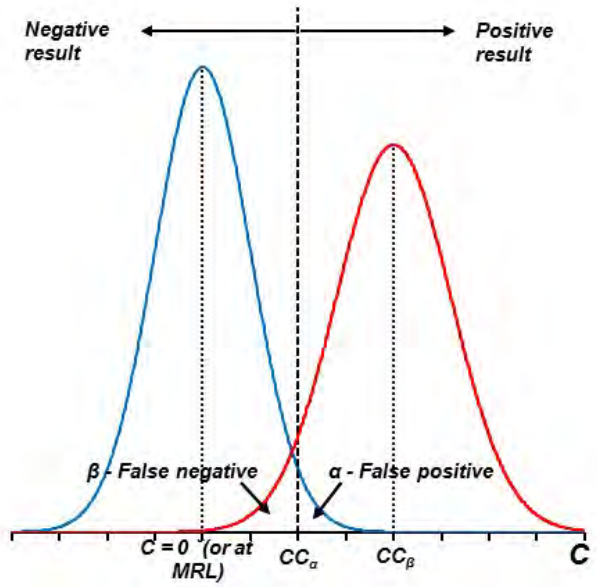


Figure 1. Relation between permitted limit, CC_α and CC_β

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. 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 to estimate CC_α and CC_β are usually provide more complex and do not make the same assumptions. Therefore, in principle LoD and CC_β could be considered equal, but their values cannot always be compared in a meaningful way.

If the 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 analyte concentration should be estimated where it can be said with known confidence (taking into account both false positive and false negative probabilities) that the concentration of the analyte is above the MRL. For the second part, CC_α and CC_β must be evaluated not for 0 concentration but at the MRL level (see Figure 1 where C = 0 or **at MRL**). 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 be used.

In Figure 1, a normal distribution is used only to clearly and simply explain the concept of CC_α and CC_β. It has been shown that often results from blank and very low concentration samples do not have a normal distribution. Moreover, it is more correct to use a t-distribution rather than a normal distribution in case if only a small number of replicate measurements are made. In addition, it should be noted that the data is heteroscedastic (width of the distribution is wider at higher concentration) in Figure 1 and homoscedastic (width of the distribution is same at both concentrations) in the video "Decision limit (CC_α) and Detection capability (CC_β)".

Interpreting results with CCα and CCβ

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

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

With CC_α and CC_β, the results must be interpreted 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, and no clear consensus exists on which approach is the best in different situations. The approach that we recommend in this course is discussed below in the video "Important aspects of estimating LoD and CC_α, $CC\beta$" and in the tutorial review [ref 19 and ref 20]. A general overview of approaches from most prominent guidelines to estimate LoD from most prominent guidelines can be found in Table 1 (NB! For more specific overview of procedures, see the specific guideline!). 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.

An Excel sheet with example calculations of LoD with approaches in Table 1 can be found at the end of this chapter.

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

Group	Ref.	What is obtained?	Equation
1	[ref 3, ref 9, ref 12, ref 15]	LoD (considers false positive and negative results – the probability of false positive and negative values depends on the 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 the replicates;
variability of the slope and intercept are not taken into account;
linearity of the 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 $\bar{y}_0$ is not necessary (i.e. $\bar{y}_0$ is equal as 0) if subtraction with intercept (or with $\bar{y}_0$) is done to all measured results before calculations.

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 the analyte concentrations from the 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 the degrees of freedom;
Recommended analyte concentration range in fortified samples is 1-5 times LoD.

Assumptions, simplifications: Normal distribution of replicates; variability of the slope and intercept are not taken into account;
linearity of the 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	$LoD = \bar{a} + t \times \frac{S(y)}{\sqrt{n}}$ **(Eq 3)**

		results – the probability of false positive and negative values depends on choice of *t*)	$LoD = t \times S(y) \times \sqrt{\frac{1}{n} + \frac{1}{n_b}}$ **(Eq 4)** where $\bar{a}$ is the average intercept; *n* is the number of repeated measurements of the sample; *S*(y) is the standard deviation of the 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 the replicates;linearity of the calibration data; variability of the 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 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 a standard deviation of the blank samples, residuals ($S_{y.x}$) or intercept. Instructions to calculate a standard deviation of the residuals and an intercept in Excel can be found in the video "Calculating LoD" below.

Read more:

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

Assumptions, simplifications: Homoscedasticity;
normal distribution of the replicates;
linearity of the calibration data; variability of the slope and intercept are not taken into account.
If repeated results at each calibration level are averaged and standard deviation of the 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 the intercept underestimates the variance of the results at 0 concentration and should not be used.
Due to the 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 the 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 the presence of a peak depends on the analyst.

Notes: This approach is very work intensive;
If repeated estimations of the 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 the 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 intermediate precision of the intercept (if MRL has been set then data at and above MRL must be used)

2. Blank matrices are analyzed to estimate the 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 intermediate precision 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 a MRL is set;
After estimating the intensity value corresponding to CC_α and CC_β, the calibration function should be used to convert them to the concentration scale;
Approach 2 for estimating the CC_α and approach 3 for estimating CC_β demand at least 20 replicates (at each level for CC_β).

Assumptions, simplifications: Normal distribution of the replicates;
linearity of the calibration data; variability of the 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	$CC_\alpha = t_{0.95}(\nu)\frac{\hat{\sigma}}{\hat{b}}\sqrt{\frac{1}{K}+\frac{1}{I\times J}+\frac{\bar{x}^2}{s_{xx}}}$ **(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)** $s_{xx} = J\sum_{i=1}^{I}(x_i-\bar{x})^2$ **(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 $\nu = I*J - 2$), 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. These equations are for homoscedastic data, for calculations in case of heteroscedastic data see [ref 18].

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 the replicates;
linearity of the 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 the 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 the concentration scale (similarly to LoD) the variance of the slope and intercept must be taken into account when estimating them.

Another property that must be considered 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 the concentration (see example in Figure 1). Analytical methods are often heteroscedastic – as the concentration increases, the standard deviation of the measurements also increases. If it is shown that the collected calibration data collected is heteroscedastic then, weighted linear regression (WLS) should be used to take the variance of the slope and intercept more accurately into account. A simplified approach that usually works sufficiently well is presented below.

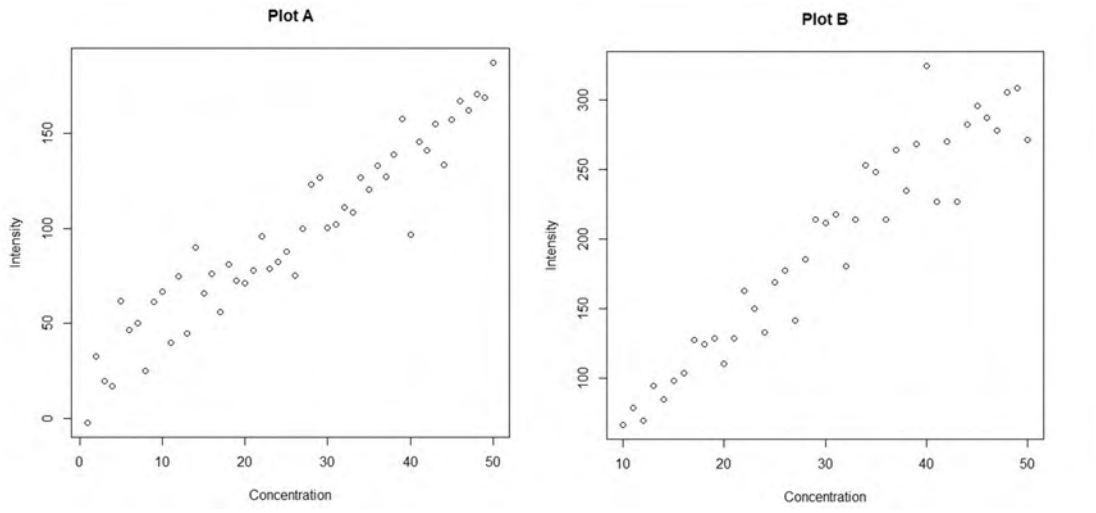


Figure 1. Data in Plot A are homoscedastic and data in Plot B are heteroscedastic. In plot A as the concentration increases the variability of results in intensity scale does not increase. However, in Plot B the variability of intensity values increases as the concentration increases.

With WLS the propagated errors of the 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 the slope and the 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 simpler.

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 a 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 a 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 complexometric titration, we do not have to worry that in some samples the concentration of calcium might be so low that it would be below a LoD for a reasonable titration procedure. However, if the LoD estimate is an important parameter used to interpret the results of an analysis, more complex and accurate approaches must be used to estimate LoD. For example, when analyzing blood samples of athletes for doping, the method must interpret the results correctly even if only very small amounts of the 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 the 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 the intercept value or from the standard deviation of the 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 the 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, the 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 the results on that day, then the 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 the LoD is used only for simple characterization of the method and not used further (see above), then the 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 the 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 the 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 the 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 the LoD and due to the 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

In case you have trouble with LINESt function in excel, we recommend you to review the following video.

lod_video_excel_solved.xlsx	21 KB
lod_calculation_approaches_example.xlsx	83 KB
lod_video_excel_unsolved.xlsx	16 KB

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 the similar approaches that are used to estimate LoD. However, in the case when a specific guideline must be followed, then the LoQ must be estimated by following the guideline.

First let's discuss the 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% at most and trueness of ±20% (i.e. the average result should stay between 80-120% of the reference value), 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 the 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:

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

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 the 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 labour-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 the range of LoQ. If a calibration function is used to estimate it, then linearity and scedasticity have similar effects on the LoQ estimation as on the 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 the 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 *Note 1

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

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

***

***Note 1**: NB! In this video we do not consider the statistical significance of intercept (intercept < 2 times the standard deviation of intercept). Although this is a simplification, the difference between the results when intercept is or is not set to 0 is not significant. i.e. the variance between the LoQ values obtained on different days is significantly larger than the difference caused by the change of the intercept value.

loq_video_excel_solved.xlsx	38 KB
loq_video_excel_unsolved.xlsx	29 KB

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 the 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 despite 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 the robustness is the stability of the method against the variability of the 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 the variability of the 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

Robustness and ruggedness introduction

Definition

- Robustness and ruggedness - the ability of an analytical method to remain unaffected by small variations in method parameters and influential environmental factors and characterize its reliability during normal usage
  - no change of the detected amount of the analyte in a certain sample in spite of the variation of the method parameter
  - no change of the critical performance characteristics (e.g. limit of quantitation) by the variation of the method parameter

Robustness or ruggedness?

- Robustness and ruggedness definitions in the guidelines are very similar
- Both terms – robustness and ruggedness – are in use.
- When used together they are treated as synonyms in most cases
- Robustness – also the feasibility to reproduce the analytical method in different laboratories

Robustness or ruggedness?

- **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

- **Robustness** should be used as a parameter characterizing the stability of the method with respect to variations of the internal factors (parameters) of the method

Robustness/ruggedness

- Sometimes robustness/ruggedness are misinterpreted and actually decision threshold, detection capability or measurement uncertainty is evaluated

- The term **robustness** for expressing the stability of the method against small variations of the intrinsic method parameters

- The term **ruggedness** for expressing the stability of the method against extraneous influencing factors

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 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 contribute 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)(1) This list is not exhaustive, there can be different other parameters that, depending on situation, could be changed.	**Likelihood of uncontrollable change (2)(2) This grading is generic, corresponding to the most common situations. In specific cases the grading can be different.**	**Possible extent of variation (3)(3) These extents of parameter variation can be considered as possible defaults. In practice different extents can be used, but they should be realistic.**	**Comments**
Liquid chromatography			
Mobile phase pH	Medium	± 0.5 units	pH will have a strong effect on the 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 the 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 the 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 the column can influence the retention time (and possibly resolution) and peak shape. There can be also influence on the 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 the analyte is taken up during the last stage of the sample preparation. This composition can influence the 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)(4) In the case of UHPLC the smaller changes in the injection solvent have stronger influence, especially if larger injection volume is used. 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 the 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 the analyte ionization efficiency in the ion source
Ion source configuration (Nebulizer position)	High (if configurations can vary) Not applicable (if fixed source)	According to the ion source design. Should be varied if the source is used in different configurations.	Ion source configuration can influence the spray and ionization efficiency in the ion source
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 the 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

Robustness and ruggedness relation to LC-MS method development and validation

LC-MS parameters

- Large number of LC-MS parameters
  - Changes in some of them result in poor reproducibility

- Complex and multi-step sample preparation procedure

LC parameters

Parameter	Likelihood of uncontrollable change	Recommended extent of variation
pH	Medium	± 0.5 units
Concentration of additives in eluent	Medium	± 10% (relative)
Organic solvent content in the eluent	Low to Medium	± 2%
Column temperature	Low	± 5 C
Eluent flow rate	Low	± 20%
Column batch and age	Medium	-

Samples and sample preparation

Parameter	Likelihood of uncontrollable change	Recommended extent of variation
Analyte extraction time; solvent amount and composition (in liquid/liquid and solid phase extraction, etc)	High	± 20%
Injection solvent composition	Low/High	± 10% (relative)
Matrix effect in broad sense (sample matrix source)	High	6 different

Mass spectrometry

Parameter	Likelihood of uncontrollable change	Recommended extent of variation
Drying gas temp	Low	± 10 C
Nebulizer gas pressure/flow rate	Low	± 5 psi / ± 1 L/min
Ion source configuration (Nebulizer position)	High (if configurations can vary) Not applicable (if fixed source)	According to the ion 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

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 the 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 even 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 the sample matrix variability (in the case of formally identical matrices) more broadly than just for the 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

Matrix robustness

Robustness

- Robustness – influence of variations of method parameters on results

- Also robustness in terms of variability of sample matrix
  - different matrices can lead to different matrix effects

Matrix robustness

- Robustness of the sample matrix is beneficial to consider and it is benefitial to distinguish it from the selectivity

- Influences:
  - selectivity
  - recovery (trueness)
  - ionization suppression/enhancement
  - limit of detection
  - limit of quantification

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 what the influence of these parameters. On the basis of this information it is possible to plan changes to the method. 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 the sample preparation and for the LC-MS analytical part.

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

For example, if the method's LoQ is very close to the LoQ required by legislation, then the changes in the LoQ value have to be monitored against small changes in the method parameters. The most influential method parameters impacting the 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 the method robustness and ruggedness – to use the experimental design or the One Factor 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

Example of the One Factor at the time approach

An Example of how an OFAT can be used is described in the table below. O represents here the optimal value, + and – are values larger and smaller than the optimal value. All three are chosen by the chemist in charge of the method development. For robustness study we change, as the name suggest, one factor at a time and monitor the response variable – the retention time – in our example. The actual experiments need to be performed in random order to prevent a bias on the response variable. After finishing the OFAT the responses can be used to calculate e.g. % deviation from the "optimal response" (experiment 7) and thereby deduct if the factor change has a significant influence on the response.

Experiments	Actual experiment order	A pH	B Additive conc.	C Column Temp.	Response Retention time
1	3	O	O	+	7.95
2	6	O	O	-	8.13
3	5	O	+	O	8.12
4	1	O	-	O	7.72
5	4	+	O	O	8.32
6	2	-	O	O	9.82
7	7	O	O	O	8.03
*O Optimal factor setting					

One Factor at a Time
https://www.uttv.ee/naita?id=32134

https://youtu.be/kBR703IK08A

In most cases experiments with one-by-one variations (One Factor 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. Furthermore, programming skills in R or Python are beneficial if the chemist is not willing to buy a software tool for the evaluation of the experimental deigns. In this course we will give an overview of both – One Factor At a Time approach and the Experimental Design approach.

Based on the common practice and for the sake of simplicity, we recommend the following:

1. Change parameters one by one (One Factor 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 the mobile phase flow rate can also lead to a 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

Different ways to evaluate robustness

Choosing the parameters

- LC parameters
  - pH
  - Concentration of additives in eluent
  - Organic solvent content in the eluent
  - Column temperature
  - Eluent flow rate
  - Column batch and age
- Sample and sample preparation parameters
  - Analyte extraction time
  - Injection solvent composition
  - Matrix effect
- Mass spectrometry parameters
  - Drying gas temp
  - Nebulizer gas pressure/flow rate
  - Ion source configuration and condition

Choosing the parameters

- How much a given method parameter can influence the critical characteristic

and

- how likely it is that this parameter will change uncontrollably

One Variable At a Time approach vs Experimental design

- One variable at the time approach
  - Simple
  - Time-consuming
  - Needs a knowledge of critical performance characteristics of the method
- Experimental design
  - Powerful tool
  - Needs a knowledge of critical performance characteristics of the method
  - Requires a knowledge and experience with mathematical statistics

One Variable At a Time approach

1) Change parameters one by one 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

3) Effects from the change of parameters should be monitored (graphical or statistical analysis)

4) Regarding the robustness tests results, if necessary, measures to improve the performance of the method should be taken

10.4 Experimental design

Additionally to the One factor at a Time approach (OFAT, explained in 10.3), robustness can also be estimated by varying multiple method parameters (factors) at the same time. Such approaches are strongly recommended if it is expected that different method parameters have *interactions*: meaning that the value of one parameter influences another value of a different parameter. In chromatography, the parameter interactions between some of the parameters are rather the rule than the exception [ref 65]. Therefore, using OFAT can lead to getting only a partial and/or flawed picture of the method robustness. For example, the retention behaviour of the acidic compounds in RP is influenced by the mobile phase pH. Additionally, changing the additive concentration can at the same time strongly influence pH and, in turn, changing pH can influence the effect of the additive. These interactions will remain unnoticed with One factor at a Time approaches.

A statistical approach to validation procedures is the Design of Experiment (DoE). This tool has been broadly applied for many years in method optimization but is also relevant for the method validation process. Even though a DoE is not yet prominently illustrated in the validation guidelines for analytical method development, publications in recent years have shown [ref 66 and ref 67] that it is on its way to becoming a valuable mainstream tool for evaluating robustness. Additionally, a DoE chapter has been added to the 10th version of the European Pharmacopoeia (Ph. Eur.) and will certainly influence the future of validation guidelines.

How to perform a DoE

Design of Experiment
https://www.uttv.ee/naita?id=32135

https://youtu.be/D71f9uFqHDA

The most basic way to study the robustness of a method is called *full factorial* design, carried out as in the following example for a robustness evaluation:

a) Brainstorm and list all factors which are likely to influence the robustness of a newly developed chromatographic method. The DoE approach is needed to determine if the factors like the eluent pH (let's call it factor A), the additive concentration of the eluent (B) and the column temperature (C) have a critical impact on the retention time (also called response variable). The three factors (A, B and C) are set on two coded levels, a high (+) and a low level (-), which are depicted in the table below. We are using one level above and one below our optimal factor set to evaluate the robustness in both directions. The factor levels themselves are chosen by the chemist in charge of the validation and should be placed in a reasonable distance from the optimal factor setting. In practice, this means considering errors in laboratory work that appear while creating e.g. a new eluent solution and then predicting based on these thoughts ΔpH or ΔAdditive conc. values, which can be used as upper and lower levels in DoE.
To calculate the number of experiments at different factor combinations, a simple formula can be used:

$$n = 2^k \quad \textbf{(Eq 1)}$$

where n – number of experiments performed at different combinations
k – number of factors.

In the current example k = 3: pH, additive concentration, column temperature. Thus, we need to perform 8 experiments in total to determine the robustness of the response variables at two levels.

b) Construct a measurement plan that would contain all possible combinations of the factor levels for one substance in the mixture.

Note 1: We are monitoring the retention times of all analyte substances, so by performing the experiments later, we can create tables like the following ones for all substances which are planned to be separated in the mixture and calculate the specific effects for every substance individually.

Note 2: If repeatability is a concern to you then it is recommended to use a central point in DoE with only "optimal" setting and repeating this point multiple times at random in the experimental plan instead of repeating the whole design itself.

In a step-by-step manner:

First, notice that in the table, we first add + signs to half of the experiments of the first factor column (pH).

Experiment number	A pH	B Additive conc. [mmol/L]	C Column Temp. [°C]
1	+		
2	+		
3	+		
4	+		
5	-		
6	-		
7	-		
8	-		

Then we fill in the next factor by splitting the first factors + signs in half.

Experiment number	A pH	B Additive conc. [mmol/L]	C Column Temp. [°C]
1	+	+	
2	+	+	
3	+	-	
4	+	-	
5	-	+	
6	-	+	
7	-	-	
8	-	-	

And we do the same for the third (all other columns that remain).

Experiment number	A pH	B Additive conc. [mmol/L]	C Column Temp. [°C]
1	+	+	+
2	+	+	-
3	+	-	+
4	+	-	-
5	-	+	+
6	-	+	-
7	-	-	+
8	-	-	-

c) Behind all the + and – signs are factor levels that are used to carry out the experiments. The decoded factors for this Full Factorial Design can be seen in the table below. Practically this means in our example to perform 8 LC runs with 4 different eluent solutions and varying temperatures in the column compartment according to the decoded values in the table below. All experiments need to be carried out in a random order to prevent a bias on the response variable. The reason for this is that unknown factors might be aligned with each other if the experiments are not performed in random order. From those 8 runs we collect all the retention times of all substances in the mixture in order to determine the robustness individually for every analyte. It is always possible to use more response variables than just the one (retention time) depicted here in this example.

Now in this step, we perform all the experiments and transfer the results, in our case, the retention time, to the table.

Experiment number	Actual order of measurement	A pH	B Additive conc. [mmol/L]	C Column Temp. [°C]	Response 1 Ret. Time [min]
1	3	9.8	5.2	31	8.31
2	4	9.8	5.2	29	8.10
3	2	9.8	4.8	31	7.24
4	5	9.8	4.8	29	7.43
5	6	9.4	5.2	31	8.32
6	7	9.4	5.2	29	8.92
7	1	9.4	4.8	31	9.84
8	8	9.4	4.8	29	9.93

d) Now it is time to calculate the actual effects of the three factors on the response variables. This is done by finding the average for the response at a high value ("+") and subtracting the average of the response at the low value ("-"). In case of a retention time, we would have:

$$Factor\ A = \frac{R_1^{A+} + R_2^{A+} + R_3^{A+} + R_4^{A+}}{4} - \frac{R_5^{A-} + R_6^{A-} + R_7^{A-} + R_8^{A-}}{4} \quad \textbf{(Eq 2)}$$

Effect of the factor A on the response would be calculated as follows:

$$Factor\ A = \frac{8.31 + 8.10 + 7.24 + 7.43}{4} - \frac{8.32 + 8.92 + 9.84 + 9.93}{4} = -1.48 \quad \textbf{(Eq 3)}$$

This is now the effect that the change in pH has on the response variable retention time. This calculation can be performed in a similar manner for all factors A, B, C, and all interactions.

Results retention time [min]		
A	B	C
pH	Additive conc.	Column Temp
-1.48	-0.20	-0.17

From this depiction, it is clear that the pH value especially has the highest influence on the retention time, in the way that a decrease in pH increases the retention time of the analyte, which would imply that the method is not robust regarding a pH change of ± 0.2 pH units.

This can also be visualized in the following plot. By the slope, it is easily deductible that the impact of the pH on the response variable retention time is greater than the effect of the additive concentration in the eluent. In general, the steeper the slope is, the more impact a factor has on the response variable. The effect size describes the numeric influence of the factor on the response variable in the set design region and can be used to distinguish the importance of effects from the less important ones.

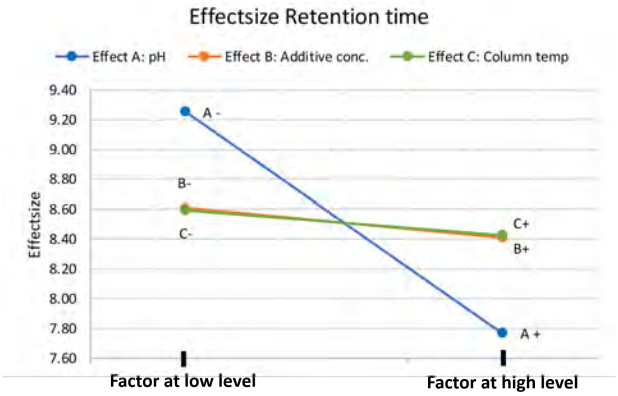


Figure 1 Main effect plot of pH, Additive conc. and Column temp. on the Retention time

The influence of the additive concentration in our chosen design region is less impactful on the retention time than pH change in our chosen design region on the retention time. Thereby, we might assume that the method is robust towards changes in additive concentration. But, as will be shown in the following section, this assumption would be wrong! In that's section we will show the importance and benefits of also monitoring factor interactions instead of only the individual effects of the factors.

e) Interactions are important

One of the major advantages of a DoE approach is the determination of the effects of interactions on the response variable. To calculate these interactions, we proceed the same way as with the normal factors A, B and C. The coded value for an interaction is deducted by the multiplication of the single effects meaning that AB= A*B which might look like this
+ * + = + or this + * - = -.

So, for the interaction of AB (pH and additive conc.), we get coded values like this:

If we do the multiplication for all the interaction combinations, it creates the following table with the interactions.

Experiment	AB	BC	AC	ABC	Ret. Time [min]
1	**+**	**+**	**+**	**+**	8.31
2	**+**	**–**	**–**	**–**	8.10
3	**–**	**–**	**+**	**–**	7.24
4	**–**	**+**	**–**	**+**	7.43
5	**–**	**+**	**–**	**–**	8.32
6	**–**	**–**	**+**	**+**	8.92
7	+	**+**	**–**	**+**	9.84
8	+	**–**	**+**	**–**	9.93

Now the effects of the interactions can be calculated with the same formula from above.

We can calculate the interaction effect the same way we calculated the single factor effect:

$$Interaction\ AB = \frac{\sum R_1^{AB+} + R_2^{AB+} + R_7^{AB+} + R_8^{AB+}}{4} - \frac{\sum R_3^{AB-} + R_4^{AB-} + R_5^{AB-} + R_6^{AB-}}{4} \quad \textbf{(Eq 4)}$$

$$Interaction\ AB = \frac{\sum 8.31 + 8.10 + 9.84 + 9.93}{4} - \frac{\sum 7.24 + 7.43 + 8.32 + 8.92}{4} = 1.07 \quad \textbf{(Eq 5)}$$

Now we have the complete picture of all effects on the response variable.

Results (factor and interaction effect sizes) For retention time						
A	**B**	**C**	**AB**	**BC**	**AC**	**ABC**
pH	Additive conc.	Column Temp.	pH * Additive conc.	Additive * Column Temp.	pH * Column Temp.	pH * Additive * Column Temp.
-1.48	-0.20	-0.17	**1.07**	-0.07	0.18	0.23

The two-factor interaction between a pH and an additive concentration has a big influence on the retention time. This implies that our method is not yet robust towards pH and additionally towards additive concentration. In our case, the conclusion comes from the comparison of numeric values of the effect size with each other. There is strong evidence towards the significance of pH and pH*Additive conc. However, easy identification of significant effects is not always given, and in these cases, it is strongly recommended to use an ANOVA approach to determine which effects and factors are statistically significant. A good ANOVA description in the context of a DoE can be found in the book [ref 68]. However, all results from a DoE should always be evaluated from a practical scientific standpoint.

In general, it is assumed that the importance of interactions will decrease with additional interacting factors, so it is unlikely that there will be very significant (three), four or five-factor interactions.
Reaching this conclusion is only possible with the help of a DoE. If the user would use the OFAT approach, the effects of the interactions on the response variables would go completely unnoticed. Only with the application of a DoE we were able to see the impact of the two-factor interaction on the response variable.

Another way of portraying interactions is with the following plots. In both plots one factor is kept constant while the other factor is compared on its influence on the response variable. The plots firstly show us that there is an interaction between factor A and B because the slope changes visibly when only one factor is changed. Secondly, we can say that the interaction of AB is more influenced by the -1 setting for B than by the +1 setting of B.

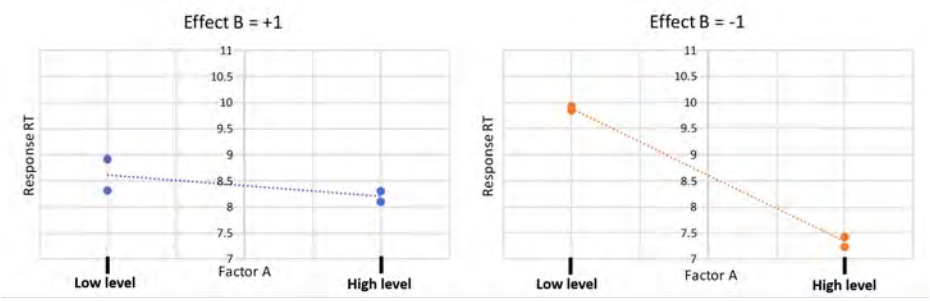


Figure 2 Interaction plots of the factors (A) pH and (B) Additive conc.

f) contour plots

A different way for the interpretation of the interactions on the response variable are the so-called contour plots, as seen in the figures below. These pots have been calculated and plotted in the "R" software with the packages pid and rsm. The lines in the plot depict the interaction between pH and additive concentration and their effect on the response variable. The black dots in the corners of the contour plot describe the design space where the measurement points have been set. The more the lines are bent, the stronger is the interaction. The examples make this evident. When looking at the figure of pH and additive concentration, we see strongly bent lines in the chosen design space, and therefore we can assume a strong interaction between the two factors. On the contrary, if we look at the interaction between pH and column temperature, then we only see a very weak interaction between those factors.

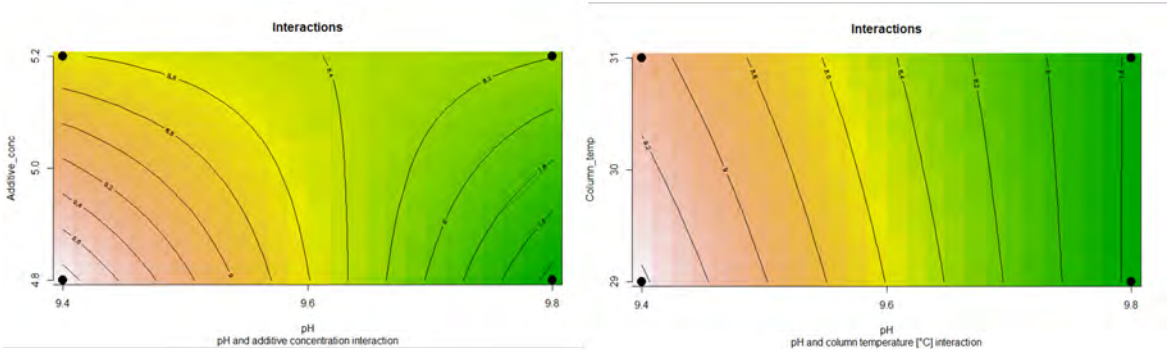
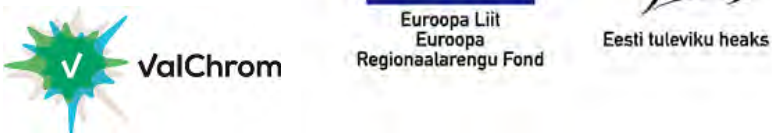


Figure 3 Contour plots of the interactions

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.

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).



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.

This course is also part of Estonian Center of Analytical Chemistry trainings.

Frequently asked questions

Questions and answers about precision, trueness and accuracy topics.

Question: Why does the replicate measurement in case of precision assume that the replicates are "appropriately independent" but in case of linearity replicates can also mean that samples are run for example in triplicate (from the same vial)?

Answer: The way you have to plan your replicates depends on the purpose that you have. Carrying out sample preparation once and thereafter injecting multiple times is a valid activity, but it does not yield the precision of the method but only the precision of the LC-MS analysis part s_r_LCMS. And importantly: s_r_LCMS is smaller (often significantly smaller) than s_r.

In order to obtain the precision estimate of the method the replicate measurements have to be carried out in such a way that all method's steps are included.

When determining linearity – carrying out triplicate injections of calibration solutions serves the purpose of reducing random scatter of data points caused by the LC-MS system on the graph. This approach does not enable reducing scatter caused by preparation of the calibration solutions. However, the scatter caused by the LC-MS is usually significantly larger than scatter caused by preparation of solutions.

Question: How to calculate representative bias when you have many bias values found in the validation eg for different fruits and vegetables or concentration levels?

Answer: The answer to this question to some extent depends on how the obtained bias data will be used. Here are some situations together with general recommendations. --- If bias is needed – either for reporting, correcting or for taking into account as uncertainty component – for a specific fruit or vegetable, use the bias for the specific fruit or vegetable. --- If bias estimate is needed for demonstrating that method's bias is sufficiently low for all matrices, then it is suggested to present the worst bias, if the worst bias is still tolerable. If it is not, then a closer look should be taken at the data. --- If bias estimate is needed for evaluating general measurement uncertainty of the analysis method (using the Nordtest approach) then it is suggested to use the root-mean-square bias. A rather good general guidance material is a recent article about bias in journal *Bioanalysis* (ref 22 in the course). Although its title specifically refers to clinical chemistry, it is actually very relevant for any chemical analysis.

Question: Section 6.3. says "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." However, one criterion has to be fulfilled in applying correction. "Useful reduction of combined uncertainty is achieved (as compared to inclusion of the possible bias into uncertainty)." It seems that even though correcting the bias for the final result, the uncertainty of the bias (before applying correction) still has to be taken into uncertainty budget. If so, the final combined uncertainty should not be reduced, and should be the same no matter the correction is applied or not. Could you please help clarify?

Answer: If there is (possible) bias in the result then there can be two situations, depending on, whether the result is corrected for bias or not.
--1-- If one corrects the result with bias then the uncertainty of the correction has to be included in the uncertainty of the result. That uncertainty can sometimes be large and if we speak about possible bias (i.e. if it is not sure whether there actually is any bias at all) then the uncertainty of the correction can be larger than the correction itself.
--2-- If one does not correct for bias then one should include the uncertainty due to possible bias into the uncertainty of the result. In both cases the (possible) existence of bias introduces additional uncertainty, whether we correct or not. The point now is that the uncertainty increase when correcting for bias should be smaller than in the case when the result is not corrected. Let us look at two examples.
--- Example 1 --- Suppose we have determined the content of a contaminant in a sample C = 120 µg/kg. Let us assume that the joint uncertainty contribution of all other (i.e. except bias) uncertainty sources is 10 µg/kg. Let us assume that we have determined bias and found that it is -20 µg/kg with standard uncertainty 6 µg/kg. In this case, if one would correct for bias one would end up with the result 140 µg/kg and combined standard uncertainty 12 µg/kg. If bias correction would not be carried out then the result would be 120 µg/kg and combined standard uncertainty would be 23 µg/kg (the uncorrected bias as well as its standard uncertainty have been included as standard uncertainty contributions). Obviously in this case bias correction is justified – uncertainty decreases by roughly two times as a result of correction.
--- Example 2 --- Suppose we have determined the content of a contaminant in a sample C = 120 µg/kg. Let us assume that the joint uncertainty contribution of all other (i.e. except bias) uncertainty sources is 20 µg/kg. Let us assume that we have determined bias and found that it is -10 µg/kg with standard uncertainty 9 µg/kg (this means that it is actually a possible bias – we do not know for sure if there is any bias at all). In this case, if one would correct for bias one would end up with the result 130 µg/kg and combined standard uncertainty 22 µg/kg. If bias correction would not be carried out then the result would be 120 µg/kg and combined standard uncertainty would be 24 µg/kg. In this case bias correction is not justified, because the uncertainty remains almost the same and we do not know whether there actually is any bias at all.

Question: Is it possible to use one sample t test instead of zeta score to evaluate whether two results are in agreement?

Answer: It is more correct to use the zeta score, because of two reasons:
--1-- The zeta score takes into account the combined standard uncertainty instead of just scatter of results. This way it accounts for both random and systematic effects.
--2-- The zeta score takes into account also the uncertainty of the reference value. As a result of these things it is possible (and not rare) that the reference value is not within the range of the individual measurement values (especially if the individual values were obtained under repeatability conditions) but at the same time the average measured value and reference value can still be considered to be in agreement.

Question: What is the difference between "process efficiency" and "recovery"?

Answer: The term "process efficiency" (*PE*) as used in this course refers to the joint effect of possible losses during sample preparation and ionization suppression/enhancement ($ME_{ionization}$) in the ion source (see section 5.1). The term "recovery" (*R*) only refers to possible losses during sample preparation. In most other techniques, this distinction is not made and if recovery determination is carried out then what is actually determined is process efficiency, in the sense that also losses after sample preparation are accounted for. This is OK, if one can assume that the losses during sample preparation are significantly higher than losses during subsequent steps, as is the case in most techniques. In LC-MS this is not so: ionization suppression/enhancement can be a very serious effect. Furthermore, depending on, whether signal loss (usually it is loss, rarely enhancement) occurs via analyte loss during sample preparation or because of ionization suppression/enhancement in the ion source, different approaches are needed for improving the situation. For this reason it is useful to distinguish between *R*, $ME_{ionization}$ and *PE* and quantify them separately.

Question: What kind of samples are best for control charts?

There are different types of control charts but X charts are addressed here only, as these are the most widespread (and useful).

In the case of X charts different quantities can be plotted on the control chart, depending on what information you want to obtain from the control chart. In most cases people use control charts for keeping the whole analytical method under control. In such a case plotting the result in the control chart is the right thing to do.

Furthermore – it is very strongly recommended not to use analyte in solvent as control sample. Please instead use real sample as control sample and carry out sample preparation every time you run the control sample. Only in such case will the control chart embrace all steps in your method. If you use analyte in solvent then any information about sample preparation (as well as possible selectivity and ionization suppression issues that occur with real samples) will be lost.

A very good source of information on control charts is the Nordtest Trollbook.

Question: How to understand performance characteristics specified eg in Drinking Water Directive (DWD)?

Directive specifies performance characteristics for uncertainty of measurement and limit of quantitation (LoQ). Note, until 31 December 2019 limit of detection (LoD), Trueness and Precision are also allowed to use as performance characteristics. The directive should be read in such a way:

Look for the parametric values in Annex I for a specific compound eg for pesticides 0.10 µg/l (commonly analysed with LC-MS).

Annex III Part B states that limit of quantitation (LoQ) should be 30% or less of the relevant parametric value. Meaning that LoQ for pesticide analysis should be 0.025 µg/l or less.

In Annex III it can be seen that uncertainty of measurement has to be 30% of the parametric value for pesticides with a note that values as high as 80% may be allowed for a number of pesticides.

In the directive the performance criterion for measurement uncertainty (k = 2) is the percentage of the parametric value stated in the table or better. Measurement uncertainty shall be estimated at the level of the parametric value, unless otherwise specified.

However, until 31. December 2019 also the following is allowed: minimum performance characteristics for trueness, precision and limit of detection can be found in Annex III, which for pesticides is 25% for each parameter meaning that in method validation for pesticides, for each compound, trueness, precision and LoD should be equal or better than 0.025 µg/l.

Question: If the blank samples are not available for the type of matrices you are working with then how can one check for the selectivity of the method?

Absence of blank matrix is relatively common situation, but there is no single solution. Some steps to take.
- Retention time and peak shapes of analyte in sample and standard solution must match. (Comparing unspiked and spiked sample peaks may be helpful.)
- If isotopically labelled standard substance is used, it must also have equal retention time and identical peak shape.
- LC conditions (e.g. different column, eluent pH, gradient profile) should be altered and all the above must still be true.
- In case of MS-detection all the recorded transitions should yield peaks of similar shape and retention time.
- Fragmentation spectra of analyte peak in sample and in standard solution must match.
- ...

Question: What is the benefit of identification points and how they can be calculated?

Answer: Identification points are related to identity confirmation and are used in the EU guidelines 2021/808 [ref 5]. Let's imagine that we register a chromatogram of a urine sample and find a peak at 5 min. From calibration experiments we know that retention time of paracetamol standard substance is also 5 min under the same conditions. Could the peak in sample be due to paracetamol? Yes, but we can not be very sure, because we only used low resolution (LR) MS for m/z 152, but there can be several compounds eluting at 5 min and yielding ions of m/z 152. (1 identification point)
For higher confidence, we could use more sophisticated MS techniques, for example MRM mode to isolate precursor of 152 and fragmentation product of m/z 110. There is relatively limited number of compounds eluting at 5 min, ionizing to m/z 152 and fragmenting to 110 m/z. There is more confidence that the peak is due to paracetamol. (1 + 1.5 = 2.5 identification points)
So, the identification points reflect how much proof we have to claim that the peak is due to specific compound. For calculation examples, please see Example 2 in chapter 2.7 (https://sisu.ut.ee/lcms_method_validation/27-identity-confirmation-examples)

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

The European Commission: Guidance document on analytical quality control and method validation procedures for pesticide residues and analysis in food and feed. Also known as SANTE (formerly SANCO). The link leads to a page where the most recent version, as well as the older versions can be found.

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.
Note that this directive is superseded by Commission Implementing Regulation (EU) 2021/808 of 22 March 2021 on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling.

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

ICH guideline M10 on bioanalytical method validation and study sample analysis. European Medicines Agency 2022. Old version of this reference was: Guidance on bioanalytical method validation, European Medicines Agency, 2011.

ref 8

U.S. Department of Health and Human Services Food and Drug Administration, Bioanalytical Method Validation, Guidance for Industry, 2018.

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 of analytical procedures: text and methodology Q2(R1), International Conference of harmonization of technical requirements for registration of pharmaceuticals for human use **2005**. A revised/renewed draft version of this guideline: ICH guideline Q2(R2) on validation of analytical procedures **2022.**

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 Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements. *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" *Clin Biochem* 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

Laaniste, A., Leito, I. & Kruve, A. ESI outcompetes other ion sources in LC/MS trace analysis. *Anal Bioanal Chem* 411, 3533–3542 (2019)

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.

ref 58

M. Hecht, R. Veigure, L. Couchman, C. I. S. Barker, J. F. Standing, K. Takkis, H. Evard, A. Johnston, K. Herodes, I. Leito, K. Kipper, Utilization of data below the analytical limit of quantitation in pharmacokinetic analysis and modeling: promoting interdisciplinary debate, *Bioanalysis* **2008,** 10(15), 1229-1248

ref 59

S. Kaupmees, R. Rebane. Analysis and validation of perfluorinated compounds in water, sediment and fish with LC- ESI–MS/MS. *International Journal of Environmental Analytical Chemistry* **2017**, 97(8), 695–709.

ref 60

R. Rebane, K. Herodes. Matrix interference in LC-ESI-MS/MS analysis of metanephrines in protein precipitated plasma samples. *European Journal of Mass Spectrometry* **2019**, 26(1), 46-54, https://doi.org/10.1177/1469066719862423.

ref 61

K. Kipper, M. Hecht, N.J. Antunes, L.D. Fairbanks, M. Levene, S. Kalkan Uçar, A. Schaefer, E.L. Blakely, B.E. Bax. Quantification of Plasma and Urine Thymidine and 2'-Deoxyuridine by LC-MS/MS for the Pharmacodynamic Evaluation of Erythrocyte Encapsulated Thymidine Phosphorylase in Patients with Mitochondrial Neurogastrointestinal Encephalomyopathy. *J. Clin. Med.* **2020**, 9, 788.

ref 62

K. Takkis, R. Veigure, T. Metsvaht, M. Hallik, M.-L. Ilmoja, J. Starkopf, K. Kipper. A sensitive method for the simultaneous UHPLC-MS/MS analysis of milrinone and dobutamine in blood plasma using NH_4F as the eluent additive and ascorbic acid as a stabilizer. *Clinical Mass Spectrometry* **2019**, 12, 23-29. https://doi.org/10.1016/j.clinms.2019.03.003

ref 63

K. Kipper, C.I.S. Barker, J.F. Standing, M. Sharland, A. Johnston. Development of a Novel Multipenicillin Assay and Assessment of the Impact of Analyte Degradation: Lessons for Scavenged Sampling in Antimicrobial Pharmacokinetic Study Design. *Antimicrobial Agents and Chemotherapy* **2018**, 62 (1), 1-11. **DOI:** 10.1128/AAC.01540-17

ref 64

R. Veigure, R. Aro, T. Metsvaht, J.F. Standing, I. Lutsar, K. Herodes, K. Kipper, on behalf of the CloSed Consortium. A highly sensitive method for the simultaneous UHPLC–MS/MS analysis of clonidine, morphine, midazolam and their metabolites in blood plasma using HFIP as the eluent additive. *Journal of Chromatography B* **2017**, 1052, 150-157. https://doi.org/10.1016/j.jchromb.2017.03.007

ref 65

S. B. Ganorkar, A.A. Shirkhedkar. Design of experiments in liquid chromatography (HPLC) analysis of pharmaceuticals: Analytics, applications, implications and future prospects. *Reviews in Analytical Chemistry*, **2017**, 36(3).

ref 66

V. Murthy, M. Naresh, K. Katari. Development and Validation of RP - UPLC Method for 2 , 6 - Dimethylaniline, Its Isomers, and Related Compounds Using Design of Experiments. *Chromatographia*, **2021**, 84(4), 359–369.

ref 67

M. R. Patil, S. B. Ganorkar, A.S. Patil, A. A., Shirkhedkar, S.J. Surana. A converged pharmaceutical analysis supported with hydrotropy & DoE with dual HPTLC and stress studies for estimation of tolvaptan. *Microchemical Journal*, **2021**, 167(March), 106328.

ref 68

Process Optimization: A Statistical Approach, By Enrique Del Castillo 2007, Volume 105 ISBN : 978-0-387-71434-9

Glossary

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.

AOAC

Association of Official Analytical Chemists

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.

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.

FDA

The US Food and Drug Administration

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).

IS

Internal standard.

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.

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).

MRL

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

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?

" I've come across many of the method validation guidelines..... however this course has provided a very nice comparative study of all these guidelines....Thanks to all who have designed and taught the course....Thank you teachers....moreover thanks to all those who have started the discussions on various topics.... It was really a wonderful experience.... and hope we'll have more such courses in future....."

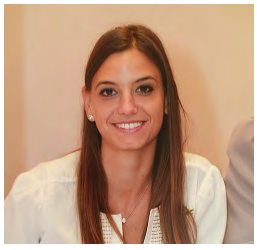
Devottam Banerjee
India
February 2023

"That was so excited and insightful course. I learned such a wonderful experience. It engaged theoritical studies with practical performance. That was a challenge for me to finish that course. I applied before for that course but failed to pass it due to the online study. But now I passed it with excellent grade as my 49 online course. Great professors and great material combined together. Thank you very much for everything."

Yahia Rashad Mohamed Mohamed Abdelhalim
Egypt
February 2023

"The course was amazing and the teachers have performed a great job in assisting us. I hope that the team will continue to make this program available for all the students/knowledge-seekers around the globe. It was a great learning experience."

Jenny Love D. Magbanua
Philippines
April 2022



"At the beginning it was a little hard for me to follow the course. There were many new things, in general, this was the first MOOC I attended and the first one that was completely on moodle platform and online. But later, when I got used to it, it was great and easy to follow. I learned a lot, not just about method validation, but also about how courses like this function, and I think this was a great experience at all.

The only better thing than course materials are teachers. I am very happy that I had the opportunity to be a part of this amazing course and to learn many new things. I would like to thank all the teachers for their help, kindness and generosity. I wish you all the best!"

Bojana Blagojevic
Serbia
April 2022

"I really liked this course. Thank you for this course. It is very organised, structured with real-life situations. I felt always motivated and interested in each parameter and I knew that the teachers are always available for any doubt. Thank you!"

Carlos André Araújo Duarte Monteiro
Portugal
April 2022

Dejan Orcic
Serbia
April 2022

"The course provided a well-rounded, systematic review of the topic, with just right amount of math to perform the necessary calculations. Even though I have over 15 years of experience with method development (and some validation), I have still learned a lot and immediately applied the gained knowledge.

The course was well-structured, providing a systematic overview of each of the validation parameters, with right amount of details (unlike practically every validation guideline I have ever read)."

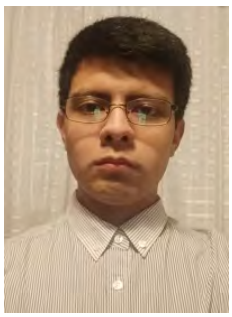
"I would like to express my sincere thanks for the opportunity to participate in a really fantastic course. The course organization was engaging, with outstanding teachers and materials, and I was able to learn a lot about this important topic. In addition, the pedagogical approach enabled me to complete the course successfully.
Thank you again for this wonderful opportunity.
I will enjoy participating in future classes if they are offered."

Luz Adriana Diaz Cano
Colombia
Feb 2022

Mina Gergis
United States of America
Feb 2021

"Thank you very much for the amount of effort and time that you and your team put in to put together a top notch course. I truly learned a lot of new and hands on information that I already started implementing at my work."

"This is an excellent course that has given me a deeper understanding of how to carry out an LC-MS method validation properly. It has great explanations of all the parameters involved in the method validation process, video lectures, and exercises about them and, the most important from my point of view, it gives you real-life examples that show the applicability of all the concepts, equations, and definitions learned through the 11 thematic sections.
Furthermore, the course is so well-thought that not only provides you with the tools to be able to solve numerical problems but also to learn how to interpret and analyze them and make decisions based on your results."



Aldo Rayas Reyes
Mexico

Feb 2021

"Hello, it was a great pleasure to be part of this course. I would like to thank the course team for sharing your knowledge and organizing all the topics for us in a really didactic way."

Adriane Costa dos Santos
Brazil
Feb 2020

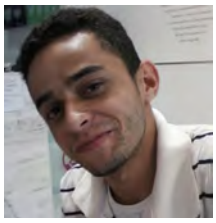
"I am very thankful to the course team for giving me the opportunity to learn something new and understand better what I already have learnt. This course was very well organised and useful for me as a recent graduated Chemist from Serbia, and I hope I will attend at some point new course at University of Tartu."

Nevena Uletilovic
Serbia
Feb 2020

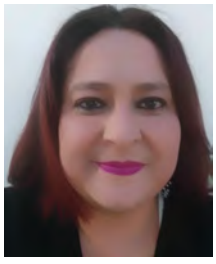
"Obviously proven working experience is the thing which employers are looking for but unfortunately actual hands-on lab courses are scarce, practically non-existent and most likely very expensive, if available. Informative videos in addition to written material and tests depicting real-life problems makes this course the closest approximation of a real work I have found so far."

Jari Juhani Aho
Finland
Feb 2020

"The LC-MS method validation course presented a complete overview of the validation of methods and in an organized way. The videos and materials are of quality and teach complex issues in an uncomplicated and interesting way. I really enjoyed the overview that was given on various validation guides. Congratulations to the University of Tartu."



Jaime dos Santos Viana
Brazil
June 2019



"First of all, I consider that this course contains a lot of information to help to solve situations when carrying out the validation of a method using LC-MS. I believe that the topics explained by the professors were clear and give us concrete examples to better understanding of how to execute or analyze the results obtained."

Monica Montańez Romero
Colombia
June 2019

"I really liked the course - it had very good organization. The lecturers presented the lecture material very well and the exercises were very well chosen."

Desislav Tomov
Bulgaria
June 2019

"Thank a lot for all your support and helping us to successfully understand and pass the course. It is a great educational course!"

Walleed Saber Abdel Aziz Ibrahim
Egypt
February 2018

"I like the discussion forums the most. I can virtually clear my doubts anytime because professors/instructors are always available to answer my questions. The immense effort the Tartu team has put together to conduct the validation course is really commendable."

Amresh Karmacharya
Nepal
February 2018

"I'm glad that I could succesfully pass this really good organized course! I'm looking forward for any other news about courses to excelence my analytical skills."

Joanna Maria Langner
Poland
February 2018

"The structure of the course is excellent and addressed all validation topics."

Isaque Wilkson De Sousa Brandão
Brazil
February 2018

"Thanks you very much for let me participate of this wonderfull course! It is very interesting, clear and complete one."

Laura Flores
Uruguay
February 2018

"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