DISSERTATIONES
CHIMICAE
UNIVERSITATIS
TARTUENSIS
154

HANNO EVARD

Estimating limit of detection for mass spectrometric analysis methods

DISSERTATIONES CHIMICAE UNIVERSITATIS TARTUENSIS
154

DISSERTATIONES CHIMICAE UNIVERSITATIS TARTUENSIS
154

HANNO EVARD

Estimating limit of detection for mass spectrometric analysis methods

UNIVERSITY OF TARTU
Press

Institute of Chemistry, Faculty of Science and Technology, University of Tartu, Estonia.

Dissertation was accepted for the commencement of the degree of *Doctor philosophiae* in Chemistry at the University of Tartu on May 27th, 2014 by the Council of Institute of Chemistry, Faculty of Science and Technology, University of Tartu.

Supervisor: Anneli Kruve (PhD), Institute of Chemistry, University of Tartu, Estonia

Prof. Ivo Leito (PhD), Institute of Chemistry, University of Tartu, Estonia

Opponent: Ms Emilia Vasileva-Veleva (PhD), Environment Laboratories, International Atomic Energy Agency, Monaco

Commencement: August 31, 2016 at 14:00, Ravila 14a, room 1021

This work has been partially supported by Graduate School Functional materials and technologies receiving funding from the European Reginal Development Fund under project in University of Tartu, Estonia

ISSN 1406-0299
ISBN 978-9949-77-137-0 (print)
ISBN 978-9949-77-138-7 (pdf)

University of Tartu Press
www.tyk.ee

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I. H. Evard, A. Kruve, R. Lõhmus, I. Leito, Paper spray ionization mass spectrometry: Study of a method for fast-screening analysis of pesticides in fruits and vegetables, J. Food Compos. Anal. 41 (2015) 221–225.

II. 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. 870 (2015) 29–44.

III. 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. 870 (2015) 8–28.

IV. H. Evard, A. Kruve, I. Leito, Tutorial on Estimating Limit of Detection on the example of LC-MS analysis: Part I, (accepted to Analytica Chimica Acta).

V. H. Evard, A. Kruve, I. Leito, Tutorial on Estimating Limit of Detection on the example of LC-MS analysis: Part II, (accepted to Analytica Chimica Acta).

Author's contribution

I. Main person responsible for planning and performing the experimental work and writing the manuscript.
II. Main person responsible for writing the respective parts of the manuscript on the topic of detection limit; critical review of the whole manuscript.
III. Main person responsible for writing the respective parts of the manuscript on the topic of detection limit; critical review of the whole manuscript.
IV. Main person responsible for planning and performing the experimental work and writing the manuscript.
V. Main person responsible for planning and performing the experimental work and writing the manuscript.

ABBREVIATIONS

3R	Electrospray ionization source with additional nebulization capillary
ANOVA	Analysis of variance
APCI	Atmospheric pressure chemical ionization
API	Atmospheric pressure interface
APPI	Atmospheric pressure photoionization
CC_{α}	Decision limit
CC_{β}	Detection capability
CF	Calibration function
CI	Chemical ionization
ESI	Electrospray ionization
IS	Isotopically labelled internal standard
LC-MS	Liquid chromatography mass spectrometry
LoD	Limit of detection
LoQ	Limit of Quantitation
MALDI	Matrix assisted laser desorption ionization
MRL	Maximum residue limit
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
OLS	Ordinary least-squares regression
PSI	Paper spray ionization
RSD	Relative standard deviation
S/N	Signal-to-noise ratio
$S_{y.x}$	Standard deviation of residuals
WLS	Weighted least-squares regression

1. INTRODUCTION

The goal of this doctorial theses is to provide a comprehensive guide on determining limit of detection (LoD) for analysis methods using mass spectrometry (MS), clarify important aspects of LoD determination and to give suggestions for practitioners on this topic.

A very large number of articles reporting new MS methods, especially liquid chromatography mass spectrometry (LC-MS) methods, are submitted for publication each year and method validation is an intrinsic part of such articles. However, in spite of the number of articles published, miscalculation and misinterpretation of validation parameters is still common due to complexity of the methods and some ambiguity in the definitions of some of the validation parameters [1]. LoD is one of these important method performance parameter that is used both for characterizing the analytical method as well as interpreting the analysis results.

There are several method validation guidelines published by prominent international organisations, which often have contrasting views on the ways of estimating (and even the meaning of) LoD. As a result, there exists a number of guidelines with diverse approaches and recommendations that, when applied in practice, lead to significantly different LoD estimates. Often standards and guidelines can leave decision on how to estimate LoD to the analyst in which case he/she needs further information. Moreover, analysts are often under the pressure from legislation, journals, community, etc. to produce as low as possible LoD estimates. This leaves room for miscommunication about the realistic capabilities of the analytical methods. In addition, sophisticated statistics must be used if LoD is to be estimated without making any assumptions. In order to be feasible for practitioners, all approaches suggested in the guidelines make some assumptions about the analytical system. It is therefore important to understand which assumptions significantly influence the LoD estimate. Therefore, for both fundamental and practical reasons, determining LoD is a complex topic.

The focus of this work was to provide a comprehensive comparative experimental evaluation on different approaches of determining LoD in MS and give recommendations to practitioners on choosing LoD determination approach.

The LoD estimation for MS analysis methods is of great importance for two reasons: (1) MS analysis methods enable detecting analytes at very low levels and are therefore widely used for trace analysis where LoD is an important parameter and (2) the LoD estimation is somewhat dependant on the analytical method being used and therefore the same suggestions might not be appropriate for different analysis methods.

In this work two MS techniques are used to estimate LoD: LC-MS, and paper spray ionization (PSI) mass spectrometry. Various approaches and important aspects of LoD estimation are considered on the basis of examples of LC-MS and PSI/MS methods. Data from these experiments were suitably analysed e.g. to study between-days LoD and to study subjectivity of some approaches to

estimate LoD. Differences in estimating LoD are emphasized when a simple LoD value for characterization of the analysis method is needed or when it is more appropriate to use complex LoD estimation approaches that make less assumptions (e.g. using decision limit (CC_α) and detection capability (CC_β) estimates). The LoD estimates found by different approaches are compared. Also simulations were made to estimate the influence of experimental design on LoD result and experiments were carried out to estimate subjectivity of data analysis.

2. REVIEW OF LITERATURE

2.1. Limit of detection

2.1.1. Definitions and use

Limit of detection (LoD, detection limit) is in most cases defined as the smallest amount or concentration of analyte in the test sample that can be reliably distinguished from zero [2].

LoD is used to make a binary decision: whether the analyte is detected in the sample or not. Therefore there is a possibility of both false positive and false negative results. A false positive result means that the analyte level in the sample is wrongly said to exceed LoD. A false negative result means that the analyte level in the sample is wrongly said to be below LoD [3]. The reliability of this decision depends on the likelihood of making a false positive and false negative decisions. However, the general definition of LoD does not specifically define whether and how these errors must be taken into account. This ambiguity problem of the LoD definition is solved by two method characteristics that have been suggested to be used in place of LoD: decision limit (CC_α) and detection capability (CC_β) [3]. CC_α is generally defined as the analyte concentration (or intensity) level found in a sample by the analytical method under question above which we can state that the probability of the signal being caused solely by noise is below α. This means that a signal above CC_α is with high probability caused by analyte and not by noise. CC_β is the concentration level of analyte present in a sample at which we can say with probability 1-β that the analytical method under question will not give false negative results (meaning results below CC_α) [3]. Graphical explanation of CC_α and CC_β can be found in Figure 1. The values for α and β are usually chosen so that the reliability ($1 - \alpha$ or $1 - \beta$, respectively) is 95% or higher. Thus, CC_α takes into account the probability of false positive results and CC_β the probability of false negative results. However, as CC_β is calculated by using CC_α both types of errors are accounted for.

CC_α and CC_β can be called differently in different guidelines and articles. For example CC_α is called critical value of the net state variable [4], critical level or critical value [5], or decision limit [3,6] and CC_β is called minimum detectable value of the net state variable [4], detection limit [3,5,7], minimum (detectable) true value [5], limit of detection [7] or detection capability [6]. CC_α and CC_β are discussed in more detail in chapter 2.3.

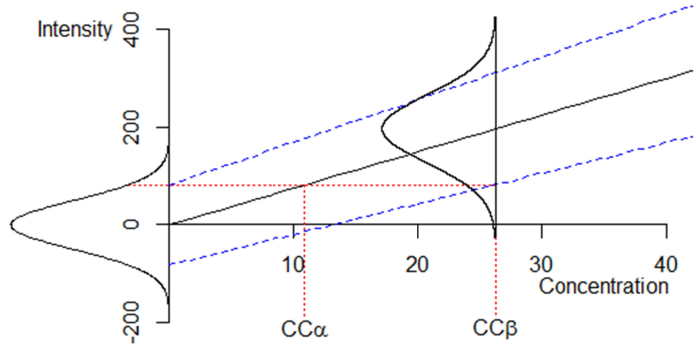


Figure 1. Graphical explanation of CC_α and CC_β. A calibration graph is shown. The dashed blue lines indicate the standard deviation of the signal multiplied by the respective one-sided quantile at 95% probability (1.64). The normal distribution curves show the distribution of values obtained from a blank sample and a sample fortified at CC_β concentration, respectively. Note that the CC_α can also be found in intensity scale from the dashed red line parallel to the concentration scale.

When analysing the different equations of different approaches to estimate LoD (see Table 1) it can be seen that in most cases when the general definition of LoD is used then it is interpreted to be equivalent to CC_β. For this reason it can be seen that in some cases the name "detection limit" is used for both LoD and CC_β. However, this is not always the case. Early in the development of the LoD concept there was a disagreement whether LoD should take into account the possibility of only false positive results or both false positive and false negative results [8]. As a result there are cases where LoD is rather considered as CC_α. However, as probability of false negative results at CC_α is 50% (half the measurements made at CC_α concentration level would give results below CC_α) making reliable decisions is not possible. Therefore, at CC_α false negative results are not taken into account and we recommend that LoD should always be considered equal to CC_β. However, care must be taken, when working with literature, to make sure what is exactly meant when using LoD.

LoD, CC_α and CC_β are used for two purposes: (1) for characterization of the method (e.g. for comparison of two different methods or different laboratories or for comparison with limits set for the method) and (2) for the interpretation of unknown sample results. For the first purpose it is more appropriate to use LoD or CC_β because samples precisely at CC_α are falsely interpreted and therefore this level does not describe the general capability of the analysis method. It should be noted here that in this work the term "method" is used with the same meaning and instead of the VIM term "procedure", because this usage is in line with the common language of the analytical community.

Interpretation of results is somewhat different when LoD is used (as opposed to CC_α and CC_β). The following rules should be followed when using LoD [9]:

(1) If a measurement result below LoD is obtained then it should be stated that the analyte content in the sample is below LoD (it cannot be stated that the analyte is not present) and the LoD value should also be given.
(2) In case the result is above LoD but below Limit of Quantitation (LoQ) then it can be stated that the analyte is present in the sample at trace level and the LoD value should again be given.
(3) In case the result is above LoQ then the determined analyte content is presented together with its uncertainty.

If only the result "below LoD" is given then information about the numerical result of the analysis will be lost [10]. However, this information can be necessary for the end user (e.g. to calculate mean and standard deviation of many samples). Therefore, it has been recommended that the result with its uncertainty should be reported irrespective whether the result is above or below LoD although the uncertainty in that region can be close to the result itself or even higher [5,11]. Giving a value with its uncertainty is necessary for further use and statistical analysis of the results.

In case of CC_α and CC_β the interpretation of the result of a particular sample should be done as follows:

(1) If the result is below CC_α then it can be stated that the concentration of the analyte is below CC_β (or that the analyte is not detected). By quoting CC_β here instead of CC_α the possibility of false negative results is taken into account.
(2) If the result is above CC_α then it can be stated that the concentration of the result is above CC_α (or that the analyte is detected).

As with LoD, when CC_α and CC_β are used for interpretation of data, the measured result with its uncertainty should also be stated.

If LoD is estimated with an approach where it is viewed as CC_β then interpreting the result of a sample using LoD instead of CC_α (i.e. for deciding whether the analyte is present in a specific sample or not) can give false negative results. In these cases there is a possibility that the obtained analyte concentration value from the measurement is by accident below LoD although the true analyte concentration value is above LoD. In the case of result equal to LoD the probability of this error is 50% and it is not taken into account.

If LoD is interpreted as CC_α (e.g. in the case of group 2 approaches in Table 1) a problem arises when using LoD value for characterizing the analysis method. In this case the LoD overestimates the capabilities of the analysis method because the possibility of false negative results is not taken into account.

For many analytes the maximum allowed concentrations, often called the maximum residue limits (MRL), in specific matrices have been set. To take into account the possibility of false positive and false negative errors when interpreting whether the sample is over or under MRL CC_α and CC_β values can be estimated at this level. It should be noted that for example in 96/23/EC CC_α and CC_β are defined so that in case the MRL has been set they should be found only for the MRL and not for the blank value [6]. The difference in this case is that

the signal should be significantly different from a sample with analyte content at MRL level instead of the background. In other words, CC_α is the analyte level detected in the sample above which there is α probability that the signal is caused by a sample with analyte concentration below MRL. To take into account the possibility of false negative results CC_β is also found as the level at which there is β probability of obtaining a result lower than CC_α. The interpretation of the results if the MRL has been set can be done as follows: if the obtained value is over CC_α we can state that the sample contains the analyte over MRL, and if the obtained result is under CC_α we can state with confidence that the analyte content is below CC_β [6]. The measurement result with its uncertainty must be reported as well.

In this work consideration is given only to the estimation of CC_α and CC_β that indicate whether the analyte is detected in the sample or not, because only in this case CC_α and CC_β are related to the topic of LoD.

Distinction is often made between two types of LoD – method LoD and instrumental LoD. Method LoD shows the detection limit of the whole method taking into account all aspects that can influence the signal strength (e.g. sample preparation, matrix effects, etc.) and cause variance in the result, and therefore influence LoD of the whole method. The samples that are used to evaluate method LoD must be matrix-matched and must go through the whole measurement procedure [2,12] which includes all the subsampling (taking sample(s) of suitable size from the initial sample), sample preparation and analysis steps. This is necessary to take into account all the sources of variability [12]. The LoD of an analytical method can be different for different matrices. It has been suggested that the matrix used to estimate the LoD should be reported with the estimated LoD value [13]. Instrumental LoD is estimated for an instrument using analyte solutions in solvent. Instrumental LoD is usually significantly lower than the method LoD, and is meant only for assessing instrument's capabilities and cannot be assigned to an analytical method [2]. In this work LoD always means method LoD.

2.1.2. LoD estimation approaches

There are many LoD estimation approaches and they are conceptually different. This work addresses the approaches that are included in validation guidelines and are most widely used by practitioners. There have been additionally many other approaches (that are often again conceptually different) suggested in the literature. Although some have been discussed and reviewed in somewhat more depth [14] many others have not found much use [15–18] and are therefore not considered here. It can be noted, however, that because of the variability (e.g. between days) of LoD the highly sophisticated approaches do not necessarily give more reliable LoD estimates [19].

In general the approaches can be divided into 3 categories: (1) approaches that need only the measurement of the sample to evaluate whether the analyte is present (e.g. visual evaluation and using signal-to-noise ratio (S/N)), (2) using

standard deviation at a single concentration, and (3) using standard deviation found from calibration data.

The first category is quite different from the other two – it is possible to use S/N and visual evaluation to interpret samples without estimating LoD. The limit set by these approaches (e.g. S/N ≥ 3) can be taken as CC_α when interpreting analysis results (this is because similar to CC_α the decision of "detected" or "not detected" is made at S/N = 3). However, when used in this way it is not possible to characterize the analysis method and compare different methods. For this multiple measurement results must be obtained to take into account both the false positive and false negative results (see Table 1, group 5). Therefore to obtain reliable results with these approaches the following cut-off approach has been suggested: multiple series of samples (whereby the analyte concentration in each series is different) are measured. Usually analysing 10 samples per series is recommended [20,21]. The lowest concentration level where the analyte is detected for all (or a high percentage e.g. 95%) the repeated sample measurements is taken as LoD. This approach takes into account the possibility of false positive results as all (or a certain pre-defined majority) of the samples must be over the limit where it is decided that the analyte is detected. Therefore this LoD estimate is more reliable to use for interpretation than the simple use of S/N or visual evaluation. Another advantage of this approach is its robust nature – no assumptions are made concerning the distribution of the results. On the other hand, this approach demands high number of replicate measurements [12] especially if LoD must be estimated on more than one day (see chapter 2.1.5). This approach is often recommended for qualitative methods [20] with binary measurement results where the decision of detected vs not detected is made (instead of reporting a continuous measurement result). However, in case of continuous results the use of binary reporting alone causes loss of information [22]. To avoid this it is possible to plot the portion of positive results against the measured concentrations. From this plot the lowest concentration at which the required amount of samples give a positive result can be taken as LoD [21].

In this work it is also suggested that results of automatic integration (the software automatically integrates the peak and also decides whether to integrate or not) can be used to decide whether the analyte is present in the sample. Automatic integration systems identify peaks by monitoring the baseline slope (the slope of baseline is over a predetermined value). However, it has been shown that automatic integration results of peaks with lower S/N can be unreliable [23,24]. Therefore all software integration results should be reviewed by the analyst in order to prevent gross errors in peak detection and integration [23].

In conclusion, these approaches are not recommended for estimating LoD in case of LC-MS/MS as large number of measurements are necessary. In case S/N or visual evaluation must be used the cut-off approach together with the S/N values is recommended. However, these approaches can still be used for quick and rough estimation of whether the analyte is present in the chromatogram. For example S/N can be used to find the lowest fortified sample con-

centration where the analyte can be seen in the sample for estimating LoD. As seen in following experiments (see chapter 4.3) knowing the approximate LoD can be useful before doing extensive measurement series for estimating LoD [Paper IV and V].

The second and third category of approaches to estimate LoD are similar to the ones used to define CC_α and CC_β. This general equation can be therefore given:

$$LoD = \frac{(\overline{Y} + k \times S(Y)) - I}{s} \quad (1)$$

where $\overline{Y}$ is the mean intensity value of blank samples, *S(Y)* is standard deviation of results, *I* is intercept of the calibration function (CF) and *s* is the slope of the CF, and *k* is a coefficient similar to the Student's *t* coefficient. Usually *k* is double the value of *t* to account for both false positive and false negative results at the same time. Also the *S(Y)* value is assumed to be same for blank samples and samples with concentration at CC_β. In some cases it is assumed that $\overline{Y}$ and *I* are equal and therefore they cancel out (e.g. group 4, Table 1). It is also possible to first calculate the concentration values corresponding to each measurement and then estimate LoD from these results.
The difference between the second and third category of LoD approaches comes from the choice of how the standard deviation is estimated: the second group uses standard deviation at a single concentration (e.g. blank or concentration close to LoD) and the third group uses standard deviation estimated from calibration data (e.g. standard deviation of intercept or residuals). In both cases a calibration function must be made to estimate the concentration value of LoD and therefore the third group needs somewhat less measurements (replicate measurements at a single concentration are not always necessary). However, in both cases still homoscedasticity and linearity are assumed. These characteristics are discussed in chapter 2.1.3.

CC_α and CC_β can be found with similar equation as equation (1). This general equation however does not take into account the fact that slope and intercept are also estimated from randomly varying measurements. This variability is not taken into account when using the equation (1). More sophisticated approaches have been suggested that use the prediction interval to estimate LoD (or CC_α and CC_β) [4] (also discussed in more detail in chapter 2.1.3).

It can be seen here that many assumptions and simplifications are made in most of the approaches. These are outlined in Table 1. In most cases when guidelines use CC_α and CC_β definitions the approaches make less assumptions than when estimating LoD [4,5]. However, in some cases the same assumptions are still made [6] (see Table 1). The approaches that make less assumptions demand more complex calculations, better understanding of the properties of the analysis method (e.g. scedasticity), and generally a larger number of repeated measurements. However, even the most sophisticated approaches sug-

gested in guidelines make statistical assumptions resulting in biased values. For example it has been shown that the approaches suggested by ISO give CC_α and CC_β estimates that are negatively biased in the case of heteroscedastic data [9].

One of the goals of this work was to study whether the results of these approaches also differ significantly from each other due to their conceptual differences. Therefore, whenever LoD (or CC_α and CC_β) value is stated the approach used to estimate it must also be stated.

Note that not all regulatory bodies have accepted the use of CC_α and CC_β and suggest approaches that do not consider the reliability problem of the LoD definition (see Table 1). The reasons why not to use CC_α and CC_β are that their estimation can be too complex [9,20] at routine laboratory level and in fact not always necessary.

It must be noted here also that LoD (as well as CC_β) estimate is only meaningful in the concentration scale – signal intensity scale (which is different in every instrument) is not appropriate for characterizing a method. CC_α (which intrinsically also refers to concentration), on the other hand, is generally used for interpretation of results obtained with the same instrument in the same laboratory and therefore using it in concentration scale is not strictly necessary [22]. In fact, if the CC_α value is estimated on the signal intensity scale from blank sample signals (e.g. as in 2002/657/EC [6]) then converting its value into the concentration scale by using a CF will introduce errors from slope and intercept (see chapter 2.1.3 on how errors in slope and intercept affect LoD). Therefore, when working with real samples it is often impractical to convert CC_α from signal units to concentration units.

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

Group	Ref.	What is obtained?	Equation	Description	Assumptions, simplifications	Notes
1	[2,12, 20,21, 25,26]	LoD (considers false positive and negative results – the probability of false positive and negative values depends on choice of *t*)	$LoD = \bar{y}_0 + t \times S(y)$ $\bar{y}_0$ 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.	Concentration of fortified samples in LoD range (e.g. lowest level where S/N > 3) or at 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;	Homoscedasticity; normal distribution of replicates; variability of slope and intercept are not taken into account; linearity of calibration data; *t* value is rounded and does not take into account the degrees of freedom; Only for single sample measurement results.	Care must be taken when integrating blank samples; Erroneous CF can lead to negative LoD results; Note that $\bar{y}_0$ is not necessary (taken as 0) if subtraction with intercept (or with $\bar{y}_0$) is done to all results.
2	[13]	LoD essentially equivalent to CC_α (considers only false positive results)	$LoD = t \times S(x)$ *S(x)* is the standard deviation or pooled standard deviation of analyte concentrations from replicate measurements.	A detailed procedure is given to choose fortified sample concentration (incl. estimating an approximate LoD first, measuring only 2 of the needed repeated samples before measuring the rest of the 7 samples); *t* is taken depending on degrees of freedom; Recommended analyte concentration range in fortified samples is 1–5 times LoD.	Normal distribution of replicates; variability of slope and intercept are not taken into account; linearity of calibration data; Heteroscedasticity is somewhat considered by careful choice of fortification concentration; Only for single sample measurement results.	LoD as equivalent to CC_α (false negative results are not accounted for); The background (mean of blank values or the inter-cept value) is subtracted from all other results. It is then suggested to iteratively check the LoD by estimating it again.

Group	Ref.	What is obtained?	Equation	Description	Assumptions, simplifications	Notes
3	[20]	LoD (considers false positive and negative results – the probability of false positive and negative values depends on choice of t)	$LoD = \overline{a} + t \times \frac{S(y)}{\sqrt{n}}$ $LoD = t \times S(y) \times \sqrt{\frac{1}{n} + \frac{1}{n_b}}$ where n is the number of repeated measurements of the sample; $S(y)$ is standard deviation of blank or fortified samples; n_b is the number of repeated measurements of blank samples. Equations give LoD in intensity scale.	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.	Homoscedasticity; normal distribution of replicates; linearity of calibration data; variability of slope and intercept are not taken into account. t value is rounded and does not take into account the degrees of freedom. Allows taking into account the averaging of sample measurement results.	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	[27]	LoD (considers false positive and negative results)	$LoD = 3.3 \times \frac{S_d}{b}$ b is the slope of the CF, S_d can be chosen as standard deviation of blank samples, residuals ($S_{y.x}$) or intercept.	Regression line must be in the range of LoD. CF is used to estimate slope and standard deviation of residuals and intercept. Number of repeated measurements not specified.	Homoscedasticity; normal distribution of replicates; linearity of calibration data; variability of slope and intercept are not taken into account. If repeated results at each calibration level are averaged and standard deviation of residuals is used for estimate LoD then the number of repeated measurements must be the same as repeated measurements for each calibration level.	The standard deviation of intercept underestimates the variance of results at 0 concentration and should not be used. Due to conservative LoD estimates, simple calculation procedure and reasonable workload (S_d is taken from residual values), this is the suggested approach if a rigorous LoD estimate is not needed.

Group	Ref.	What is obtained?	Equation	Description	Assumptions, simplifications	Notes
5	[20,21]	LoD (considers false positive and negative results)	-	Cut-off approach; number of repeated measurements (usually 10) are made at different concentrations near LoD; The lowest concentration at which all the samples are „detected“ is used as the LoD; The detection threshold can be established for example based on S/N, visual evaluation or automatic integration for chromatographic methods.	Uses robust statistics. This approach does not assume normal distribution [5]. Visual evaluation of presence of a peak depends on the analyst.	This approach is very work-intensive; If repeated LoD estimations are needed then this approach is not recommended for LC-MS/MS methods; It has also been suggested to plot the portion of positive responses against concentration to find the lowest concentration at which necessary number of samples give the decision „detected“; Each sample should be independent of the others.
6	[6,21]	CC_α and CC_β	CC_α: 1. $CC_\alpha = \bar{a} + 2.33 \times S_{lab}$ $\bar{a}$ is the average intercept and S_{lab} is the within-laboratory variability of the intercept (found from data at and above minimum required limit) 2. Blank matrices are analyzed to estimate noise in the analyte time window. $S/N > 3$ can be used as CC_α.	Some simple approaches suggested to estimate CC_α and CC_β; Similarly CC_α and CC_β estimation approaches are suggested in case an MRL is set; After estimating the intensity value corresponding to CC_α and CC_β calibration function should be used to convert them to the concentration scale;	Normal distribution of replicates; linearity of calibration data; variability of slope and intercept are not taken into account. Possible heteroscedasticity is considered to some extent: CC_α and CC_β are not found using the same variance. In these approaches the α value is 1 % and the β value is 5 %.	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).

Group	Ref.	What is obtained?	Equation	Description	Assumptions, simplifications	Notes
			CC_β: 1. $CC_\beta = CC_\alpha + 1.64 \times S_{lab}$ S_{lab} is the within-laboratory variability of the mean value at CC_α. 2. $CC_\beta = CC_\alpha + 1.64 \times S_{CC\alpha}$ $S_{CC\alpha}$ 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.	Approach 2 for estimating CC_α and approach 3 estimating CC_β demand at least 20 replicates (at each level for CC_β).	The coefficients in equations do not take into account the degrees of freedom.	
7	[4]	CC_α and CC_β	$CC_\alpha = t_{0.95}(\nu)\frac{\hat{\sigma}}{\hat{b}}\sqrt{\frac{1}{K}+\frac{1}{I\times J}+\frac{\bar{x}^2}{s_{xx}}}$ $CC_\beta = \delta\frac{\hat{\sigma}}{\hat{b}}\sqrt{\frac{1}{K}+\frac{1}{I\times J}+\frac{\bar{x}^2}{s_{xx}}}$ $s_{xx} = J\sum_{i=1}^{I}(x_i-\bar{x})$ $\hat{b}$ is the estimated slope, $\hat{\sigma}$ is the estimated residual standard deviation, $t_{0.95}$ is	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	Normal distribution of replicates; linearity of calibration data; It is suggested to estimate whether the data are heteroscedastic based on prior knowledge and visual evaluation of the data; In heteroscedastic approach standard deviation of results is assumed to increase linearly with concentration	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.

Group	Ref.	What is obtained?	Equation	Description	Assumptions, simplifications	Notes
			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, $\overline{x}$ is the mean value of the con-centration levels, x_i is the concentration if *i*th calibration level.	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.		

2.1.3. Influence of calibration function on the LoD estimate

CF can be used for several different purposes when determining LoD: (1) to convert the estimated LoD values from signal intensity scale to concentration scale, (2) to estimate the background intensity (from intercept), (3) to estimate the intensity variance over the CF via the standard deviation of residuals ($S_{y.x}$). LoD is therefore strongly affected by the correctness of constructing the CF. To simplify calculations most LoD estimation approaches assume homoscedasticity and all approaches assume linearity of the calibration data (see Table 1 in Chapter 2.1.2). Scedasticity is a property of an analytical method to give measurement results with either constant variance within the used concentration range (homoscedasticity) or increasing variance with the increasing concentration (heteroscedasticity) [28]. LC-MS/MS methods, as a rule, give nonlinear and heteroscedastic data [Paper III]. However, for such methods that give nonlinear and heteroscedastic data often a linear and homoscedastic range can be found in a narrower concentration range [11,28]. It must be noted here that in a narrow range the data is still heteroscedastic but the heteroscedasticity is not significant. In order to estimate the appropriate range of data for estimating LoD the following chapters examine more thoroughly the testing of linearity (chapter 2.1.3.1) and homoscedasticity (chapter 2.1.3.2).

If a linear CF is used for nonlinear data then the obtained results will be biased. A possible consequence of biased slope and intercept values is obtaining negative concentration values for low intensity signals [10]. In case of nonlinear CF the calculations to estimate the necessary parameters become more complex [11] and in general assuming linearity in a narrow concentration range can be considered safe especially in the low concentration range [28]. Therefore in order to estimate LoD a concentration range must be found where the response is linearly related to concentration so that a linear CF can be used. Testing linearity of data is discussed in chapter 2.1.3.1.

Conversion of the analysis result of a sample from intensity scale to concentration scale is influenced by the inaccuracy of slope and intercept. Moreover, the variability of the measured intensity value for the sample affects the results [11,28]. The reliability of slope and intercept of the CF can be described by standard deviation values because they are estimated from calibration points that are affected by random variability. The accuracy of the predicted concentration along the CF can be described by the prediction band (prediction interval of all results in the CF range). The prediction band shape shows that the variance is smaller when working closer to the middle of the CF and wider at the edges (see Figure 2). The width of this band is also influenced by the total number of calibration measurements made (number of calibration levels and number of repeated measurements at each level) and the location of the calibration points [11,28]. Larger number of measurements leads to lower variation of slope and intercept and therefore lower variance of the concentration value. As the LoD estimate depends on the variance of the measurement results (and therefore can be calculated from the prediction band) lower variance leads to lower LoD values. These sources of error are usually not taken into account

when estimating LoD. However taking these sources of variance properly into account in the calculation of LoD (or CC_α and CC_β) is very complex (e.g. see the approach suggested by ISO 11843-2 [4]).

Replicate measurements of sample also reduce the width of prediction band as in this case multiple measurement results can be averaged and the standard deviation of the mean is found. As the number of measurements increases the standard deviation of the mean decreases and therefore the procedure that measures the sample large number of times will have a lower LoD. However, note that this is only the case when more than one measurement is made for the sample and the mean result from these measurements is used to interpret whether the signal is above or below LoD. This is because standard deviation of the mean must be used to describe the result and therefore the same parameter must be used to estimate LoD.

Note that in case only one sample measurement is made the prediction interval of a single measurement is used and in case more than one measurement is made (and therefore the mean value is used for further calculations) the prediction interval of the mean is used [11].

The use of the prediction interval is to be preferred over confidence interval, which does not adequately account for the random variability of a single future sample measurement. However, the prediction intervals do not take into account all sources of variance [29,30]. Even more accurate LoD estimation approaches, that use tolerance interval, have been suggested [30]. Tolerance interval takes into account the fact that the parameters that are used to calculate the prediction interval are only estimates. For example a standard deviation value calculated from replicate measurements of a sample is only an estimate because information about the whole population is not collected. Tolerance interval takes this into account by setting a confidence level for limits within which a certain portion of the population falls. It must be noted that as the number of measurements increases the difference between tolerance interval and prediction interval decreases. The approaches that use the tolerance interval to estimate LOD have not been used in guidelines and it has been shown that other statistical aspects are more important (e.g. scedasticity) when estimating LoD [30]. However, the importance of using the tolerance interval rather than the prediction interval to calculate LoD in analytical systems with different characteristics should be studied further.

Another important aspect of calibration data that influences LoD is scedasticity. Many LoD approaches assume homoscedasticity (see Table 1). If this assumption does not hold then the obtained LoD estimate may be significantly erroneous [30,31]. Heteroscedasticity influences LoD estimation because: (1) slope and intercept given by ordinary least-squares regression (OLS) can have biased estimates if data are heteroscedastic, (2) if the LoD estimation approach takes into account the errors of slope and intercept (see discussion above) as in the case of CC_α and CC_β estimation in ISO 11843-2 (see Table 1) then assuming homoscedasticity gives overestimated results, and (3) the LoD estimation approaches that use standard deviation at one specific concentration (or using

$S_{y.x}$) assume homoscedasticity and therefore the LoD is easily over- or under-estimated.

First let us consider the influence of scedasticity on slope and intercept values. If the analytical system yields homoscedastic data then OLS can be used to estimate the slope and intercept of the CF. If (1) the standard deviation at each calibration level is not constant (the data are heteroscedastic), (2) the calibration range is wide, (3) the calibration points are equally (as opposed to having calibration points more densely in the low concentration region) distributed and (4) the measured sample concentration is in the lower end of the range, then the weighted least-squares regression (WLS) should be used to calculate the CF parameters [32]. If all these stipulations are true then it is likely that the CF will not fit the data accurately in the low concentration region when estimated with OLS. This is because the higher concentration levels "skew" the regression line so that it does not pass close to the lower concentration data. As the slope and intercept are used in LoD calculation the LoD estimate will be influenced as well.

In WLS the weights are calculated for each calibration point so that higher variance leads to lower weights. These weights are then used to assign "importance" to the calibration points when calculating the CF parameters: the points with lower variance influence the CF more and the regression line moves closer to these points. WLS however requires more complex calculations and may require more measurements compared to OLS [28].

In addition, different approaches have been proposed for calculating weights and it can be complicated to determine which approach should be used. For example the following equation can be used to calculate the weights:

$$w_i = \frac{s_i^{-2}}{\frac{1}{p}\sum_{i=1}^{n} s_i^{-2}} \tag{2}$$

where w_i is the weight of the given concentration level, s_i is the variance of the concentration level i, n is the number of measurements at each concentration level, and p is the number of calibration levels [11,28]. The weights should not have units and therefore scaling is necessary (the s_i^{-2} are divided by the average of all s_i^{-2} values in the given equation). These approaches demand that more than one repeated measurement is made at each calibration level and can therefore be too demanding for everyday use. Often simpler approaches are used for calculating weights (e.g. $1/(x^z)$ and $1/(y^z)$ where x is concentration of the calibration level, y is the measured signal of the calibration level, and z can be chosen as 1, ½ or 2) where multiple measurements in each concentration are not necessary [33]. The assumption forming the basis of using these simplified weights is that it is previously known how the repeatability of calibration points depends on the concentration/signal. The choice of z depends on how fast the variance of results increases as concentration increases: the change in the

variance should correspond to the change in the weight meaning in case of faster increase of variance larger z value should be used. When using the simpler approaches for weighting then scaling of the weights is still necessary.

To choose which weighting approach is best for use or if weighting is necessary at all (and therefore to test whether the data are heteroscedastic or not) it is possible to calculate the relative residual values for the different models. Although often relative residuals calculated from concentrations are suggested to be used for this purpose [33,34] the same conclusions can be made using the relative residuals calculated in intensity scale. Here it is suggested to use $1/x^2$ values as weights (normalisation of these weights is again necessary).

Second, if WLS is used then the prediction interval for concentrations at the lower end of the calibration line becomes narrower (similar to the variance of intensity values in case of heteroscedastic data) [11,28]. It must be noted here that WLS itself does not change the scedasticity of the data or the prediction interval but using the appropriate approach to estimate the prediction interval that takes into account the weights at each concentration gives different prediction intervals. This is because when weights are taken into account the narrowest region of the prediction interval is not in the middle of the CF but is moved towards the lower concentrations (see Figure 2). Therefore this prediction interval gives a more correct description of variability, especially in the lower concentration region, if the data are heteroscedastic. This in turn leads to a more correct LoD (or CC_α and CC_β) estimate. However, it must be noted here that in most cases LoD approaches do not take this information into account.

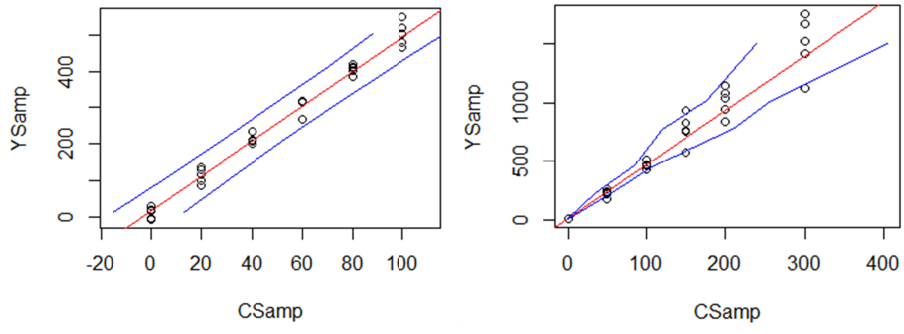


Figure 2. Simulations of data in R: the left figure shows prediction interval of the mean (in blue) in the case of homoscedastic data calculated with OLS (regression line is shown in red); the right figure shows prediction intervals of heteroscedastic data calculated with WLS, same legends are used as in both. In both cases the prediction interval is calculated for samples (sample data points shown as circles) that are simulated separately of the calibration data. It can be seen from the figure with heteroscedastic data that when a small number (in this case 5) of repeated measurements are made for each calibration level then the prediction interval width can change irregularly with concentration because of random variations in the data points.

Third, most of the LoD estimation approaches (except the ISO 11843-2 approach and the robust cut-off approach) in Table 1 either explicitly or implicitly assume homoscedasticity. A narrower range, in which the data are nearly homoscedastic (heteroscedasticity is not significant), can always be found [11] when estimating LoD to give an accurate estimate of the standard deviation of the result. When such a range is found then the use of WLS becomes unnecessary and therefore for LoD estimation only OLS is needed.

2.1.3.1. Testing linearity

Many different approaches can be used to evaluate the linearity of the data [11,28]. Here we discuss the following approaches: visual evaluation of the calibration graph and residuals, and Lack-of-Fit test [11,28,35].

First it is possible to evaluate whether the data is linear by constructing a calibration graph and fitting a linear CF to the data. If the data points scatter randomly around the CF the data can be assumed to be linear. However, this approach only gives a rough estimate and linearity cannot be confirmed because random scatter as such cannot be confirmed. Moreover, the evaluation of linearity is subjective as the results depend only on the evaluation of the analyst. Therefore, it is suggested here that absolute residuals or (if possible) some other approaches should be used to confirm linearity.

Absolute residuals are calculated by the following equation:

$$e_{abs,i} = y_i - (b \times x_i + a) \tag{3}$$

The obtained residuals are plotted against their concentration values. The random scatter of residuals around 0 refers to the linearity of the data. Although this approach is simple it may not be entirely objective and the evaluation of the result can be difficult if the data are heteroscedastic or each calibration level is measured only once. Experiments were conducted to test the effect of this subjectivity (see Chapter 3.3 and 4.1).

Another alternative to test linearity of data is the Wald-Wolfowitz runs test. In this test specific signs are given to data points that are above or below the calibration line. If there are many data points with the same sign in a row then the data can be said to be nonlinear. However, this test demands many data points to show that nonlinearity is significant and it cannot be used in this work [28].

SANCO also suggests using relative residuals (absolute residual values divided by the estimated signal at the given concentration) to estimate linearity and sets allowed limit of ± 20% for residuals [36]. However, high relative residual values will be obtained not only if linear model is used for nonlinear data but also if the data are heteroscedastic and OLS is used in place of WLS to estimate the CF parameters [34]. As WLS is complex and not often used it must be known that the range being tested for linearity with relative residuals is homoscedastic. Evaluation of scedasticity demands more than one replicate measurement at each concentration level (see Chapter 2.1.3) and therefore

compared to using absolute residuals more measurements must be made for using this approach. As with absolute residuals the final evaluation whether the data are linear is made subjectively by the analyst. OLS, WLS and testing heteroscedasticity of data are discussed in more detail in chapter 2.1.3.

The other two approaches use the principles of Analysis of Variance (ANOVA) to estimate linearity. The linearity of data is tested by testing whether a linear calibration model fits the data. In ANOVA the F test (a simple division the two variances that are being compared) is used to compare the random variance of measurement result to the variance caused by systematic factors (such as nonlinearity). The result of the test therefore shows whether the variance of the analysis results caused by the change in the factor is significantly different from the random variance of the measurement results. In other words it is possible to estimate whether the change in the factor has a significant influence on the result [28]. After the F value has been received from the equations it is compared to critical values in a table (the table can be found for example in [28]). If the F value surpasses the critical value (for chosen p value which shows significance) then the two variances are significantly different: the random variance is significantly smaller than variance of results when taking into account the change in the factor. This means that the change in the factor significantly changes results. Both of the tests discussed below should be interpreted as one-sided tests as the results show whether one variance is larger (not whether one is different than the other).

In the Lack-of-Fit test the variance of mean values against estimated values of the model (error of the model) at each concentration level is compared to the random variance of measurement results. The equation for this test is therefore:

$$F = \frac{s^2_{mean(y.x)}}{s^2_y} = \frac{\frac{1}{p-2}\sum_{i=1}^{p} n_i(\bar{y}_i - \hat{y}_i)^2}{\frac{1}{n-p}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2} \tag{4}$$

where n_i is the number of measurements on each calibration level and n is the number of all measurements, p is number of calibration levels, y_{ij} is the result of the j-th measurement at the level i, $\bar{y}_i$ is the mean value of the measurement results on level i, and $\hat{y}_i$ is the measurement result estimate from the model for level i [11]. A CF where $s^2_{mean(y,x)}$ is significantly larger than s^2_y means that the variance of the calibration points around the proposed calibration model is larger than the variance of the measurement results of calibration points. Therefore, it can be concluded that the given model does not fit this data and some other model should be tested. If this model is a linear model then a conclusion can be made that the data is not linear (or that a simpler model should be used by e.g. removing the intercept value from the model).

The second approach compares the variances of residuals given by two different models. In case simpler model (e.g. linear model) gives a significantly

greater variance the more complex model should be used (e.g. nonlinear model) and therefore nonlinearity of data can be assumed. However, if the variance does not change significantly the simpler model should be used [35]. For this test the following equations are used:

$$SS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5}$$

$$dSS = \frac{SS(simple) - SS(complex)}{\nu(simple) - \nu(complex)} \tag{6}$$

$$F = \frac{dSS}{\frac{SS(complex)}{\nu(complex)}} \tag{7}$$

where ν is the degrees of freedom calculated by subtracting the number of parameters in the model from the number of measurements made. *SS*(*simple*) and *SS*(*complex*) are the sum squares of residuals of the simpler model and the more complex model respectively [35]. Therefore the F value is actually measured from the decrease in variances of the residuals when using a more complex model. If the improvement achieved with a more complex model for the explanation of data points is negligible (F value will be low) then the simple model should be used. If the variance of the residues for the complex model is significantly smaller than for a simple model then a large F value will be received and the complex model should be used. This approach, however, demands that a nonlinear CF must be fitted for data. As choosing the correct nonlinear function is not always obvious and fitting it to the data can be complex and is rarely used in LC-MS this approach is not discussed further here.

It must be mentioned here that the correlation coefficient (or its squared value) is not a good indicator of whether the data are linear or not [28] and therefore this statistic cannot be recommended for this purpose.

More thorough discussion on the topic of linearity of data has been given in many statistics books (e.g. by Miller, Danzer, Mandel [11,28,35]).

2.1.3.2. Testing scedasticity

In this chapter different approaches are discussed that can be used to estimate whether the data are homo- or heteroscedastic and therefore to estimate the homoscedastic range if necessary.

First, the simplest and often suggested approach is to visually evaluate whether the variability of the results increases with increasing concentration [4,11]. This is usually evaluated from a plot with absolute residuals against concentration. Although this approach is simple it is hypothesized that it can be somewhat subjective. In order to study this, different plots were simulated in R and 14 analysts were asked to evaluate whether a plot was homo- or heteroscedastic (see Chapter 3.3 for plots and 4.1 for results). As previously ex-

plained, if the data are heteroscedastic then the CF estimated by using OLS can produce strongly biased concentration predictions at lower analyte levels. All the relative residuals at lower concentrations may therefore be significantly higher than at high concentrations (meaning that either positive or negative residuals strongly dominate) [34]. If WLS is used the data points at low concentrations have higher influence on the CF parameters and the CF will therefore be less biased. As a result the relative residuals also fall closer to 0. This effect on relative residuals can be used to estimate whether WLS should be used in place of OLS and also whether some WLS models using different weighting equations give significantly better fit to the data [34]. The CFs found with different approaches can also be compared by comparing the sum of squares of relative residuals – lower sum of squares means better fit to the data [33,34]. However, it can happen by random chance that the OLS CF will pass through the data points at low analyte levels, and therefore this cannot be considered a demonstration of homoscedasticity of data. Moreover, if the data are nonlinear and a linear model is used then the relative residuals will also have large relative residuals at low concentrations and therefore it cannot be distinguished whether the data is heteroscedastic or nonlinear.

A more complex approach to estimate whether the data are heteroscedastic is to use the Hartley test (also called F_{max} test) [11]. F_{max} test is an F test where the measurement result of the concentration level with the largest variance is compared to the result of the level with the smallest variance. In the case of calibration the following equation can be used:

$$F_{\max} = \frac{s_{\max}^2}{s_{\min}^2} \tag{8}$$

where s_{max} is the standard deviation of the calibration level with the highest variance and s_{min} is the standard deviation of the calibration level with the lowest variance. The F value is then compared to a critical value from the table for Hartley test. Note that this test is not the same as the usual F test as not just the variance of two datasets is compared. As the number of levels increases the likelihood of the datasets with smallest and largest variance having values that are significantly different only by chance with the usual F test increases. This is taken into account by the critical values of the Hartley test. This test, however, requires that calibration levels have a number of repeated measurements (preferably at least 4). Other approaches are available to estimate heteroscedasticity but are more complex [11] and are therefore not discussed further here.

2.1.4. Experimental design

To conclude the previous chapters it is now appropriate to ask how the parameters of the calibration data should be chosen for estimating the LoD and how they influence the LoD estimate: (1) including blank values in the calibration data; (2) concentration of the lowest calibration level used; (3) concentration of the highest calibration level used; (4) number of calibration levels and number of repeated measurements at each calibration level; (5) distribution of the calibration levels in the concentration range.

(1) The use of blank values when estimating LoD (or CC_α and CC_β) is a complex topic in general [37] especially when estimating LoD for an LC-MS/MS method. For estimating LoD an estimate of mean and standard deviation of blank values must be made. In order to obtain comparable results from the chromatograms with peaks and blank samples the same integration approach should be used for both. The following is therefore suggested: the integration is done so that all the noise over the baseline (the mean value of noise around the peak area) is integrated meaning the "integration line" is drawn on the baseline. However, the data analysis software of the instrument used in this work does not take into account the data points that are below the integration line. As a result the integration result of a blank sample is always positive and with a lower standard deviation than the integration results from chromatograms with peaks (standard deviation decreases because the variance of points below the integration line are not taken into account for blank samples, but these points are above the baseline if a peak is present). However, if only chromatograms with visible peaks are integrated then the decision to integrate (or not to integrate) will be subjective as it depends on whether the analyst decides that the peak is present. Moreover, information about the analyte might already be present in a chromatogram where a peak could not be visibly seen by the analyst therefore leading to erroneous results. Also extrapolation is needed in this case (mean and standard deviation of blank samples must be found from data that contain the analyte). In conclusion direct estimation of standard deviation and mean of blank samples could not be done properly for the given data. Estimating these parameters from intercept and $S_{y.x}$ has been suggested [37] but as discussed above is also not without assumptions.

In this work the following approach is used: the blank chromatograms are integrated and the obtained values used in the calibration data. Although, as explained above, this cannot be done in an ideal way, this is nevertheless better than leaving them completely out. This approach is used for several reasons. First, integration of blank chromatograms avoids loss of information because integration of the chromatogram is always performed. Including blank measurements into the calibration data is suggested by ISO [4]. Moreover, this approach is found to be acceptable because using the estimated intercept or mean of blank values to estimate LoD does not give significantly different results (see chapter 4.3.2) and therefore the diffe-

rence between the two approaches is not significant for the given data. The standard deviation of blank values is used to estimate LoD and the results are also compared to other approaches (see chapter 4.3.1). However, further discussion on whether to integrate blank samples and to include the results in the calibration data is necessary.

Also it can be concluded here that statistical model with intercept taken as 0 (as simpler model can be preferable) cannot be used here. This is because the blank and intercept values do not have a mean value of 0, and the blank or intercept values are also not subtracted from the measurement results.

(2) In order to choose the lowest concentrations, simulations of calibrations with different concentration levels were made to study the effect of using concentration levels below CC_α and CC_β for estimating $S_{y.x}$ and the slope for calculating LoD. Results of simulations can be found in chapter 4.2.

(3) The highest concentration value that can be used in the CF depends on the linearity and heteroscedasticity of the analytical method. All suggested approaches that use CF to estimate LoD assume linearity and therefore the highest calibration level that can be used should still allow linear fit. At least 6 calibration levels have been suggested to be used for estimating linearity [Paper III] in addition to the blank values. However, when validating a method with unknown LoD or linear range more levels should be planned as many can fall below LoD (and give only noise) or above the linear range. If the used approach assumes homoscedasticity then the highest concentration level can be chosen so that (in addition to linearity) the data can be shown to be homoscedastic in that range (see chapter 2.1.3.2 for testing scedasticity). If the results from linearity or scedasticity test shows that the current data are not consistent with our assumptions then the highest calibration level data can be removed and the tests repeated. It is therefore useful to know the approximate range in which LoD is in so that not too many calibration levels should be left out of the data due to nonlinearity or heteroscedasticity (see discussion in chapter 2.1.3).

(4) The number of calibration points and calibration levels should be chosen so that at least relevant tests can be made (e.g. test for homoscedasticity and linearity). For example at least 4 repeated measurements at each level are suggested when estimating homoscedasticity with the F test. More replicates also increase the reliability of the results. Moreover, with a larger total number of calibration points LoD can be estimated more reliably (larger number of measurements leads to less uncertainty about the parameter we try to estimate, including LoD). Therefore this number also depends on the practical need of the LoD application.

(5) It is suggested that the calibration points should be equally distributed when estimating linearity [Paper III]. However, for practical purposes it is advisable to choose the calibration points so that their concentration differs e.g. 3 times in case an approximate LOD value is not known beforehand.

The choice of calibration points influences the LoD estimate differently depending on the approach used and the scedasticity of the data and therefore different approaches are reviewed separately in the following discussion.
In case of approaches that use the variance at only one concentration (e.g. blank or fortified samples, see Table 1 group 1) the slope and intercept influence the LoD estimate. The slope does not depend on the distribution of calibration points when the data are homoscedastic and linear. When the data are heteroscedastic the slope is strongly influenced by the choice of calibration levels only for some certain cases [32] in which case using WLS would be more appropriate but often not practical. If the LoD estimation approach uses also the variance found from the CF (e.g. the $S_{y.x}$, see Table 1 group 4) then it must also be considered how the choice of different calibration points influences this variance. For example if $S_{y.x}$ is used and the data are homoscedastic then the distribution of calibration points has little influence on LoD. However, if the data are heteroscedastic then using more calibration points at higher concentration levels will increase the $S_{y.x}$ and therefore increase LoD.

For approaches that take into account the variance of CF-s parameters (e.g. ISO 11843-2, see Table 1 group 7) moving the mean concentration of the calibration points to a lower value (meaning using rather calibration points with lower concentrations) the LoD can decrease but only in the case of homoscedastic data. However, this change is insignificant (see chapter 4.2). In case the data are heteroscedastic and therefore WLS is used the weights associated with calibration levels influence the results so that LoD does not depend on the concentrations chosen for calibration levels.

It is not advisable to prepare the calibration solutions with consecutive dilutions due to accumulation of error. Also the experiment should be planned so that the solutions would be measured in random order to so the instrumental drift could not influence the results in systematic way [Paper III].

Some important notes must be made about calculating the $S_{y.x}$ from the measured calibration graph data. First if more than one replicate measurement is made at each calibration level and the results in each level is averaged before calculating the $S_{y.x}$ then the same should be done to the future samples. For example 4 replicates must be made and averaged for all the future samples if 4 repeated measurements are made at each calibration level and their results are averaged before estimating the $S_{y.x}$. Secondly it is important to note that the weighted $S_{y.x}$ estimated when using WLS instead of OLS cannot be used similarly to the unweighted $S_{y.x}$, because it does not describe the standard deviation in any specific location in the calibration graph.

2.1.5. Day-to-day variability of LoD

LoD varies between days [2,20] and this variability can be more significant than with most other validation parameters (e.g. precision, trueness, sensitivity). The same holds for CC_α and CC_β [19,38]. It has been shown that when estimating CC_β on only 3 separate days the obtained values can have a relative standard

deviation (RSD) of more than 50% [19]. The experiments made in the course of this work show that in case of LC-MS/MS LoD estimates of some approaches can differ by up to 10 (!) times between days [Paper IV and V].

The reason for the high variability of LoD values under the same conditions is that LoD is estimated from parameters that are random variables and are strongly dependent on the (sometimes subtle) variations in experimental parameters. This effect is amplified in the case of LC-MS (and also for MS alone), because it has a large number of adjustable parameters and some of them are not easy to control [Paper II]. Some of the sources of variability are reasonably constant within day and cause only day-to-day variability of LoD (e.g. cleanliness of the MS system and ion optics, small differences between batches of chemicals, mobile phase pH). Most vary also within a day and cause both within-day and day-to-day variability of LoD (e.g. small fluctuations of temperature and gas pressures). A well-known manifestation of between-days variability of parameters is the variability of the calibration graph slope, which should be remeasured for every day/sequence [Paper II and III]. As a result, the "true" LoD value can be significantly different on different days. Therefore, LoD estimated only once for a method with an MS detector can be used only if the LoD estimate is not used for any significant decision and the working range relevant for the samples under study starts more than an order of magnitude higher than the found LoD.

In the course of this work we have suggested to divide the LoD values into the following categories: within-day LoD, between-days LoD and between-labs LoD. The differences between these LoD values are explained in Table 2.

In most cases LoD value is not estimated every day. However, if it is known that the LoD significantly changes between days and if the LoD is critically important for the interpretation of the results then the safest approach is to estimate LoD every day. If the value of LoD is not critically important (e.g. LoD is known to be well below the analyte concentration in samples) then LoD determined on a single day is sufficient [20]. However, even in this case it is strongly advised to periodically re-evaluate LoD [20].

Although this topic has not been discussed in necessary detail in the literature some suggestions can be found in different guidelines about estimating the between-days LoD can be found in different guidelines. It is recommended to estimate LoD on separate days using different batches of reagents and materials that normally change as the method is used [25] for a more reliable estimate. It has been suggested to evaluate the between-days LoD from single-day LoD values by taking a median of the results [4], however, this would mean that in roughly 50% of the cases the laboratory is unable to detect the analyte at LoD level and cannot be therefore recommended. For this reason we have suggested a more reliable approach by using a quantile (e.g. 95% level). It has also been suggested that monitoring CC_α and CC_β values between days via a control chart can be used for keeping the performance of the analytical method under control [38]. Eurachem has suggested using low concentration quality control results measured on different days (under intermediate precision condi-

tions) to estimate LoD [20]. In this manner it is possible only to use some simpler LoD estimation approaches and only between-days LoD is estimated which might not be always fit for interpretation of the results.

The between-days LoD should be used in two cases: (1) the LoD is first of all necessary for characterising the analytical method but not as much for interpreting results (for example, results in the LoD range are far below the MRL), or (2) LoD (or CC_α and CC_β) is an important parameter that is used for critical interpretation of results but the random variation in *determination* of LoD is the main reason for difference between days (meaning the true LoD value itself does not change between days). In both cases LoD should at least be estimated on separate days while validating the method, or regularly over a longer time period and the data should be used to estimate between-days LoD. In the second case more frequent estimation of LoD is suggested as the between-days LoD becomes more reliable and therefore best for use. Moreover if the LoD (or CC_α and CC_β) are critical for evaluation it should be evaluated whether the parameters that LoD is calculated from change significantly between days in order to estimate whether it is more correct to use within-day or between-days LoD. If yes then the variability does not stem from determination variability but the LoD is indeed different on different days. To test this it is possible to compare the within-day repeatability of LoD values and day-to-day variability of LoD values with F test [28]. If the latter is significantly larger, then it can be concluded that LoD varies significantly between days. For this test LoD should be estimated minimum on 3 separate days and twice per day.

In conclusion between-days LoD is suggested for use instead of within-day LoD in most practical cases. As more data are collected over a longer period to estimate between-days LoD the estimate becomes more reliable.

The between-labs LoD should only be used for characterization of the method. Therefore, between-labs LoD can be used to compare different analytical methods and to determine whether they are fit for purpose. The between-labs LoD can be evaluated from between-days LoD measurements with similar approaches as when estimating between-days LoD values from within-day LoD results (e.g. taking the median value of between-days LoD-s of different labs). It must be made sure that the approaches used to estimate LoD are the same as LoD estimates of different approaches are not comparable [31,39]. These experiments would be complicated and costly, and in many cases it is possible to compare the methods by experience.

The LoD estimates can vary between labs not only due to random variability of uncontrolled conditions but also due to differences in e.g. equipment. For example, the same LC-MS/MS method with different MS instruments can have different LoD values. For this reason the between-labs LoD estimate should be accompanied with information about the minimum and maximum LoD estimates for this analytical method. Therefore, when comparing two different analytical methods boxplots (robust statistics) can be used where it can be seen for example that although the between-labs LoD is higher for one method but with appropriate measures similar or even lower LoD values can be achieved.

Only when comparing these parameters of different methods can it be stated with some reliability that it is possible to reach a lower LoD value with one of the methods. It can also be concluded from there that if a new analytical method with significantly lower LoD values is proposed it cannot be assumed that the between-labs LoD value will be significantly lower than for the other methods.

Table 2. Appropriate uses of different proposed LoD values for interpretation of the analysis result and for characterizing the analytical method in different situations.

	Within day LoD	**Between-days LoD**	**Between-labs LoD**
For interpretation of results	If LoD is an important parameter and the between-days variability is large; It is recommended to use CC_α and CC_β	Should be used if within-day experiments are not practical	Should not be used
For characterization of the method	Should not be used	Should be used for characterization of the methods performance within the lab	Should be used for comparison between methods and evaluation of what LoD can be expected

2.1.6. Qualitative and quantitative analysis

Analysis methods can be divided into two groups: (1) methods that provide a discrete measurement result, and (2) methods that provide a continuous numerical result (see Figure 3). The focus of this work is on MS methods, which produces continuous measurement results and therefore the suggestions in this work only apply for this group of methods. The continuous result given by the method however does not mean that the end result must also be presented as such. For example from the results of continuous methods information can be extracted about what compounds the sample contains (e.g. NMR or IR spectroscopy, or high resolution MS), or into which group the sample falls (e.g. based on principle component analysis of data, is the analyzed wine a Merlot and in which region is it produced), or binary information about whether the analyte is present in the sample or not. In fact it can be seen that the interpreted result around LoD is binary and therefore qualitative but the results that are given by the measurement can still be continuous. Only the interpretation gives the result a discrete value.

Qualitative analysis methods are defined somewhat differently in different sources. Here the following definition is used: qualitative analysis methods are methods in which substances are identified or classified on the basis of their chemical or physical properties [20]. This also includes binary “yes” or “no” answers [20] and therefore analysis methods that give results that are interpreted

using LoD are qualitative methods. Further in depth discussion about analysis methods with binary responses (from both continuous or binary measurements results) and their quality assurance can be found can be found in the literature [22].

Quantitative analysis methods on the other hand are defined as method which determines the amount or mass fraction of a substance so that it may be expressed as a numerical value of appropriate units [6]. Therefore, if a sample concentration is high enough for quantitation the analysis method can be considered quantitative but it the concentration is near LoD the method can be considered qualitative.

Another important definition for this work is the screening method which can be defined as methods that are used to detect the presence of a substance or class of substances at the level of interest [6]. The important aspects of screening methods are that they should be simple to use, have high throughput and be positively biased. After a sample is found to contain the substance at the level that is important then the sample can be analyzed again with a validated quantitative analysis method. Therefore only the samples that are more likely to contain the analyte in the range of interest go through the quantitative analysis. Therefore this allows laboratories to analyse more samples quickly and expand the analytical scope to analytes that are in most samples not likely to be present [6,36]. Because the screening method only has to say whether the analyte is detected in the range of interest LoD is an important parameter for these methods. Also as LC-MS/MS methods are complex, expensive and time demanding a screening method done prior to this analysis can be useful. Ambient ionization methods provide a good possibility to develop screening methods based on MS (see chapter 2.3).

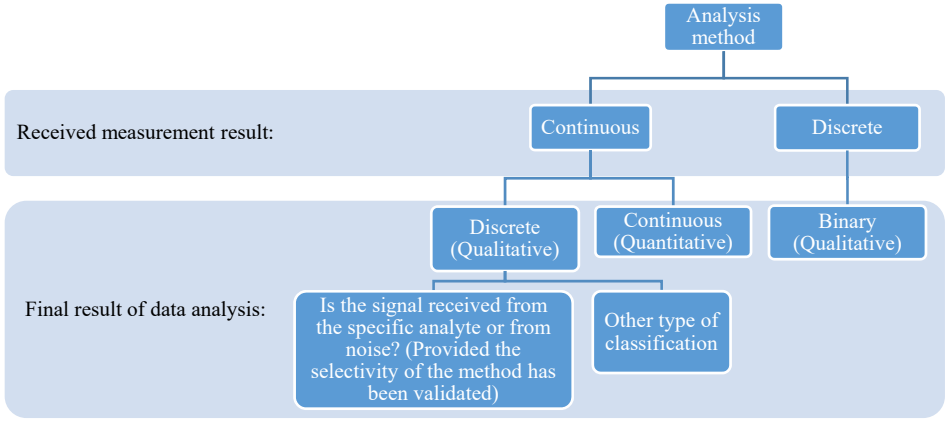


Figure 3. Classification of analysis methods based on the type of results.

2.1.7. Bayes theorem and its use in estimating LOD

Common statistical tests (including CC_α and CC_β estimates) do not let the analyst take into account all sources of information. As an example let us consider a lab that measures a biomarker to identify the presence of a disease in patients of one hospital. The probability of a false positive result of the test is 5%. The hospital orders 200 tests in a year. On an average 10 of them are false positive. From previous years it is known that the probability of a true positive result is 1/200. Therefore if 200 tests are carried out in one year and 11 positive results are discovered then it is possible that on an average only one of them is truly positive. In this example the probability of false positive test results is therefore 10/11 which equals roughly 91% and not 5%. This *a priori* knowledge of the probabilities of true positives and true negatives can be taken into account by using the Bayesian statistics, which enables calculating the probability of a positive sample being truly positive. In general the true positive and true negative values can be viewed as different hypotheses. Although this example is for discrete measurements the same principles can be applied to continuous analysis results.

Recently some possibilities to estimate LoD that take the prior knowledge into account have been suggested [40,41]. Moreover, an approach has been suggested on the basis of information theory, which takes prior knowledge (and non-normal distribution of noise, see chapter 2.4) into account [18]. These approaches are not widely used, however. The main difficulty when using Bayesian theorem is the numerical estimation of the prior probability of the result [42], which can be very difficult. Therefore we recommend to consider using Bayesian statistics to estimate LoD only when reliable prior knowledge is available. It is clear that this prior knowledge significantly influences the result, e.g. if the probability of obtaining a false positive result is high and the probability of the sample being truly positive is low. Further information on using the Bayes theorem can be found in review articles by Armstrong and Hibbert et al. [42,43].

2.1.8. Choosing LoD estimation approach

In this section important practical considerations of which approach to use for the estimation are highlighted.

First it should be considered if it is at all necessary to estimate LoD. For example LoD does not need to be estimated when it is known that the samples measured by the analytical method are always significantly higher than the methods LoD.

If there is a requirement to estimate LoD with some specific approach then the approach must be followed. The regulatory bodies often make guidelines for specific applications (e.g. pesticide measurements) and take into account the suitability of the approach for that particular field. However, often standards

and guidelines can leave many decision on how to estimate LoD for the analyst in which case the suggestions in this work can still be useful.

Before moving forward to LoD estimation the stability of the analyte should be considered. Stability of the analyte can influence the measurement results and therefore can influence the LoD estimate. If there is a good reason to believe that the analyte stability influences the measurement results significantly its impact should be estimated before validation of the method (including LoD estimation) and if necessary appropriate measured must be taken to account for this influence [Paper III].

As shown in previous works [31,39] different approaches give different LoD estimates. Moreover, it can be seen from the previous discussion that it is simple to "manipulate" with the LoD estimates (e.g. by using different experimental design) and therefore the precise description of the approach used to estimate the LoD value must be given whenever the LoD value is reported. Also the units of the result should be presented together with LoD [12].

If it is possible to choose between different LoD estimation approaches then it should be first made clear, which approach is compatible with the characteristics of the given analysis method. If LoD is estimated for characterizing the analysis method and the working range is significantly above the LoD then it should be estimated with a simple and conservative approach (otherwise inappropriately large effort is spent for LOD estimation). Conservative LoD estimates are suggested here to assure that the stated LoD can indeed be achieved when necessary. In case of LC-MS/MS we suggest using the ICH approach that uses a CF and $S_{y.x}$ to estimate LoD (see Table 1, group 4). The CF should be determined in the LoD range and the data must be sufficiently linear to use a linear CF. If the data are heteroscedastic the LOD values will be overestimated. If a lower LOD value is needed for the analysis method then finding the narrow homoscedastic range for calculating LOD is possible. The reasons for suggesting this approach and more detailed suggestions on the experimental design to estimate LoD are further discussed in Chapter 4.3 and 4.4.

In some cases important and costly decisions might be made on the basis of analysis results where LoD is used for interpretation. In such cases simple approaches for LoD determination might not be sufficient [30] and more complex approaches should be used. Multitude of different parameters and assumptions must be considered when a reliable LoD is needed – complex experimental design (that demands large number of measurements), and complex tests and calculations must be used. In this case our suggestion is to use CC_α and CC_β estimated by the equations given in ISO 11843-2 [4]. Relevant tests should be performed to assess which assumptions can be made (e.g. heteroscedasticity should also be tested in addition to linearity). In addition, the differences in LoD values between days must be estimated and it should be considered whether LoD needs to be estimated daily. For LC-MS/MS both the day-to-day variability and heteroscedasticity are likely to be important.

It is suggested that after calculating LoD the results can be validated by analyzing samples near the LoD [27]. For example it has been suggested to check

whether an estimated LoD is achieved within an analytical run it has been suggested to add two spiked samples with concentrations at LoD as a first and a last sample to be analysed [36].

It is likely that most applications that use LoD do not need the complex approaches for its estimation. For these applications the simple approaches suggested in guidelines fit well unless significantly underestimated values are obtained due to large bias in the approach used to estimate LoD. However, the simple approaches described in the guidelines make assumptions and simplifications that can significantly influence the estimated values. For example tolerance interval should be used in place of prediction interval.

As a final action it should be considered if the obtained LoD estimate is fit for the purpose of the analysis method. If it is not then improvement of the analysis method is necessary and thereafter the LoD must be estimated again.

2.2. LC-MS/MS

2.2.1. Overview of LC-MS/MS

The extensive use of LC-MS has become possible largely due to the advent of the atmospheric pressure ionization (API) methods [44]. The API sources are able to produce gas-phase ions with little or no spontaneous decomposition from delicate and high molecular weight analytes. This, combined with the intrinsic sensitivity of mass spectrometers, has revolutionized large areas of chemical analysis where traces of organic analytes are determined in complex matrices. Among the ionization methods electrospray ionization (ESI) has proven especially versatile [44]. As a result, almost all fields of chemical analysis (bioanalytical and medical, environmental, food, drug discovery [45], etc.) have experienced big changes.

The success of the LC-MS technique arises from its ability to give three-dimensional data. First, the compounds are separated in time by LC. Ions generated in the ionization source are then separated according to their m/z ratios in the mass analyzer of MS. Finally, the MS detector measures the abundance of each ion. In addition to these dimensions tandem mass spectrometry (MS/MS) enables significant increase in selectivity (and decreased noise) by fragmentation of ions to monitor specific product ions that are produced from the analyte. Compared to the traditional LC detectors, such as ultraviolet – visible spectrophotometry or fluorescence, the MS detector therefore enables significantly more reliable identification of the compounds eluting from LC. Moreover, MS often allows detection of analytes at lower concentrations than other detectors and therefore is often used for trace analysis.

The extensive additional possibilities, however, come at a cost: LC-MS systems are complex and a large number of parameters have to be at or near optimal values in order to get the desired performance [46,47]. This automatically means that whenever an analytical method based on LC-MS is developed, its performance has to be carefully checked and monitored. There-

fore, method validation is a key activity in LC-MS analysis, indispensable for obtaining reliable results [48].

Although there are many benefits to using MS as a detector, the accuracy of the obtained results may be strongly influenced by the ionization suppression/ enhancement occurring in most API sources. Both ionization suppression [49] or ion source contamination [50,51] result in variability of both MS signal and the obtained results. As seen in previous chapters LoD depends on the repeatability (and therefore accuracy) of the analytical method. Therefore, it is important to estimate LoD (CC_α and CC_β) with matrix matched samples and for all the samples to go through the whole method procedure (see chapter 2.1.1 about method LoD).

2.2.2. Properties of LC-MS/MS

Due to difference in the way that analytical techniques give signals not all LoD estimation approaches are appropriate for all analytical methods. In the case of LC-MS/MS it can often be seen that the baseline is at 0 and its standard deviation is 0 or very low [52]. This occurs if signal processing (e.g. thresholding) is used by the instrument software to improve the S/N of the peak [53]. This kind of data "censoring" by the instrument software (or hardware) means that some data are lost because the same value (e.g. 0, if the signal strength is below some threshold) is always given to results below it. This may cause non-normal distribution of analyte signals from low-level samples [2,54] and erroneous mean and standard deviation values as well as calibration graph parameters may be obtained [10]. If LoD is estimated assuming normal distribution of results the estimate will be erroneous even in case the standard deviation of the distributions is the same [18]. Therefore, assumptions made by most of the LoD estimation approaches may not be valid.

Robust statistics can be used in cases when normal distribution cannot be assumed (due to thresholding or some other reason). Approaches have been suggested to estimate LoD from these data [16,17,55]. EP17-A [24] recommends estimating LoD with the approach proposed by Linnet et al [55]. The approaches using robust statistics can be also used if the data contain outliers [16,55]. Although these approaches can be simple to use they can give LoD values that are higher and vary significantly between days due to properties of robust statistics. An approach often suggested in guidelines that uses robust statistics (see Table 1, group 5) sets a cut-off threshold to classify detected and not detected analytes. Although robust, a large number of measurements is required by this approach. Therefore these approaches are only suggested if alternatives are not possible to use or give significantly erroneous LoD estimates.

An important aspect to consider when using LC-MS/MS is the method of obtaining analyte signal from the chromatogram: measuring the analyte's peak height or measuring its peak area. The peak area takes into account more information, does not assume identical peak shape in standard and sample

solutions at different concentrations, and has a larger dynamic range [23]. Peak height and area are not comparable in characterizing an analytical method and if peak areas are used for quantification then LoD must also be estimated from peak area data. However, due to difficulties in identifying the start and end of a peak at low S/N the area can have lower precision than height at low S/N [23]. This lower precision can lead to higher LoD values. However, when using the area of a peak the intensities of the data points are averaged and therefore this average value of a peak must be above the averaged intensities of blank sample for the peak to be detected. Averaging always increases precision and in this case will lead to decrease of LoD. Therefore it cannot be concluded that using peak heights would provide different LoD results. In this work peak area is used due to its significantly wider use.

Another problem that arises here is that it can be difficult to extract reliable data from samples at low concentrations [12] because close to LoD noise blurs the beginning and the end of a peak leading to reduced confidence in the result of integration [23]. Moreover, if blank samples are used to estimate LoD an integration method must be used that is capable of providing comparable results for samples with and without peaks. In order to obtain these comparable results the same integration approach should be used for both. In the article published in the course of this work [Paper IV and V] we have suggested integrating of all the noise over the baseline (the mean value of noise around the peak area) so that the "integration line" is drawn on the baseline. However, the data analysis program used in this tutorial does not take into account the data points that are below the integration [Paper IV and V]. As a result the integration result of a blank sample is always positive and with a lower standard deviation than chromatograms with peaks. Standard deviation decreases because the variance of points below the integration are not taken into account, but these points are above the baseline if a peak is present. However, not integrating blank and low concentration samples (because there is no visible peak) means that information about these samples is discarded based on the subjective decision made by the analyst. Information about the analyte might already be present in a chromatogram where a peak could not be seen by the analyst therefore leading to erroneous results. As a possibility to solve the problem of estimating the mean and standard deviation values of blank samples they can be assumed to be equal to the intercept value and $S_{y.x}$ [37]. However, including blank measurements into the calibration data has also been suggested [4]. In this work the blank values are integrated and used in the calibration data. Therefore no information is lost because integration of the chromatogram is always performed and extrapolation of data is not needed (mean and standard deviation of blank samples should not be found from samples that contain the analyte). However, from the results it can be seen that using the estimated intercept or mean of blank values to estimate LoD does not give significantly different results (see chapter 4.3).

From the discussion above it can be concluded that practical limitations of analysis method (instrumental or other) must be considered when choosing an approach to be used for LoD estimation.

2.2.3. Monitoring of fragments produced by MS/MS near LoD

So far false positive and false negative results due to quantitative variability of measurement results have only been considered. However, errors in results can also be qualitative, i.e. caused by interfering components in the sample that give identical signal to the analyte. This problem is amplified at LoD level by the low analyte concentration and may lead to the signal apparently exceeding the LoD, while in reality the signal is caused by an interferent and analyte content may be below LoD in the sample. To avoid this error the selectivity of the analytical method must be assured at the LoD level.
In LC-MS/MS selectivity (and correct analyte identity) is assured by observing the retention time on the chromatogram, and by recording the signals of more than one fragment ion specific to the analyte. In addition, peak shape can offer useful information. Furthermore, the relative abundance of the fragments is measured and compared to the relative abundances found with the standard substances. The article published in the course of this work addresses the different approaches of identity confirmation in LC-MS/MS [Papers II and III].

The fragment ions used for identity confirmation (qualifier ions) might have significantly lower signal intensity than the most abundant fragment ion, which are usually used for quantitation (quantifier ion), and therefore might not be observable at LoD level if LoD is determined using the quantifier ion. To avoid false positive results reliable confirmation of identity is important and therefore it is often required that LoD is defined in such a way that one or more qualifier ions are also detectable at LOD. For example 2002/657/EC [6] demands that the S/N value of all observed ions should be above 3. If the qualifier ion(s) have low intensity signals then the consequence of this requirement is that method's LoD increases significantly. An approach has been suggested to defining LoD in such a way that the need of observing the relative abundance of ions at LoD level is accommodated [56]. However, this topic has not been discussed thoroughly and needs further research.

2.3. Paper spray ionization

2.3.1. Ionization methods of MS

One of the main difficulties in using MS as a detector is the production of gas-phase ions from analyte molecules. Many ionization sources have been developed that function with different mechanisms and are able to produce analyte ions from different phases. Another important characteristic of an ionization method is the internal energy transferred during the ionization to the compound. For example electron ionization transfers large amount of energy to gas phase compounds and as a result the compounds fragment extensively. However, in case of chemical ionization (CI) and ESI the energy transfer is small and the gas phase analytes produce molecular ions. Ionization methods have also been developed to produce gas phase ions from solids, most important

of which is matrix assisted laser desorption ionization (MALDI). In addition to ESI atmospheric pressure chemical ionization (APCI) and atmospheric pressure photoionization (APPI) are two approaches that can ionize compounds from solvents: first the solvent with the sample is nebulized after which ionization of the compounds is achieved with a coronal discharge (similar ionization mechanism to CI) in case of APCI or with ionizing ultraviolet radiation in case of APPI. [57]

However, the most widely used ionization method for analytes in solvents is ESI. ESI has allowed the possibility to ionize analytes that are not volatile (e.g. large molecules like proteins) expanding the capabilities of MS greatly. ESI works by spraying the eluent from a needle to which electrical potential has been applied. As a result electrochemical reactions take place and excess positive charge is created into the liquid phase in case positive potential is applied to the needle (positive ionization mode). The excess ions collect to the liquid surface meniscus and due to the repulsive forces between the ions the liquid surface expands and forms a Taylor cone. At some point the repulsive forces overcome the surface tension and droplets with excess positive charge are formed. In case of ESI however pneumatic nebulization is also used to help the formation of droplets. [58]

These droplets start losing solvent as it evaporates. As a result the excess positive charge will once again overcome the surface tension and produce a stream of smaller droplets. This process is repeated several times as the same happens with the produced smaller droplets. [58]

As with MALDI, ESI ionization mechanisms are not understood to full extent. Two theories have been offered about the production of gas phase ions from the droplets neither of which are universally accepted. In the ion evaporation model it is described that the analyte moves to the surface of the droplet and desolvation of the analyte ion takes place [58]. This can happen only from small droplets with diameter of approximately 10 nm or less as only then the repulsive forces are strong enough for the desorption to happen [59]. In the charge residue mechanism droplet evaporates until only the analyte is left and a gas phase ion is formed. The analyte is usually ionized by protonation as there is an excess amount of H^+ ions in the droplet. However, the possibility of forming adducts with other ions is also possible (e.g. Na^+) [58]. These models describe the ionization of different molecules – large sphere shaped proteins are likely to ionize by the charge residue model and smaller molecules by the ion evaporation model. Another model that has been suggested is the chain ejection model where a disordered polymer leaves the surface of the small droplet similarly to ion evaporation model [59].

Strong matrix effects can influence ESI as other compounds beside the analyte compete for the charge and the surface of the droplets by supressing the analyte ionization. It must be noted here however that other matrix compounds can also produce ionisation enhancing effects [60]. Therefore, extensive sample preparation is necessary prior to ESI/MS analysis to remove the other matrix compounds [Papers II and III]. This can make analysis methods significantly

more complex. These matrix effects can be somewhat decreased by using nanoESI ionization where a needle with much smaller diameter and lower eluent flow rates are used. As a result droplets are formed only due to repulsion of ions (no pneumatic assistance is necessary). These droplets are significantly smaller and therefore there is more surface from where the analytes can desolvate. However, nanoESI is somewhat capricious and the needle can clog easily. [58]

Due to these problems a new family of ionization methods have been developed called ambient ionization methods. Ambient ionization methods are generally defined as methods that ionize analytes from samples under ambient atmosphere conditions and where the sample needs little or no pre-treatment prior to analysis. By now over 30 different ionization methods belong to this family and can be divided into 3 groups: (1) direct ionization, (2) direct desorption/ionization, and (3) two-step ionization. In the first ions are created from solvent (eluent on a surface or in a needle or from droplets) by high electric fields [61]. The ionisation mechanisms here are similar to ESI and nanoESI. In the second group the liquid or solid sample is on a surface which is then "bombarded" with charged droplets, ions, atoms or photons and as a result the compounds in the sample are brought into the gas phase and ionized. In the third and largest group the sample is first brought into gas phase or droplets of sample solvent are created after which ionization takes place by using charged droplets (created by ESI mechanism), plasma or similar mechanisms to APCI and APPI. A wide variety of applications have been shown to be fitting for ambient ionization techniques (from analysing different samples from surfaces to monitoring chemical reactions and medical analysis) [61]. Although the analysis can be made with little sample pre-treatment in the course of this work it was found that the matrix effects can still strongly affect the analysis [Paper I] and therefore further development of these methods is necessary. The most promising of this wide variety of ionization sources are desorption electrospray ionization, low temperature plasma ionization and PSI [61,62].

2.3.2. PSI mechanisms

In PSI [63] the sample is first transferred onto a paper triangle (e.g. with base of 5 mm and height of 10 mm) with a sharp tip in front of an MS entrance. Voltage of approximately 3000–4500 V is then applied between the paper and the MS entrance, and finally the whole triangle is wetted with an eluent (in some cases eluent is added before the voltage is applied). As soon as this is done, the liquid starts spraying and the sample components are ionized and directed toward the MS. PSI therefore falls under the direct ionization group of ambient ionization methods. Paper has several advantageous characteristics for this type of ionization: well-known properties in chemical analysis, possibility of chemical modification, availability, low cost and ease of fabrication. Moreover, paper can be used as a substrate on which the sample is collected and then transported to a laboratory. For example analytes in dried blood spots on paper have been shown

to be more stable (and easier to handle) than in blood itself [64]. Therefore, with PSI it is possible to connect the sample pre-treatment and ionization into one step [63].

The mechanisms of the different processes occurring in PSI have been studied. It has been shown that a saturation limit for the amount of sample exists starting from which adding more sample does not increase the signal intensity [65]. This amount depends on the size of the paper used for PSI. The mechanism by which the liquid moves on the paper has also been studied. It was found that capillary forces and the movement of excess liquid on the paper toward the spray are responsible for the majority of the liquid's movement. Movement due to the electrophoretic forces was not found to be significant [65]. It has been also found that two different ionization mechanisms work in PSI. During the first period of spraying when there is still enough eluent on the paper to form a Taylor cone, the ionization mechanism is similar to nanoESI. However, after most of the eluent has been used up and visually no more spray is seen, higher electrical current is measured and spectra similar to APCI can be seen. It is therefore theorized that an electrical discharge and desorption of analyte molecules occur [66].

Because of its ease of use, low cost and fast analysis time PSI analysis methods are good candidates for screening methods. It has been shown that a great variety of analytes including amino acids, peptides, proteins, herbicides, therapeutic drugs and fatty acids can be ionized with PSI. Also PSI has been applied for analyses of different matrices, such as urine and, especially, blood [63]. Out of the different possible applications of PSI, the possibility of measuring therapeutic drugs from dried blood spots has gained the most attention [67–70]. Moreover, when a whole piece of tissue is placed on the paper, hormones, lipids and therapeutic drugs can be identified [71]. PSI can also be used in food analysis. When a cola drink was used as eluent, caffeine could be identified from the spectrum. Also, thiabendazole and imazalil could be identified from the peel of an orange with wiping method where the orange is wiped with the paper that is later used for PSI [63].

However, as commonly no sample pre-treatment is made strong matrix effects and large repeatability of results can be seen. In the course of this work it was found that the matrix of the sample not only influences the ionization mechanism at the tip of the paper but can also influence the movement of eluent (and therefore also the analyte and other matrix compounds) to the tip [Paper I]. As a result isotopically labelled internal standard (IS) are commonly used for quantification of results. A known amount of IS is added to the sample and when measurements are made the intensity value of the IS and the analyte are recorded. The calibration function is made in the scales of intensity value of analyte divided by the intensity value of the IS (in place of intensity value of the analyte on the y-axis) and the concentration value of the analyte divided by the concentration value of the IS (in place of concentration value of the analyte on the x-axis). After the results are received the concentration of the analyte can be calculated by multiplying the calculated concentration ratio by the concentration

of the IS in the sample. This method takes into account many different variabilities that are caused by the matrix and systematic effects on that specific analysis as the intensity of the IS also changes due to these errors. However, it has been noted that even though repeatability of results significantly improve the LoD of the method still suffers from matrix effects [68].

2.3.3. Estimation of LoD for PSI approaches

Similar approaches as brought in Table 1 have also be applied to PSI methods [72–75]. The only difference, however, is that as IS is used the LoD found from these approaches is not in the scale of analyte concentration in the sample but in the analyte and IS concentration ratio scale. Therefore, to receive LoD in the analyte concentration value the unitless LoD must be multiplied with the concentration of the IS in the sample.

It must be noted here that the concentration of the IS used in the samples must be the same for standard and samples – it is not possible to decrease the LoD value by decreasing the concentration of IS. However, small changes in the concentration of IS (e.g. when weighting is used to produce the samples and standards with known IS concentration the concentrations will never be exactly the same) are allowed. Large concentration changes can cause changes in LoD value due to heteroscedasticity and nonlinearity. Moreover, the IS can influence the ionization efficiency of the analyte (by supressing the ionization of the analyte due to competition for charge and surface in the droplet) [76] and therefore using a different concentration of IS can change intensity values given by the analyte causing systematic error between measurements. This also means that the concentration of the IS should not be chosen to be too high so that the analyte ionization would not be significantly supressed.

In conclusion whenever an LoD value is estimated for an analysis method it is assumed that the IS concentration will be the same in the sample that will be measured in the future.

3. EXPERIMENTAL

3.1. LC-MS/MS

The following conditions were used for all the LC-MS/MS measurements. Agilent Series 1100 LC system was used for the chromatographic separation. An aqueous buffer with 1 mM ammonium acetate and 0.1% acetic acid (eluent A), and methanol (eluent B) were used as mobile phase components. For both analysis methods the eluent flow rate was 0.8 ml/min and sample injection volume was 5 μL. The column temperature was set at 30 ^{0}C. Agilent LC/MSD Trap XCT ion trap mass spectrometer was used for MS/MS detection. DataAnalysis for LC/MSD Trap Version 5.2 (Build 374) program made by Bruker Daltonik GmbH was used for data processing (including calculation of S/N ratio) and analysis. The analytes were recorded in segments where specific *m/z* values for the analytes were recorded around the time of its elution from the column.

All the stock solutions and following dilutions were made by using the Sartorius ME235S GENIUS balance (with 0.01 mg resolution).

Water was purified with MilliQ Advantage A10 system. For the buffer solution acetic acid (Sigma-Aldrich, ≥ 99.8%) and ammonium acetate (Fluka, BioUltra, ≥ 99.8%) were used. HPLC grade methanol was acquired from Sigma-Aldrich.

It must be noted that due to the high number of replicate measurements by some of the LoD estimation approaches all the requirements of all approaches cannot be followed.

3.1.1. Pesticides

For pesticides a 250 mm (4.6 mm inner diameter) Agilent Eclipse XDB-C18 column with 5 μm particles was used and the following LC gradient was used: from 0 to 20 minutes component B content was increased from 20% to 100%, from 20 to 25 minutes the component B was kept constant at 100%, from 25–27 minutes the component B was decreased from 100% to 20% after which 7 minutes of post-run time was used. The retention times of pesticides were the following: spiroxamine 18 minutes, imazalil 16.2 minutes, triazophos 20.3 minutes, propamocarb 5.7 minutes, thiabendazole 10 minutes, carbendazim 8.7 minutes.

Commercial Agilent ESI ionization source was used. The nebulization gas pressure of 50 psi was used. Drying gas with flow rate of 12 l/min was used at temperature of 350 ^{0}C. The following transitions were used for the compounds: spiroxamine 298 to 144 *m/z*, imazalil 297 to 255 *m/z*, triazophos 314 to 162 *m/z*, propamocarb 189 to 144 *m/z*, thiabendazole 202 to 175 *m/z*, and carbendazim 192 to 160 *m/z*. The fragmentation amplitudes of the compounds were optimized using the software's built-in function.

The following chemicals and materials were used for QuEChERS sample preparation: acetonitrile (acquired from Sigma-Aldrich, HPLC grade), acetic

acid (Sigma-Aldrich, ≥ 99.8%), magnesium sulfate anhydrous (Lach-ner, assay purity 99.2%), sodium acetate (Reakhim, the former Soviet Union, dried at 60 ^{0}C for at least 48 h), PSA bonded silica (Supelco). All the pesticide standards were acquired from Dr. Ehrenstorfer GmbH.

Stock solutions of all pesticides (kept at -20 °C) were made in acetonitrile which were used on 6 separate days for preparing the calibration samples. All the pesticide measurements were made from homogenized tomato matrix that was pre-treated with the QuEChERS method [77]. Before spiking the blank samples were analysed to check whether they are truly blank. Two separate dilutions were made for each calibration sample (including the blank) and 2 parallel experiments were made for each sample. In total of 5 measurements were made with blank samples. In total of 10 different calibration levels (together with blank) were made. The concentration of each subsequent calibration level was approximately 2.3 times higher than the previous level. However, for two lowest calibration levels the difference was 12 and 5 times. The total concentration range was therefore below 5 orders of magnitude.

3.1.2. Antibiotics

For antibiotics Phenomenex Synergy Hydro-RP 250 x 4.6 mm column with 4 μm particles of C18 stationary phase was used. The following LC gradient was used: from 0 to 10 minutes component B content was increased from 20% to 90%, from 10 to 15 minutes component B content was kept constant at 90%, from 15 to 20 minutes component B content was decreased from 90% to 20% after which 7 minutes of post-run time was used. The retention times of imipenem, doripenem, meropenem and cilastatine were 3.8, 5.6, 6.6, and 9.8 minutes, respectively.

For LC-MS interface two different ESI nebulization systems were used. First, the commercial Agilent ESI nebulizer originally designed for the used MS system and, second, a novel ESI nebulizer 3R [78] with an added inner capillary that directs additional nebulization gas to the tip. The parameters of the 3R nebulizer have been optimized in earlier works. The outer and inner diameters of the capillaries of the 3R nebulizer were respectively 4 and 2 mm, 0.8 and 0.55 mm, 0.203 and 0.089 mm [78]. The nebulization gas pressure of the commercial ESI was 50 psi and the gas pressures for the inner and outer capillary in case of 3R nebulization were 14 bars and 2 psi respectively. The drying gas flow rate and temperature of ESI and 3R nebulization were 10 L/min at 350 ^{0}C and 10 L/min and 325 ^{0}C respectively. The MS/MS data analysis program segment times and each segments parent to precursor *m/z* were respectively: 0 to 5.1 minutes 300 to 256 *m/z*, 5.1 to 6.1 minutes 421 to 342 *m/z*, 6.1 to 7.5 minutes 384 to 340 *m/z*, 7.5 to 12 minutes 359 to 202. The fragmentation amplitudes of the compounds were optimized separately for the two different nebulization methods.

Doripenem was purchased from AK Scientific Inc. (Union City, CA, USA), meropenem was kindly donated by AstraZeneca Limited (Macclesfield, United

Kingdom). A mix of imipenem and cilastatine was purchased from Merck Sharp & Dohme Corp. (New Jersey, USA).

LoD measurements of meropenem, doripenem and cilastatine were repeated on 5 separate days and imipenem on 3 separate days with both commercial Agilent ESI and 3R nebulizer. Stock solutions of all compounds were prepared with concentrations of approximately 1000 mg kg^{-1} in water and were stored at -80 °C. The calibration solutions were prepared from the stock solutions by successive dilutions with water.

For all antibiotics 8 calibration solutions were measured. In the case of meropenem, doripenem and cilastatine 10 separate samples were prepared for 4 calibration points with lowest concentration and one sample was prepared for 4 calibration points with the highest concentrations. In the case of the commercial nebulizer the used concentration range was 0.25 to 100 μg kg^{-1} for meropenem and 1 to 500 μg kg^{-1} for doripenem and cilastatine. For 3R nebulizer the concentration range was 0.25 to 50 μg kg^{-1} for meropenem and 0.5 to 250 μg kg^{-1} for doripenem and cilastatine. Due to imipenem's instability (see below), all calibration points of this analyte were measured with one separate sample. For imipenem the concentration range was 60 to 500 μg kg^{-1} for both ESI and 3R ionization. 10 separate blank samples were prepared and measured in each sequence.

3.2. PSI

Whatman Grade 1 filter paper (GE Healthcare, Little Chalfont, UK) was used for all measurements with PSI. Solvent containing 20% of 0.1% formic acid with 98.0–100.0% purity (Fluka, Buchs, Switzerland) in ultrapure water prepared with MilliQ Advantage A10 (Merck Millipore, Billerica, MA, USA) and 80% of HPLC grade acetonitrile (Avantor Performance Materials (JT Baker), Center Valley, PA, USA) by volume was used as eluent for PSI. Imazalil standard substances was acquired from Dr. Ehrenstorfer GmbH (Augsburg, Germany) and the imazalil-d5 isotopically labelled IS was acquired from Sigma-Aldrich (St. Louis, Missouri, USA). Acetonitrile was used as the solvent for standard and stock solutions.

The tomatoes used for testing were all acquired from a local supermarket. Retsch Grindomix GM200 (Retsch, Haan, Germany) was used for homogenization of the samples at 5000 rpm for 15 seconds and an additional 20 seconds more at 10 000 rpm.

The homogenated samples were spiked so that the analyte concentrations were approximately 0.0081, 0.035, 0.074, 0.26, 0.58, 0.90, 1.16 and 1.58 ppm. All the samples were also spiked with the IS so that its concentration was approximately 1.6 ppm. An additional blank sample was prepared where only the IS was added. All the fortifications were prepared by weighting by using the Sartorius ME235S GENIUS balance (with 0.01 mg resolution). Samples were carefully mixed on a Vortex mixer (VWR International, Leuven, Belgium) to guarantee the homogeneity of the spiked pesticides in the sample. No other

sample preparation steps were taken, so as to keep the sample pre-treatment quick and simple, and therefore more fitting to be used as a screening method.

Varian 320 triple quadrupole MS model number: MS0906A002 (Varian Inc., Palo Alto, CA, USA) with nanoESI housing in the positive ion mode was used for measurements. The specific equipment necessary for PSI measurements were developed [Paper I]. A specifically designed template was used for fabricating isosceles paper triangles with height of 10 mm and base of 5 mm. The paper triangle was placed into a holder, made from stainless steel, which was fabricated to fit into the nanoESI housing. The holder with the paper was placed in front of the MS inlet so that the distance between the inlet and the paper tip was approximately 5 mm. The nanoESI housing allows for precise positioning of the paper tip in all three axes.

Capillary voltage and collision energy were optimized for the analyte and IS with the native nanoESI ion source of the same MS instrument. The emitter voltage was set to 3500 V and shield voltage to 300 V when PSI experiments were conducted (the shield is a part of the ion optics in front of the MS inlet that helps transmittance of ions into the capillary). Drying gas pressure was set to 15 psi and temperature to 150 °C. Argon gas pressure of 1.5 mTorr was used for fragmentation. The electron multiplier detector voltage was set to 1300 V.

After fixing the triangle in the holder the sample was applied on the paper triangle. This is done by pipetting approximately 2.5 µL of sample onto the paper using a 2.5 µL automatic micropipette (precise pipetting is not possible due to high viscosity and heterogeneity of the sample). The sample is dried in ambient air after which the holder with the paper is placed into the nanoESI housing. Then the voltage was applied between the triangle and the MS inlet. The recording of the spectrum was started and thereafter 20 µL of eluent was added using a 20 µL automatic pipette (Eppendorf AG, Hamburg, Germany). This sequence of operations enables recording of the whole spectrum and avoids partial vaporization of the eluent before the spray is initiated. The multiple reaction monitoring mode was used to record MS responses of the analyte and IS. The monitored transition for the analyte was 297 to 159 m/z and for the IS was 302 to 255 m/z. Optimized capillary voltage and collision energies were for the analyte 68 V and 20.5 V respectively, and for the IS 64 V and 13 V respectively.

5 replicate measurements were made for the blank sample and sample with analyte concentration of 0.58 ppm. Other samples were measured once. Integration of the IS and analyte signal was made over the same length of time (over the time where IS signal was visibly observable over the background noise) from the recorded chronogram for each sample. In some cases the IS signal was not observed due to poor repeatability. Therefore, usable data could not be obtained and the data was not used in the analysis. These experiments were repeated on 3 separate days. The shield, inlet and holder were cleaned between every measurement. Regular measurements were made without any sample to check for carryover. No carryover was noted in the course of these experiments.

3.3. Demonstration of subjectivity of statistical tests

It was hypothesised that in the following cases the interpretation of data can be subjective (depending on the analyst): (1) testing linearity based on absolute residuals plot, (2) testing scedasticity based on absolute residuals plot, (3) interpreting whether the analyte is present or not in the chromatogram by visual evaluation, and (4) integration of chromatograms. To test these hypothesis, and the extent of the influence that the subjectivity can have on the results, a test was conducted. 14 analysts with experience in LC were given the same instructions and the same data for interpretation.

In the first two cases data was simulated in R with known properties (e.g. linear or nonlinear for first and homoscedastic or heteroscedastic for second case) using random number generator function in R. When testing linearity the simulated data was homoscedastic and when testing scedasticity the data was linear. Absolute residual plots were created from this data. In both cases 6 different plots were chosen for the analysts to interpret: two were chosen so that their correct interpretation would be straightforward (e.g. homoscedastic data was plotted that could be easily interpreted as homoscedastic), the other four were chosen so that they would be difficult to interpret correctly (e.g. linear data was simulated but a plot was chosen so that it might be easy to interpret the result as nonlinear). The plots can be seen in Figure 4. These plots were shown to the analysts and it was asked to interpret whether the data was linear or nonlinear in case of linearity testing and whether the data was homoscedastic or heteroscedastic in case of scedasticity testing.

In the last two cases chromatograms from LC-MS/MS analysis of samples containing pesticides (see description in chapter 3.1.1) at 5 different concentrations were given to the analysts for interpretation and integration. The concentrations of the samples were close to the LoD so that in some samples analyte could not be identified but in others the analyte would be clearly present. All samples contained 5 different pesticides – carbendazim, imazalil, triazophos, spiroxamine, propamocarb. The presence of the analyte was interpreted in 4 separate replicate measurements at each concentration and integration was done in only one replicate measurement of each concentration. The analysts were asked to evaluate every chromatogram and state whether the peak of the analyte is present in the chromatogram or not. The integration “technique” explained in chapter 2.1.4 was explained to all analysts.

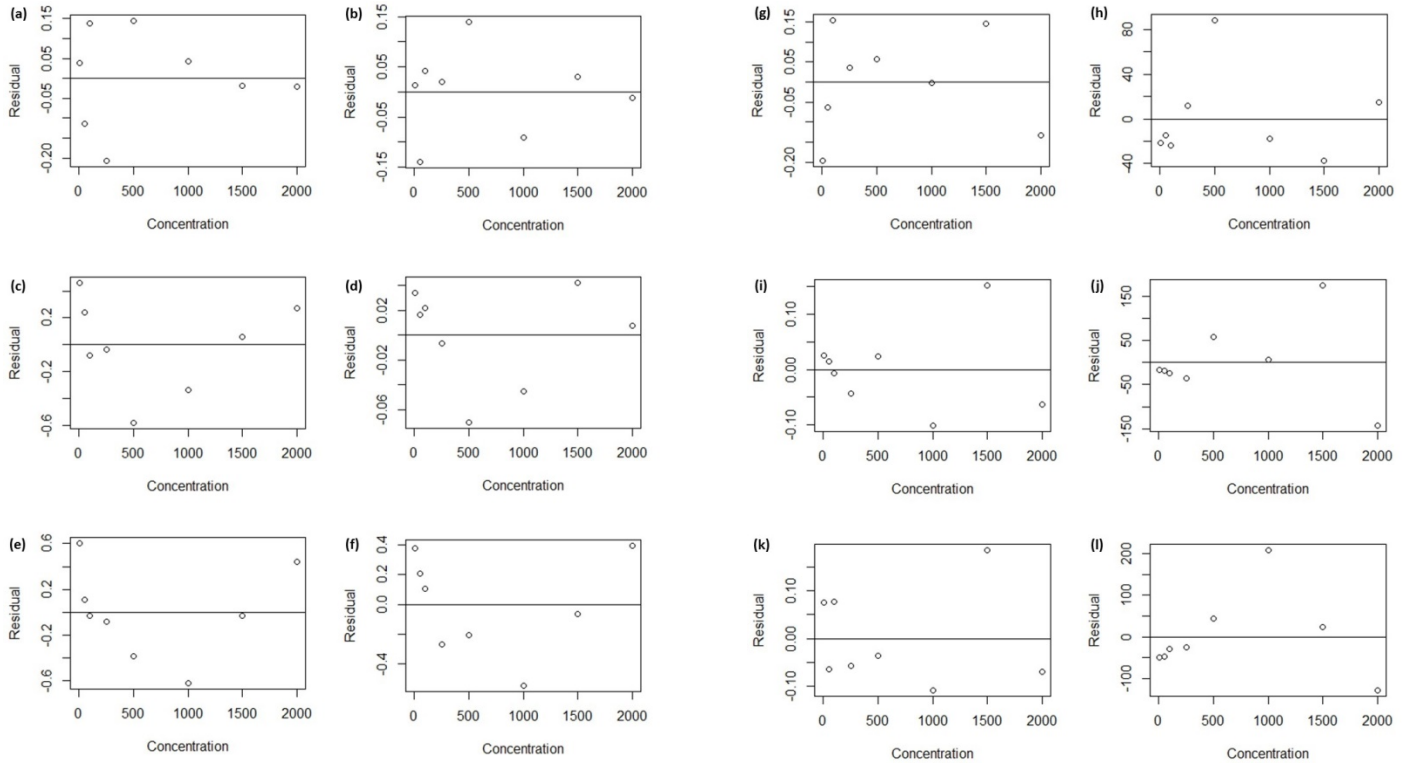


Figure 4. Plots (a)–(f) were used for linearity testing: the data are truly linear in plots (a), (b) and (d) and nonlinear in plots (c), (e) and (f). Plots (g)–(l) were used for scedasticity testing: the data are truly homoscedastic in plots (g), (i) and (k) and heteroscedastic in plots (h), (j) and (l).

3.4. Simulations

Simulations in R were carried out to study how strongly the LoD value is affected if calibration points used to calculate it are below CC_α and/or CC_β. For this a script was written that simulates a linear calibration function data with known variance and therefore the correct value for CC_α and CC_β was known. Both homoscedastic and heteroscedastic data were simulated. This R script can be found in Appendix 1. In these specific cases the CC_α and CC_β are 32.9 and 65.8 for homoscedastic data, and 32.9 and 73.11 for heteroscedastic data. This script is then used to simulate data measurement results 1000 times at each calibration level (producing 1000 separate CF-s) at each following concentrations: (1) 0, 75, 100, 125, 150, 175, 200, (2) 0, 35, 60, 85, 110, 135, 160, (3) 0, 5, 25, 50, 75, 100, 125, and (4) 0, 5, 10, 15, 20, 25, 30. Notice that in the last case all calibration levels are below CC_α. The true value of slope is 5 and intercept is 0. From each calibration function that is simulated the following parameters were calculated: slope, intercept, $S_{y.x}$, standard deviation of intercept, LoD calculated from ICH suggested approaches (see Table 1, group 4) using the $S_{y.x}$ and standard deviation of intercept. To analyse these data the mean and standard deviation of these calculated parameters was found. From this data it can be seen how strongly the LoD values are affected by the choice of calibration levels below LoD and what is the reason for the deviations. Discussion of the results can be found in Chapter 4.2. To assure that the results can be reproduced set.seed(1) was used in the case of data presented in this work.

Moreover, simulations were also made to study how moving the mean concentration of the calibration points to a lower value affects the prediction interval at low concentrations (and therefore affects CC_α and CC_β calculated by using approaches suggested by ISO 11843-2, see Table 1 group 7). For this linear homoscedastic data was simulated and prediction interval values at 0 concentration were calculated. Two different sets of calibration levels were used in one of which most of the calibration levels are significantly lower: (1) 0, 200, 400, 600, 1000, and (2) 0, 5, 10, 15, 20, 1000. For given experiment slope and intercept were both set as 5, 4 replicate measurements of each calibration level were made and 5 replicate measurements was given for future sample analysis. These calibration sets were simulated 1000 times and the mean and standard deviation of prediction interval at 0 concentration was found. For the results to be replicable set.seed(1) was set for the script. The script with further details can be seen in Appendix 2.

4. RESULTS AND DISCUSSION

4.1. Subjectivity tests

The results of the interpretations of absolute residual plots by the analysts are presented in Table 3. It can be seen that in both tests some plots were interpreted incorrectly by the majority of the analysts. Also the interpretation of the data between the analysts differs. Therefore, in some cases the analyst should make a conclusion that a definite interpretation cannot be made from the plot and always keep in mind that the interpretation can be subjective. Moreover, it can be seen from the plots (see Figure 4) that linear heteroscedastic data could be interpreted as nonlinear data. However, if the data are nonlinear and strongly heteroscedastic the nonlinearity might be visually insignificant and can be overlooked.

Table 3. Interpretation results of absolute residual plots. The plots are designated with the letters as in Figure 4 and their actual linearity or scedasticity is given. Incorrect interpretations show the number of analysts who interpreted the plot contrary to the true linearity or scedasticity.

Plot	Linearity	Incorrect interpretations	Plot	Scedasticity	Incorrect interpretations
(a)	linear	4/14	(g)	homoscedastic	0/14
(b)	linear	0/14	(h)	heteroscedastic	13/14
(c)	nonlinear	2/14	(i)	homoscedastic	14/14
(d)	linear	11/14	(j)	heteroscedastic	0/14
(e)	nonlinear	0/14	(k)	homoscedastic	6/14
(f)	nonlinear	1/14	(l)	heteroscedastic	3/14

From Figure 5 it can be seen that the analysts often do not agree about the presence of the peak (and therefore whether the analyte is detected) around LoD. Therefore, interpretation of the presence of the peak around LoD is subjective. As a result determining LoD by using this interpretation depends on the analyst. It can be seen from the results that that the differences between the LoD values estimated by (using the visual evaluation approach taking into account all the 4 repeated measurements) the analysts were up to two times. However, it should be noted that (1) the difference was largely determined by the choice of concentrations for the samples (the two times difference comes from the difference of the two consecutive concentrations in the series) and (2) only 4 repeated measurements were evaluated at each concentration level. If more than 4 replicates were to be used the difference between the results might be decreased. As the analyte peak's shape and height changes between measurements, replicates are needed to find LoD. In order to take the repeatability of measurements into account it is possible to use the cut-off approach with visual evaluation. However, it is possible that interpretation of the results could also be made clearer if more data points were to be collected over the same peak. The

MS instrument used in these experiments allowed only approximately 5 to 10 number of data points to be collected per peak (due to relatively long cycle time needed for a measurement in an instrument using an ion trap).

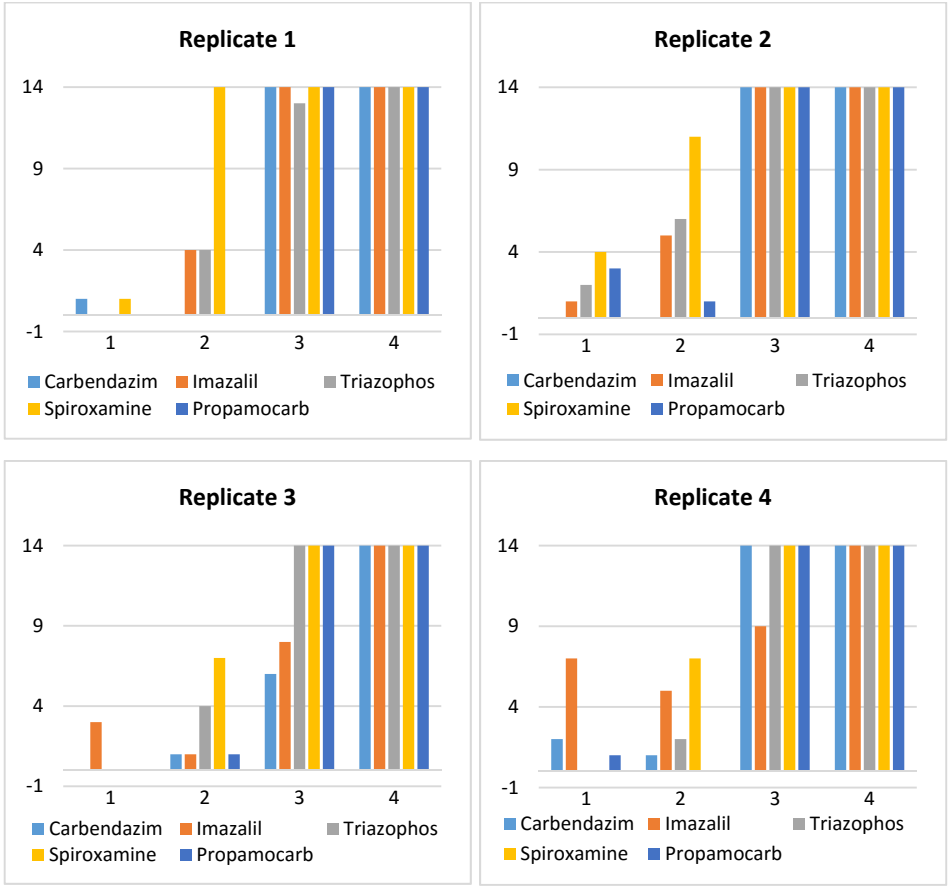


Figure 5. 4 replicate measurements were interpreted at 5 different concentration levels (on x-axis) by 14 analysts (on y-axis). Highest concentration level is not shown – similar to level 4 all analysts agreed that the peak is present at that level. On y-axis it is shown how many of the 14 analysts agreed that an analyte peak is present in that chromatogram. This was done for 5 separate compounds (shown in different colours). The first calibration level has the lowest concentration of analytes. For all compounds the concentration difference between the first and second calibration level was approximately 12 times, between the second and third was approximately 5 times and between the rest approximately 2 times.

In Table 4 the RSD of integration results can be seen. The RSD of results between the analysts integrating a sample, where a peak was not visually present (calibration level 1), was over 60% but only over 6% for a sample with

a clear peak (calibration level 4). The RSD values of analyte peak areas (from different chromatograms of the same sample solution) obtained by one analyst are similar to (when no peak is present) or higher (if the peak is present) than the RSD values of analyte peak areas on the same chromatogram between different analysts. Therefore the difference between analysts when integrating the results is significant only in case of samples with peak intensities close to the noise level.

Table 4. RSD of integration results between all different analysts depending on the calibration level. The concentrations of the calibration levels are similar as explained in Figure 5.

	Calibration level				
	1	2	3	4	5
Propamocarb	64%	49%	11%	3%	1%
Carbendazim	39%	44%	23%	6%	4%
Imazalil	13%	25%	10%	5%	2%
Spiroxamine	34%	26%	7%	2%	1%
Triazophos	14%	20%	3%	2%	1%

4.2. Simulations

The results of simulations can be found in Tables 5 and 6. The parameters given in different columns of the table are given in the caption of the table. The 4 different calibrations given in rows correspond to the different calibration levels are given in Chapter 3.4. From the results it can be seen that even in case all calibration points were below CC_α the $S_{y.x}$ was not significantly affected. However in these cases it is seen that a value for slope that is close to 0 or negative can be obtained which causes unrealistic LoD values. It can be concluded that when using simple LoD estimation approaches (that do not account for the deviation of slope) and low number of repeated measurements the calibration points should be taken at or above CC_β value so that the variance of slope has less random effect on the LoD estimate. Therefore, here it is suggested that at least 3 calibration levels should be used at concentrations where peaks are present in the chromatogram.

It can also be seen from simulations that if standard deviation of intercept is used in place of $S_{y.x}$ the LoD values are always lower. In case the data are strongly heteroscedastic (see Table 6) both of the parameters overestimate LoD and therefore standard deviation of intercept might be somewhat less biased. However, the difference between the two can be small (in practical terms). Moreover, because conservative results are rather needed when using these approaches it is advisable to use $S_{y.x}$.

Table 5. An example result is given (set.seed(1) is used to obtain this simulation data). The concentration levels used in the calibration 1 to 4 can be found in the same as in the script and in Chapter 3.4. The following parameters are brought out in the table: (1) Bcalc – estimated slope, (2) Syx – standard deviation of residuals, (3) Acalc – intercept, (4) Aerror – standard deviation of intercept, (5) Blank – simulated value at 0 concentration, (6) LOD.res – LoD calculated from Syx, and (7) LODint – LoD calculated from Aerror. If "mean" is written in from of the name of the parameter the mean of the 1000 results was found and if "sd" is written in front of the parametes the standard deviation of the 1000 results was found.

Calibration	meanBcalc	sdBcalc	meanSyx	sdSyx	meanAcalc	meanAerror
1	5.01	0.60	97.10	31.25	-1.51	78.56
2	5.03	0.72	93.79	30.44	-3.08	66.90
3	5.02	0.83	95.68	30.58	-0.02	57.16
4	5.09	3.75	96.16	31.79	-0.48	65.52

Calibration	meanBlank	sdBlank	meanLOD.res[a]	sdLOD.res[a]	meanLOD.int[b]	sdLOD.int[b]
1	-1.26	99.46	64.86	22.09	12.79	1.62
2	-1.29	100.09	62.94	22.88	12.83	2.01
3	2.23	97.42	64.76	23.89	12.95	2.33
4	2.00	99.10	285.96	7861.68	38.11	1068.16

[a] LoD calculated by using the standard deviation of residuals (group 4, Table 1). [b] LoD calculated by using the standard deviation of intercept (group 4, Table 1)

Table 6. Example results of simulated heteroscedastic data (set.seed(1) is used). Same denomination is used as in Table 5.

Calibration	meanBcalc	sdBcalc	meanSxy	sdSxy	meanAcalc	meanAerror
1	5.00	0.91	159.21	52.08	-1.38	128.81
2	5.03	1.04	135.01	44.69	-3.81	96.30
3	5.03	1.13	121.82	40.15	-0.08	72.77
4	5.11	4.03	103.42	34.21	-0.70	70.47

Calibration	meanBlank	sdBlank	meanLoD.res	sdLoD.res	meanLoD.int	sdLoD.int
1	-1.26	99.46	108.82	42.60	88.04	34.46
2	-1.29	100.09	93.16	41.25	66.45	29.43
3	2.23	97.42	84.78	37.35	50.64	22.31
4	2.00	99.10	111.86	1544.37	76.22	1052.31

The results from the simulation testing whether the prediction interval changes significantly if the calibration levels are significantly lower were the following: for the first and second calibration set the mean of prediction interval were 14.7 and 13.5 with standard deviations of 2.6 and 2.1 respectively. It can be therefore concluded that although the prediction interval is somewhat narrower in case the mean value of calibration levels is lower but this difference is not significant. Therefore, the choice of calibration levels does not significantly influence the CC_α and CC_β results calculated by ISO 11843-2 if the data is homoscedastic and linear.

4.3. LC-MS/MS

The following two chapters (chapters 4.3.1 and 4.3.2) review different approaches for estimating LoD for analytical methods that need LoD for characterising the method but where scrutinizing the analytical results of specific samples against LoD is not critical. This is the typical situation e.g. determination of most pollutants – if the result is near LoD then it is far below the MRL and consequently the sample is compliant with requirements. If scrutinizing of the results at LoD level is critically important (e.g. in doping analysis) then CC_α and CC_β should be used and this is reviewed in chapter 4.3.3.

In the case of LC-MS/MS the data are likely heteroscedastic and possibly nonlinear (depending on the range of calibration points used). As these analytical method characteristics strongly influence the LoD estimate their evaluation is reviewed here on the example of experimental data. These characteristics are tested separately for data with one (Chapter 4.3.1) and multiple repeated measurements (Chapter 4.3.2) at each calibration points as these cases are somewhat different. The results of these two experimental designs are then also compared.

As discussed above (see Chapter 2.2.3) in case of MS/MS it is possible to observe different fragments for quantification and identification. Our recommendation is the following: LoD should be estimated from the signal intensity (or by using some other approaches, e.g. S/N) of a qualifier ion only (1) if LoD is used for making critical or costly decisions or (2) if LoD does not increase significantly due to using qualifier ion instead of quantifier ion. Otherwise LoD can be estimated using the quantifier ion signal. However, selectivity of the analytical method must be studied thoroughly before LoD estimation when validating the method.

4.3.1. Data with single measurement at each calibration level

Linearity of data with single measurement at each calibration level was estimated by first inspecting the calibration graph and then the absolute residual values. In order to find the linear range the highest concentration values were removed one by one until linearity could be confirmed. It should be noted here that the calibration graph samples should be matrix matched (see chapter 2.1.1).

In all cases the linear range estimated by the absolute residuals is the same or narrower than estimated by visual inspection of the calibration graph. The visual interpretation of calibration graph is therefore used rather as a guide to find the highest concentration from which it is not obvious whether the data are linear or not. Then the absolute residuals are plotted and highest concentrations are removed (and CF parameters recalculated) until linear range is found. An example can be seen in Figure 6 where seemingly linear data in calibration graph are in fact nonlinear according to the absolute residuals plot. It can also be seen from the calibration graph plot that the R^2 value cannot be used to estimate linearity of the data.

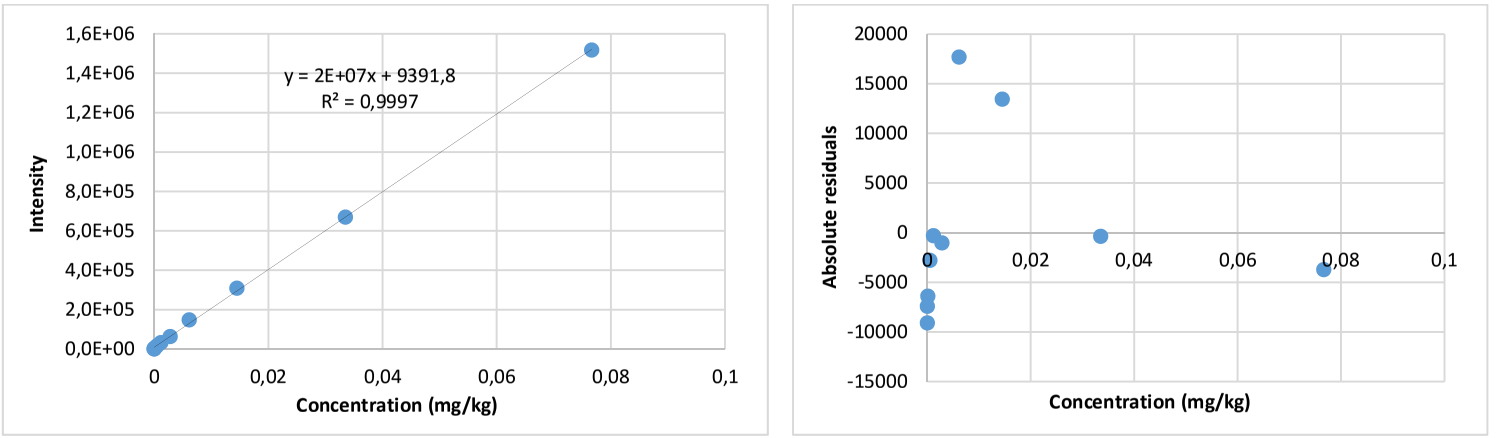


Figure 6. Example of apparently linear calibration data (on the left) that according to absolute residuals analysis (on the right) are not linear.

Relative residuals (these can only be used to estimate linearity if we know that the data are homoscedastic) and Lack-of-Fit test (this can only be used if more than one measurement is made at each calibration level) are not appropriate for the given data. Blank sample and fortified samples with intensities below LoD were kept in the calibration data as CF is not significantly influenced by these points if a sufficient number of data points are above LoD (see chapter 4.2). When integrating peaks of the samples (including blank samples) the guidelines offered in chapter 2.1.4 were followed.

When estimating LoD heteroscedasticity can be taken into account by removing higher calibration levels until homoscedastic range is found. However, in this case removing further calibration levels due to heteroscedasticity can often lead to too few data points (e.g. 3 or 4) for LoD estimation. This is due to the heteroscedastic nature of typical LC-MS/MS data and the wide range of concentrations used in this experimental design. Furthermore, estimating heteroscedasticity for calibration data with single measurements can be untrustworthy (see chapter 4.1).

Although heteroscedasticity is not evaluated it is still important to evaluate linearity of the data. Otherwise the slope and intercept values can be unreliable and not fit for LoD estimation. For example an incorrectly high intercept value can be obtained causing the LoD to become negative. In case this occurs it can be said that the calibration range that was used to estimate LoD was not appropriate. Moreover, if significantly nonlinear data are fitted with a linear model then the error of the model will be large.

OLS was used for estimating the CF parameters. Moreover, the calibration model was not forced through the origin (see chapter 2.1.4).

Next, LoD values were estimated. For the LoD estimation approaches where standard deviations of repeated measurements of blanks or fortified samples are used, the Student's *t* coefficient was chosen as 3.3 and 4.65 (see Table 1). These values correspond to different p values (α and β values) when calculating LoD. However, in some cases (e.g. approach suggested by EPA) the LoD estimates take into account only the false positive results, thereby effectively corresponding to CC_α. Therefore the *t* value is used to choose only the α value (e.g. *t* value of 1.64 and 2.33 leads to CC_α with 95% and 99% confidence against false positive values respectively). If one would need to calculate CC_α and CC_β with 95% confidence level then the multiplier 3.3 (two times 1.64) would be used. The *t* value also depends on the number of degrees of freedom, which in the guidelines is usually assumed as very large.

In order to estimate the LoD via standard deviation of the results of fortified samples (approaches in groups 1, 2 and 3 in Table 1) it is assumed here that the analyst has some prior knowledge of the approximate LoD so that the samples are fortified at a level near LoD. Here the fortified sample results for calculating LoD are taken from samples with lowest concentration at which all 4 repeated measurements still have S/N > 3.

Automatic integration was not used because the used data analysis program (DataAnalysis, LC/MSD Trap software 5.2) does not allow adjusting the inte-

gration parameters. Because only 4 measurements are made at each calibration level the S/N values and visual evaluation results can be unreliable because more repeated measurements at all calibration levels are necessary for these approaches.

Systematically lower LoD was obtained when the standard deviation of intercept was used in place of $S_{y.x}$ (see Figure 7) in the case of Group 4 approaches (Table 1) confirming the results of simulations (see Chapter 4.2). Due to more conservative LoD estimate it is preferable to use the $S_{y.x}$.

It can be seen from the comparison of LoD estimates (Figure 7) that the approach using $S_{y.x}$ gives results similar to the other conservative approaches (e.g. approaches using standard deviation of fortified samples at a single concentration). However, due to strongly heteroscedastic data, if higher concentration levels are used to estimate the standard deviation of fortified samples then the LoD estimate from these results will be higher (e.g. in Figure 7 in the case of propamocarb the LoD value calculated using the $S_{y.x}$ is significantly lower than the LoD estimates from approaches using standard deviation of blank or fortified samples). The LoD estimates calculated from the standard deviation of blank samples are lower also due to heteroscedasticity of data or due to the systematically low variability in integrated blank values. Therefore, it is easy to over- or underestimate LoD when unsuitable concentration level is chosen for fortification in the case of LC-MS/MS analysis methods. Prior knowledge of an approximate LoD value is necessary (e.g. see the LoD estimation approach suggested by EPA).

The approaches using $S_{y.x}$ rely on the assumption of linearity of data because in case of using a linear model on nonlinear data the errors of the model increase the standard deviation of the residuals. It can also be noted that the standard deviation of the mean LoD value is higher than for approaches using the standard deviation of fortified and blank values. In the case of approaches 3 to 6 in Figure 7 only the slope and intercept vary, but in the case of approaches 7 and 8 also the standard deviation varies.

In conclusion, the approach using $S_{y.x}$ (group 4 in Table 1) is recommended for general use for the following reasons: (1) a small number of measurements is needed, (2) little prior knowledge of LoD is necessary, (3) conservative LoD estimates are obtained, (4) the approach suggested in ISO 11843-2 also uses $S_{y.x}$ and therefore the results of these approaches can be readily compared.

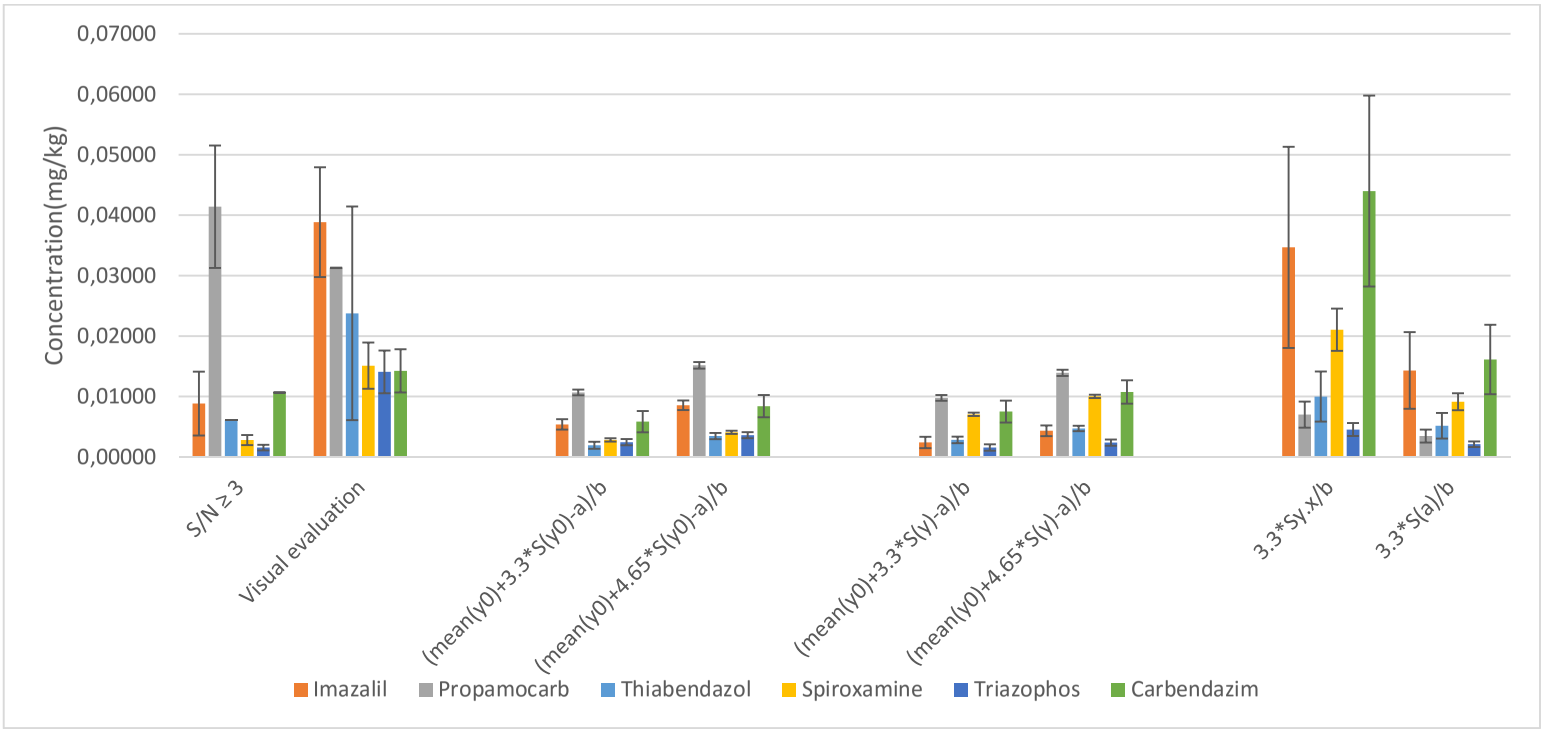


Figure 7. LoD estimates from different approaches (equation for estimation given on x-axis) for the different pesticides (represented by different colours) can be seen. The error bars represent standard deviation of the mean of four replicate LoD estimates within the day. The LoD values (and their corresponding standard deviations of the mean) of imazalil, spiroxamine and triazophos are multiplied by 5, 10 and 20 respectively for better comparability on the graph.

4.3.2. Data with repeated measurements at each calibration level

Linear range was found for data with 4 replicate measurements at each calibration level on 6 separate days. Similarly to the data with single measurement at each calibration level the calibration graph and absolute residuals were used to find the linear range by removing the highest calibration levels one by one until linearity could be confirmed. Lack-of-Fit test was then used to confirm that the linear model found fits the data. Only in two cases (two different compounds on separate days) out of 36 (six analytes on six days) the results of Lack-of-Fit test demonstrated that the linear range had not been correctly found by using the absolute residual plots. When one more calibration level was removed for these cases the Lack-of-Fit test showed that the data are linear. As an example, in Table 7 the highest calibration levels are given for each compound that is found to be in the linear range with the specific test. These data are collected on one single day. It can be seen that in most cases estimating linearity from a visual inspection of calibration graph gives a similar result as absolute residuals. However, in case of spiroxamine the absolute residuals show significantly narrower linear range than visual evaluation of the calibration graph. The Lack-of-Fit test gives either the same or wider linear range than when estimated using the absolute residuals of the calibration graph. It can therefore be concluded that in most cases absolute residual plots can give critical evaluation of linearity of data in case more than one measurement is made at each calibration level.

Table 7. The highest calibration level (mg/kg) of each compound with each linearity test is given that is found to fit the linear range together with all the lower concentration data. The data were collected on a single measurement day.

	Calibration graph	Absolute residuals	Lack-of-Fit
Spiroxamine	0.053	0.010	0.121
Imazalil	0.056	0.056	0.056
Triazophos	0.015	0.015	0.015
Propamocarb	0.843	0.843	1.981
Thiabendazole	0.183	0.183	0.183
Carbendazim	1.522	1.522	1.522

For the following reasons, OLS was used: (1) it avoids the ambiguity of choosing the weighting scheme, (2) in the lower part of the calibration graph an approximately homoscedastic range usually exists and can be used by the majority of LoD estimation approaches, and (3) the LoD approaches used here do not take into account the standard deviation of slope and intercept and it can thus be expected that using WLS will not give significantly different results.

Outliers (significantly deviating results that are not appropriate to be kept in the data) can possibly be identified here by visually inspecting the data on a calibration graph. If a data point clearly deviates more than the other replicate measurements at that level then an outlier can be suspected. These outliers can

easily be left unnoticed on calibration graph with only single measurement. However, data points must not be removed without a sufficient justification – unusual large deviation of data point is not enough to allow removal of the data point. In this work 8 data points out of over 800 were identified as outliers and were removed: 6 due to clear variance of data of all compounds in a single injection from the usual intensity (intensities for all compounds were unusually low) and 2 were removed due to abnormal peak shape.

Contrary to the data used in the previous chapter here heteroscedasticity can be readily observed. Homoscedastic range can be found for LoD estimation by removing the highest calibration levels one by one until it can be confirmed that the data are homoscedastic.

Three approaches to estimate the homoscedastic range of the data were compared: absolute and relative residual plots, and using the Hartley's test. From the absolute residuals plot it was visually observed whether the data are homoscedastic. To estimate the homoscedastic range with relative residuals 3 of the lowest concentration levels were plotted and the highest calibration levels were removed until the observed relative residuals did not significantly deviate from 0 towards positive or negative values. Hartley's test (see description in chapter 2.1.3.2) was applied to the data and the highest concentration levels were removed until an F value below the critical level was obtained. From the results (see Table 8) it can be seen that the relative residual plots allow higher concentration levels to be left into the data than the other tests. This is because the relative residuals do not take into account the standard deviation of the results and it is only observed whether the regression line passes close to the data of the lowest calibration levels. However, in absolute residuals the standard deviation is indirectly observed by visual evaluation and in Hearty's test standard deviation is used in calculations. Therefore relative residual plots overestimate the homoscedastic range and are not recommended for use.

Table 8. The highest concentration (μg/kg) of the homoscedastic range estimated by the respective approach.

	Absolute residuals	Relative residuals	Hartley test
Spiroxamine	0.35	10	0.35
Imazalil	2.0	23	2.0
Triazophos	0.52	6.1	2.8
Propamocarb	31	840	72
Thiabendazole	33	77	33
Carbendazim	0.16	670	1.9

The Hartley's test is a general test for assessing whether groups of data (in this case: sets of replicate measurement results at different calibration levels) all have the same variance. This test therefore does not take into account that in a calibration graph the growth direction of variance is known. As a result, it is

seen in the data that due to low number of replicate measurements at each calibration level it can happen by random chance that higher calibration levels have lower variance than the results from blank samples (see Table 9). The following is therefore suggested: if the standard deviation of a calibration level is statistically significantly lower than that for the lowest calibration level (in this case: the blank value) this level should not be taken into account when estimating the homoscedastic range. This is because this level is most likely in the homoscedastic range.

Table 9. The square of standard deviation at 0.0086 µg/kg (highlighted with bold) is significantly lower than for blank samples. As a result when Hartley's test is carried out homoscedastic range cannot be found. However, when removing the standard deviation results of the 0.0086 µg/kg sample a homoscedastic range can be found from blank samples to samples with range from blank samples to samples with concentration of 2.89 µg/kg.

C (µg/kg)	0	0.0086	0.104	0.535	1.23	2.89	6.33	14.7
$S(y)^2$	556835	**7852**	908155	4892298	1917186	9604625	60769665	147550168

In spite of these difficulties and that it is somewhat more complex to use the Hartley's test, this test is to be preferred over the use of absolute residuals as the latter can be subjective. Furthermore, Hartley's test is more conservative when analysing the data: the estimated homoscedastic range is always the same or narrower compared to one estimated by absolute residuals (see Table 8). This is important because the collected experimental data are strongly heteroscedastic and in most cases only 4 or less calibration levels are in the homoscedastic range. As these data include the blank solutions and solutions with low analyte concentration level (where the peak is not clearly present) the slope value estimated from these data has large variance and is therefore not reliable (e.g. if only calibration levels are used where no peak is detected then the slope value is 0). Therefore, for further LoD calculations the results from Hartley's test were used so that LoD could be estimated from a homoscedastic range. The strong heteroscedasticity was not caused only by the nature of the LC-MS/MS technique but also by the wide range of concentrations of the calibration levels. It is therefore concluded again that it is useful to approximately know the range where LoD is so that more concentration levels would fall in the homoscedastic range.

It must be noted here that evaluating a concentration range as homoscedastic does not necessarily mean that this range is insignificantly heteroscedastic in the case of possible future datasets in the same concentration range. This means the range is evaluated as homoscedastic only because the used statistical test do not have the statistical power to distinguish heteroscedasticity. It is still reasonable to assume that this range is homoscedastic as the influence of heteroscedasticity in this range is less important than in wider range. However, it can be concluded that the LoD values estimated in the previous chapter from data with a single

point at each calibration level may be overestimated as heteroscedasticity was not properly taken into account.

Because the analytical methods under consideration in this chapter do not need an LoD estimate where all assumptions and simplifications are taken into account and rather need a conservative LoD estimate that can be with high probability routinely achieved by the analytical method the following recommendation is made. At least 5 calibration levels (of which 3 should clearly have the analyte peak present, i.e. S/N > 3) should be used for LoD estimation. If a smaller number of calibration levels are found in the homoscedastic range then a new experimental design to estimate LoD with a different calibration range should be made. If this is not possible then the lowest calibration levels left out due to heteroscedasticity should be added to the data until the above requirements are met. If possible the 3 calibration levels with analyte peak present should be close to LoD, e.g. the concentrations of the calibration levels should not differ from LoD by more than an order of magnitude. However this last suggestion is a general guideline as linearity and heteroscedasticity are the parameters that determine the range that could be used. Using calibration levels in the heteroscedastic range means that the LoD values can be somewhat overestimated and therefore conservative.

The LoD values estimated by different approaches are compared in Figure 8. Calculations by all the approaches assume that future samples will be analysed only once (standard deviation is used in place of standard deviation of the mean and the intensity values for calibration levels are not averaged before calculation of $S_{y.x}$ and of the intercept). In general similar trends can be seen here as with LoD results where data contained only one measurement for each calibration level: the LoD values are lower when calculated from blank standard deviations, similar results are obtained for approaches that use $S_{y.x}$ and of fortified samples, and the standard deviation of intercept provides a lower LoD estimate than residuals. Although, here 4 replicate measurements could be used for the cut-off approach using S/N and visual evaluation, this is still too few for obtaining reliable results and significantly below the usually required 10 measurements (see Table 1). Therefore, in this case no significant difference exists between used approaches if single measurement or if replicate measurements at each calibration level are made.

LoD estimate was also calculated using the Eurachem (group 3, Table 1) approach where it is taken into account that the sample is measured repeatedly. This was calculated from fortified samples and 4 replicate measurements were assumed for the future samples. The LoD estimate has a similar value to the ones calculated from the standard deviation of blank values.

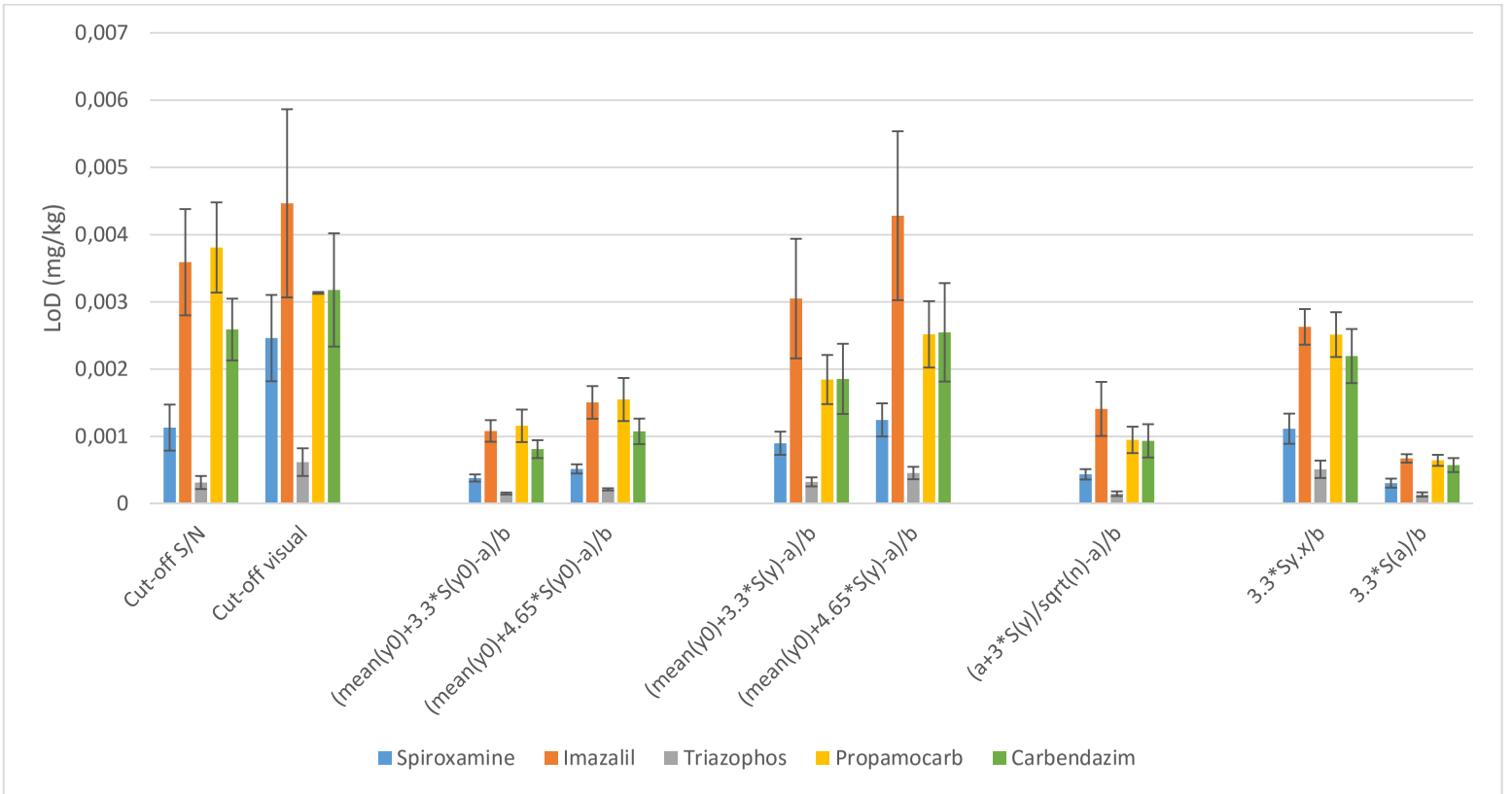


Figure 8. LOD estimates of different approaches (equation for estimation are given on x-axis) for 5 different pesticides (in different colours). The error bars represent standard deviation of the mean of six replicate LoD estimates between the days. The LoD values and standard deviation of mean values for propamocarb and carbendazim are divided by 10 and 5 respectively for better comparability on the graph.

It should also be noted that the difference between the LoD values from different approaches can be (depending on the compound) only 2 times. This difference can be statistically significant but might not be important for the given application. However, it is still suggested to use the results of more conservative approaches.

The results were compared to the data used in the previous chapter to identify whether making replicate measurements at each calibration level significantly changes the results and therefore whether it is useful. For this the paired t-test was performed as in this case the variation due to changes between days does not influence the results. The test is done for each compound and each LoD estimation approach separately. It must be noted that for the approaches using standard deviation of blank and fortified values only the slope and intercept values change between the single and multiple calibration level measurement data and not the standard deviation as the same data is used in these cases. The data indicate that in most cases the LoD values obtained using the same approach with and without replicates of calibration points are not significantly different (see Table 10). Only one compound the results are significantly different for multiple LoD approaches. The main difference between these results can come from the fact that the linear range and homoscedastic range are evaluated differently leading to the use of different data for LoD estimation.

Table 10. Results of paired t-test comparing LoD values in case only one measurement was used at each calibration level in comparison to results when 4 replicates at each calibration level were used to estimate LoD. The critical t value is 2.57. The t values higher than the critical level (i.e. results are significantly different) are highlighted.

	t value					
Approach	Spirox-amine	Imazalil	Triazo-phos	Propamo-carb	Thiabenda-zole	Carben-dazim
S/N ≥ 3	0.26	0.10	1.99	0.60	1.58	1.00
Visual evaluation	0.90	1.12	0.27	0.00	1.82	0.45
(mean(y0)+3.3*S(y0)-a)/b	0.25	0.37	1.10	1.70	***3.65***	0.29
(mean(y0)+ 4.65*S(y0)-a)/b	0.49	1.04	0.93	1.24	1.35	1.26
(mean(y0)+ 3.3*S(y)-a)/b	0.44	0.12	1.27	1.90	***3.51***	0.15
(mean(y0)+4.65*S(y)-a)/b	1.14	1.31	0.94	0.59	***2.62***	1.22
3.3*Sxy/b	1.82	1.54	1.46	0.38	***4.75***	2.56
3.3*S(a)/b	2.16	1.73	1.61	2.41	***2.66***	***2.91***

In conclusion, if only a simple (i.e. not for critical decisions) LoD estimate is necessary then a single measurement at each calibration level is sufficient. Only if more reliable LoD estimates are necessary (e.g. for a thorough validation of an analytical method for publication of results) then performing more than one measurement at each calibration point increases the reliability of the LoD estimate and decreases its variability of LoD estimates within a day.

A comparison was also made between LoD results when using the mean value of blank samples and the intercept value for approaches where standard

deviation at a single concentration is used to estimate LoD (Table 1, group 1). These parameters are used in these approaches to estimate the LoD in the intensity scale after which the corresponding concentration can be found (using the intercept and slope of the CF). It can be seen from Figure 9 that the results between the approaches when using the mean value of blank samples or the intercept is not significantly different. Therefore it is recommended here that blank LC-MS/MS results should be integrated similarly to chromatograms with peaks and the data should be used to estimate the CF. This is important because not integrating blanks and low level samples (and using the intercept value estimated from higher concentration samples) will result in losing information at low concentrations.

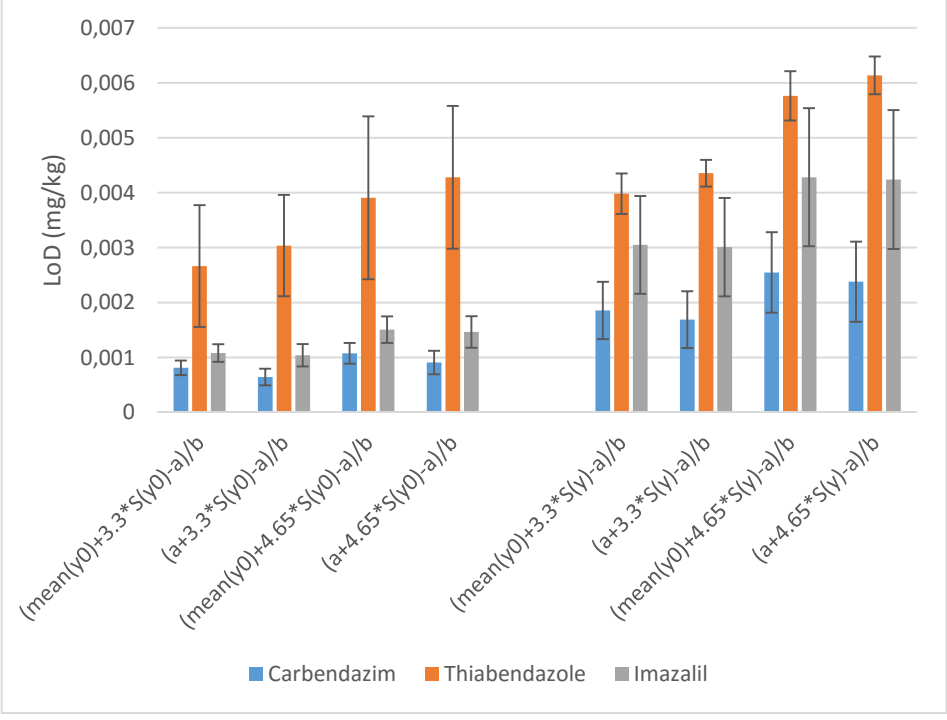


Figure 9. LoD estimates of different approaches (equation for estimation given on x-axis) for 3 different compounds (in different colours). The error bars represent standard deviation of the mean of six replicate LoD estimates between the days.

4.3.3. CC_α and CC_β estimation approaches

If an analytical method is used that demands a reliable parameter for interpretation of results in terms of whether the analyte is detected or not and for characterization of the method then it is recommended to use CC_α and CC_β. Here the approaches suggested in the guideline ISO 11843-2 [4] are discussed and the obtained CC_α and CC_β values are compared. The data collected for these

experiments follow the specific requirements given in that guideline. ISO 11843-2 allows calculation of CC_α and CC_β for homo- and heteroscedastic data with separate approaches and these results are compared. The approach that takes into account the heteroscedasticity of the data assumes that the standard deviation is linearly dependant on concentration. Moreover, the estimation procedure for this approach uses WLS in place of OLS [4]. As the CF uses multiple independent calibration solution preparations (in this case 2) separately (i.e. not averaged) for each calibration level then it is assumed that the same number of independent replicate sample preparations and measurements will be done for the future samples. To the authors' knowledge these approaches are the most sophisticated approaches recommended in guidelines and take into account considerations that other LoD approaches do not (see Table 1). In addition, these approaches are quite widely used.

In comparison to these approaches the simple approaches for estimating CC_α and CC_β (group 6 in Table 1) can be recommended only when high reliability is not required, because they make similar assumptions to the LoD estimation approaches reviewed in the previous chapters. Keeping in mind that CC_α and CC_β are usually used for making critical decisions, these simplified approaches are not discussed here.

The data range found in the previous chapter is used for calculations: (1) for the approach assuming homoscedasticity the homoscedastic and linear range is used, and (2) for the approach that assumes heteroscedasticity the linear range is used. Ideally in the first case the data is homoscedastic, but can in practice be somewhat heteroscedastic as some higher calibration level data must in some cases be added so that enough calibration levels could be used. Figure 10 presents the comparison of the two ISO 11843-2 approaches and the ICH approach using the $S_{y.x}$ (group 4, Table 1). From paired t-test (similar to test used in chapter 4.3.2) of the results it is seen that for the given data the two approaches suggested by ISO 11843-2 do not give statistically different CC_α and CC_β ($p = 0.05$). The results of the t-test can be seen in Table 11. Although significantly more complex the approach that takes heteroscedasticity into account is more appropriate due to possible heteroscedasticity of the data used here. Moreover, the LoD value estimated from $S_{y.x}$ is significantly different from the estimated CC_β value only for one compound. This can be due to the fact that the most important assumptions on linearity and scedasticity have been taken into account in both approaches. Therefore fairly accurate interpretation can be done with this LoD approach. However, in case reliable interpretation of results is needed the ISO 11843-2 approach is still preferable.

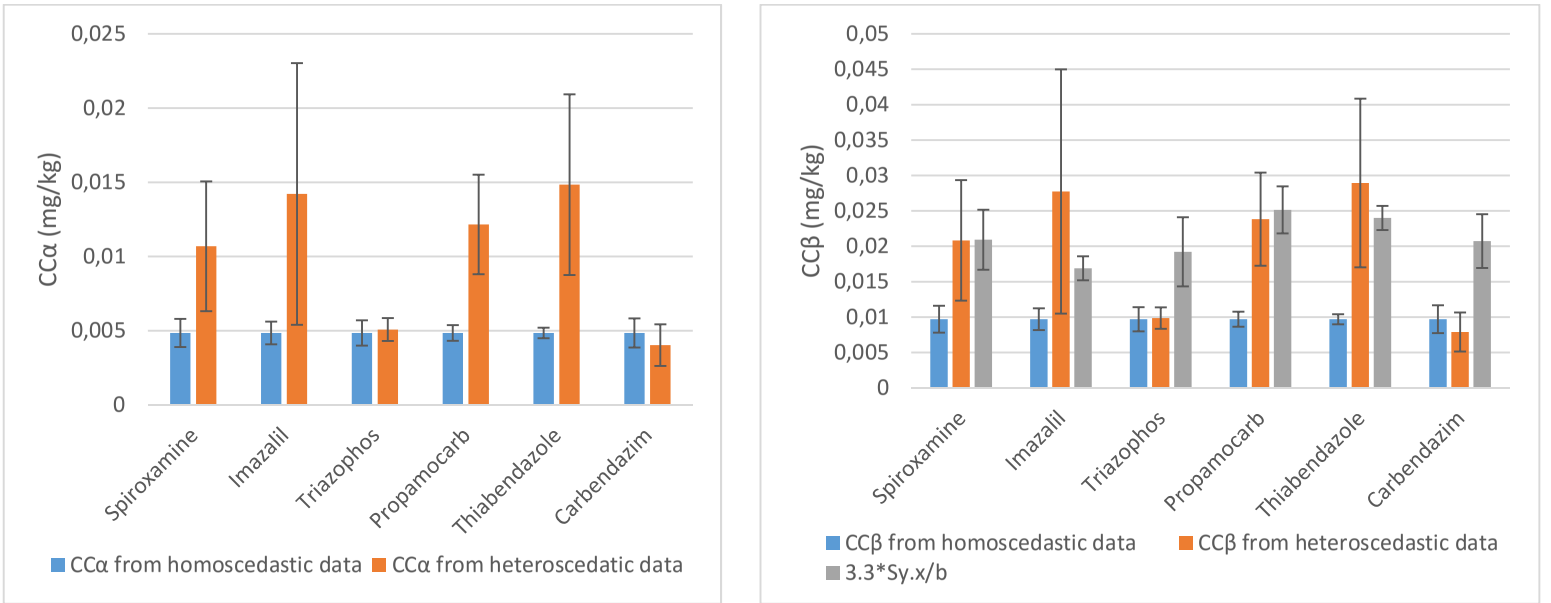


Figure 10. The ISO 11843-2 approaches to estimate CC_{α} are compared (on the left) and the CC_{β} estimation approaches are compared to the approach using $S_{y.x}$ (group 4, Table 1) (on the right). In both graphs the results are normalised to the propamocarb value obtained using the ISO 11843-2 approach for homoscedastic data. The LoD is estimated with the assumption of homoscedastic data with repeated measurements at each calibration level. In this graph the average LoD of the between-days results is compared to CC_{β} values. The error bars show the standard deviation of the mean.

Table 11. Three different pairs of estimates are compared to each other: (1) CC_α estimates from the two different approaches given in ISO 11843-2, (2) CC_β estimates from the two different approaches given in ISO 11843-2, and (3) the CC_β value from the ISO 11843-2 approach that uses heteroscedastic data and WLS is compared to the approach using $S_{y.x}$ value (Table 1, group 4). These comparisons are numbered in the rows of the Table similarly.

	t value					
	Spiroxamine	Imazalil	Triazophos	Propamo-carb	Thiabenda-zole	Carben-dazim
Comparison 1	1.35	1.14	0.17	2.22	1.63	0.44
Comparison 2	1.32	1.12	0.06	2.19	1.61	0.49
Comparison 3	0.01	0.66	1.78	0.18	0.42	2.89

4.3.4. Between-days LoD

The LC-MS/MS system parameters may significantly vary between days. Thus, the variability of LoD between days should be tested and between-days LoD should be used if necessary (see Chapter 2.1.5).

In order to estimate whether LoD changes significantly between days more than one estimate of LoD is necessary on each day. Therefore 4 separate CF-s were obtained with a single calibration point at each calibration level for obtaining 4 independent LoD estimates within a day.

With the data collected in the experiments, only the approaches in group 4 Table 1 that uses $S_{y.x}$ or standard deviation of intercept can be used for finding several LoD estimates within one day (there are not enough data to use other approaches). Here only the approach using $S_{y.x}$ was used. LoD values were calculated with the approach from 4 different calibration graphs on each 6 days (altogether 24 LoD values) for each compound. The same data was used as in chapter 4.3.1 where single measurement was made at each calibration level. ANOVA (see general description of ANOVA in chapter 2.1.3.1) was performed on these data: the changing factor was taken to be time (meaning different days) and it was tested whether the variance of LoD results is significantly larger between days than within a day.

The results show that the between-day variance of LoD estimates is similar to within-day variance (Table 12). Therefore, from these data it cannot be said that LoD significantly differs between days. In this case it is therefore appropriate to use the between-days LoD estimated because the LoD does not significantly differ between days and using a between-days LoD (calculated from data collected in many days) gives a more accurate representation of the methods capabilities. As discussed in chapter 2.1.5 if it would have been shown that the LoD significantly changes between days then the between-days LoD might not be best for use when interpreting result. This is so because in that case we would ignore the conditions of that specific day that significantly influences the LoD.

Table 12. Here the F values for different compounds calculated by ANOVA are given. The critical F value for one-sided F test at p = 0.05 for this ANOVA (accounting for degrees of freedom) is 2.773 [28]. It can be seen that for all the compounds the F value is below the critical value and therefore the between-days variance is not significantly larger than within day variance (meaning LoD between days does not change significantly compared to random change of LoD within a day).

	F
Spiroxamine	0.85
Imazalil	0.82
Triazophos	1.73
Propamocarb	1.23
Thiabendazole	0.65
Carbendazim	0.89

In this work it is recommended to estimate the between-days LoD using the following equation:

$$LoD_{btw-day} = mean(LoD) + 1.65 \times S(LoD) \quad (8)$$

where mean(LoD) is the mean value of the LoD estimates found on different days, S(LoD) is the standard deviation of LoD estimates from different days and 1.65 is the one-sided Student's *t* value for 95% (assuming a large number of measurements, more accurate *t* values can be chosen based on the degrees of freedom). This estimate should then be used for both interpretation of results and characterization of the analysis method.

Based on equation 8 between-days LoD values for different approaches were calculated from pesticide measurement data where repeated measurements at each calibration level (the data from which calculations are made can be seen in Figure 8). The results are presented in Figure 11.

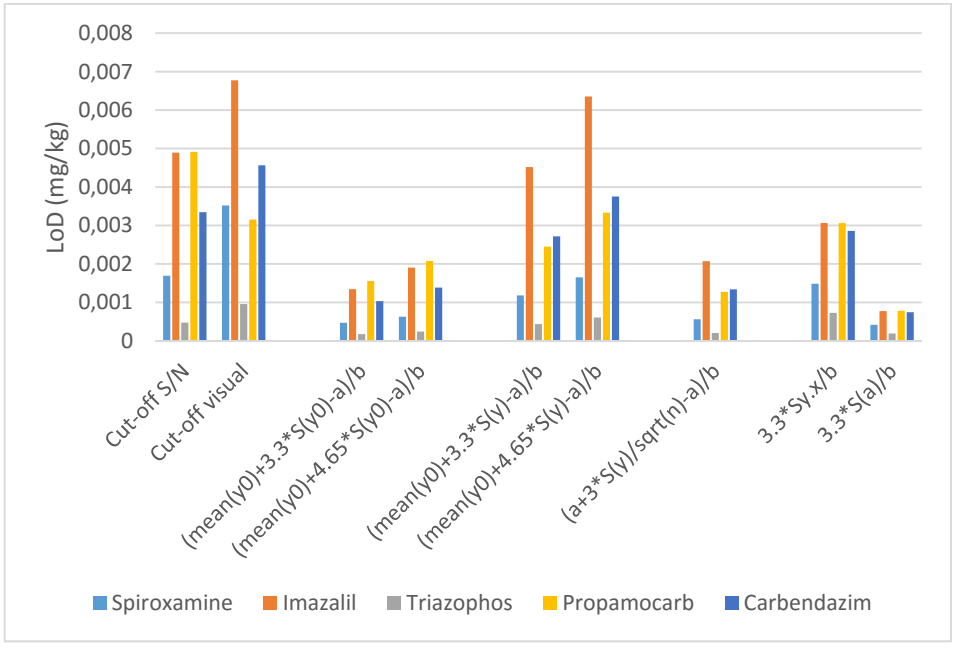


Figure 11. Between-days LoD values calculated for different compounds and approaches. The LoD values and standard deviation of mean values for propamocarb and carbendazim are divided by 10 and 5 respectively for better comparability on the graph.

4.3.5. Comparing LoD values of different LC-MS/MS methods

An important use of LoD estimates is comparison of analytical methods. It was therefore important to test whether the same conclusions (which are presented above on the basis of the analysis method for measuring pesticides in tomato) about the LoD estimation approaches can be made when using other LC-MS/MS methods.

For this 3 LC-MS/MS methods were compared: (1) determination of pesticides in tomato, (2) determination of antibiotics in blood, and (3) determination of antibiotics in blood using a different ion source (3R). The 3R source has an additional capillary that directs additional nebulization gas to the tip (see Chapter 3.1.1) and has been shown to give lower LoD values as compared to the classical nebulizer design [78].

The LoD values used for this comparison are estimated for pesticides from the data with more than one replicate measurement at each calibration level and the data used for antibiotics is described in chapter 3.1.1. The linear and homoscedastic range for the antibiotics data were found using absolute residual plots. However, similarly to pesticide data, due to the wide concentration range, additional concentration levels were required in the calibration data to meet the requirement that at least 5 calibration points must be included in the data of which at least 3 levels clearly have the analyte peak present in the chromatogram.

Figure 12 demonstrates that different approaches give rather consistent results with the analytical methods. It can also be seen that that the 3R nebulizer gives consistently lower LoD values than the commercial ESI source. This can be due to its better nebulization ability, leading to higher ionization efficiency and therefore lower LoD values [78].

Different analytical methods should not be compared to each other in case the LoD values are estimated with different approaches. This is why a clear description of the approach used to estimate LoD (and the concentration units of the LoD) should be given with the result. This also means that if a lab obtains (for a certain method) a higher LoD than reported for the same method in the literature, then the information on the LoD estimation approach has to be taken into account. The conclusion can often be that the lab's method is not necessarily inferior to the one reported in the literature.

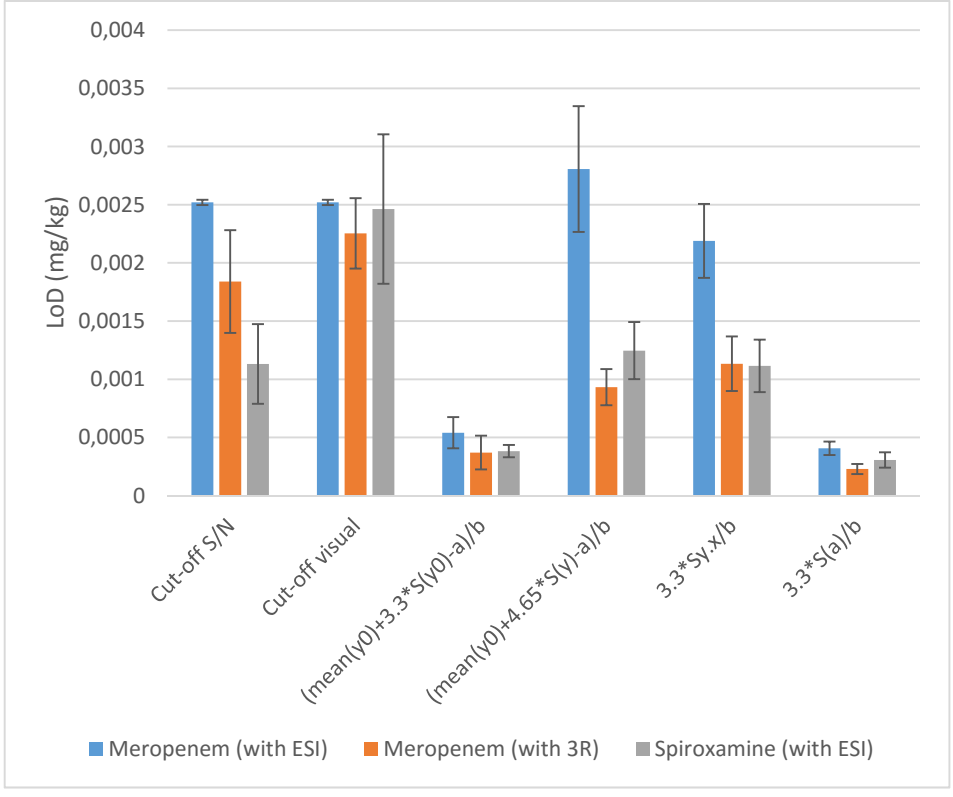
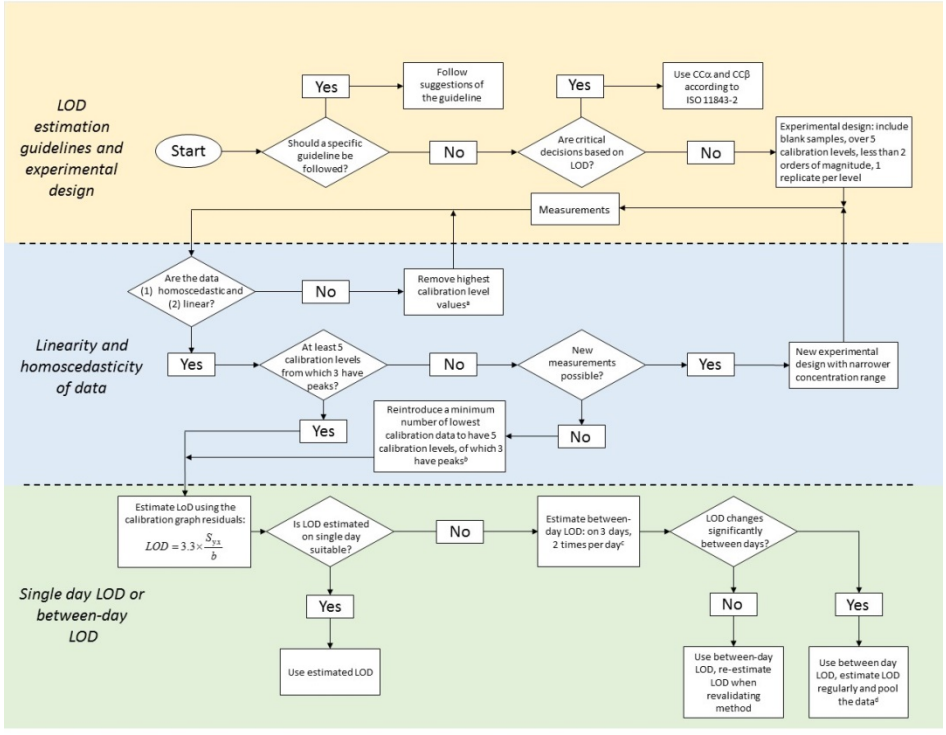


Figure 12. LoD estimates of different approaches (equation for estimation given on x-axis) for different compounds estimated with different analytical methods (in different colours) can be seen. The error bars represent standard deviation of the mean of replicate LoD estimates between the days.

4.3.6. Conclusion of LC-MS results

As a summary, the steps of estimation of LoD, formulated as a general flow-chart is presented in Figure 13. It takes into account the experimental results, discussed in the previous chapters. In this flowchart the LoD estimation is divided into 3 parts. All steps for estimating LoD found to be important in this work are included.



[a] See chapter 3.3.2 on removing the highest calibration levels.
[b] The resulting LOD can otherwise be over- or underestimated.
[c] All the steps in the above flowchart must be followed each time.
[d] The data used to estimate between-day LOD should be representative of the method's capabilities at the moment, i.e. irrelevant data can be left out of pooling.

Figure 13. Flowchart (decision tree) for estimating LoD using suggestions made in this work.

4.4. PSI

The fast screening analysis method for five pesticides – thiabendazole, aldicarb, imazalil, methomyl and methiocarb – in oranges, grapes and tomatoes was developed [Paper I] and the LoD of these pesticides were evaluated to be below 5 mg/kg on the basis of S/N values (see Table 13). Further details of these experiments can be found in Paper I.

Table 13. S/N values calculated from PSI/MS/MS measurement results of homogenized tomatoes spiked with 5 mg/kg of pesticides in the multiple reaction monitoring mode. The noise was measured from the blank homogenized samples. The MRL values of the pesticides for tomatoes are shown in mg/kg.

	Methomyl	Thiabendazole	Aldicarb	Methiocarb	Imazalil
S/N value	522.1	81.1	5.0	106.1	173.3
MRL (mg/kg)	0.02	0.05	0.02	0.2	5

It must be noted, however, that this method (even with IS) is not robust enough for routine analysis because in some cases no signal (of analyte or IS) is received from the measurement. This poor robustness can be explained by matrix effects that strongly depend on the portion of the homogenized tomato sample that is pipetted onto the PSI paper. As a result in the cases when very low signal is received from the IS the data cannot be used in further calculations. The number of measurements among the data that cannot be used for this reason is about 20%. As the decision whether the IS signal is present or not was made by the analyst the standard deviation of results is rather subjective.

As explained in chapter 3.2 contamination of the MS was regularly checked and the MS was cleaned after each measurement, so that contamination of the MS can be ruled out as a reason.

In Figure 14 a typical chronogram of IS and analyte can be seen. As soon as the eluent is added to the paper and the eluent reaches the paper triangle tip the signal is created. However, a visual spray (an example can be found in Figure 15) can often be seen for more than 20 second which is significantly longer than the strong signal from analyte. This decrease in signal strength could be explained by the gradual increase of ionization suppression due to more matrix eluting to the tip of the paper [79].

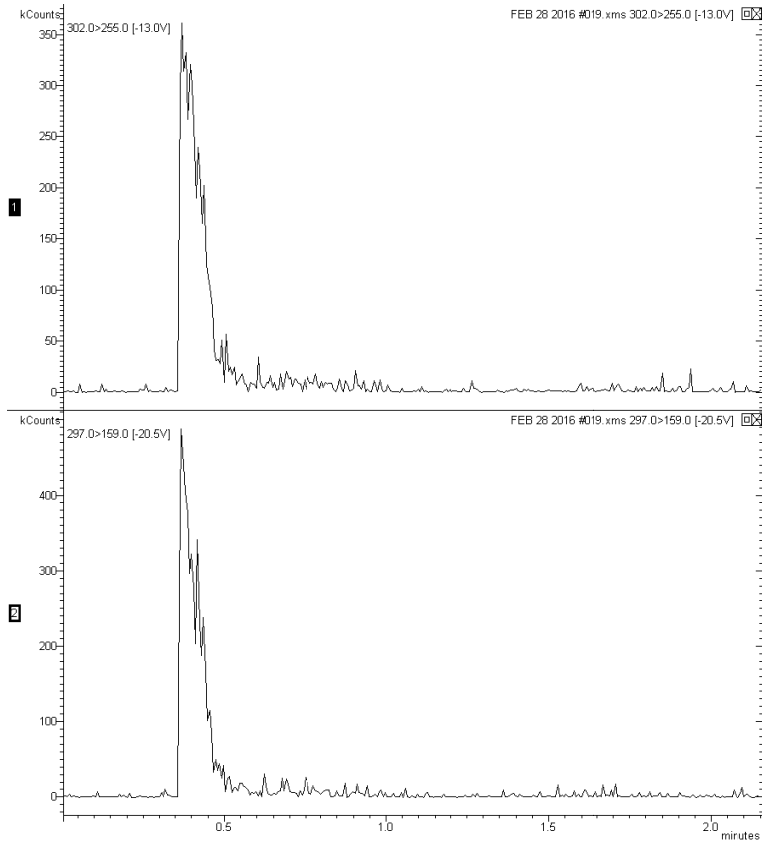


Figure 14. Here chronograms of imazalil (bottom chronogram) and the IS (top chronogram) collected with PSI/MS/MS can be seen.

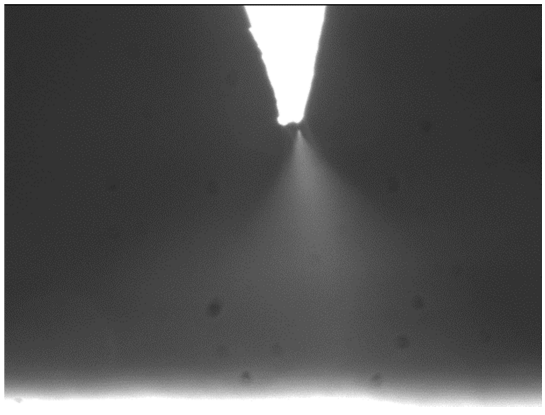


Figure 15. After eluent is added and its front reaches the tip of the paper visual spray is initiated. This picture is taken by a microscope attached to the nanoESI housing. The width of this picture is approximately 5 mm.

Three of the lowest concentration levels were left out of the calculations as the analyte signal was not observed (signal was difficult to distinguish from noise) and they were considered below LoD. However, data from blank sample were left in the calibration data. The linearity of the data was verified using absolute residuals. Heteroscedasticity could not be confirmed from the absolute residual plots. F-test (Table 14) was performed on the replicate measurement data collected from blank sample and fortified sample at 0.58 mg/kg. From this data it can be concluded that between these two calibration points the data can be treated as homoscedastic.

Table 14. F-test results of 3 separate days performed on replicate measurement data of blank sample and sample fortified at 0.58 mg/kg concentration. 5 replicate measurements were performed for both samples. The standard deviation was calculated for the values of analyte intensity divided by IS intensity. The critical value of this F-test is 9.605 (two-tailed, p = 0.05) [28].

	S(blank)	S(fortified)	F
Day 1	0.078	0.068	1.32
Day 2	0.071	0.038	3.47
Day 3	0.057	0.054	1.10

The LoD values were first calculated in the scale of analyte and IS intensity ratio. After this the calibration function was used to estimate the value in analyte and IS concentration ratio scale. The standard deviations and mean blank values of measurements used in the calculations are therefore also in the scale of analyte and IS intensity ratio. The LoD was then estimated in analyte concentration scale by multiplying the average IS concentration in the sample used in the calculations with the LoD value found in the scale of analyte and IS concentration ratio.

The LoD estimates calculated by different approaches can be seen in Figure 16. It can be seen that the approaches that use standard deviation values at single concentration can give somewhat higher LoD values that when $S_{y.x}$ is used. However, this difference is generally not significant when the variance of the LoD values between days is taken into account. Therefore, here we suggest using the latter LoD estimation approach as it is simpler and requires less measurements. Because critical decisions are not made with screening methods then using CC_α and CC_β is not necessary here as these can be overly complex and time consuming for estimation. However, the LoD should be significantly below the MRL level for the analysis method to be useful. In case of imazalil in tomatoes the MRL is 0.5 mg/kg [80] and therefore the LoD value of the method is below the MRL. Therefore this approach could potentially be used as a screening method.

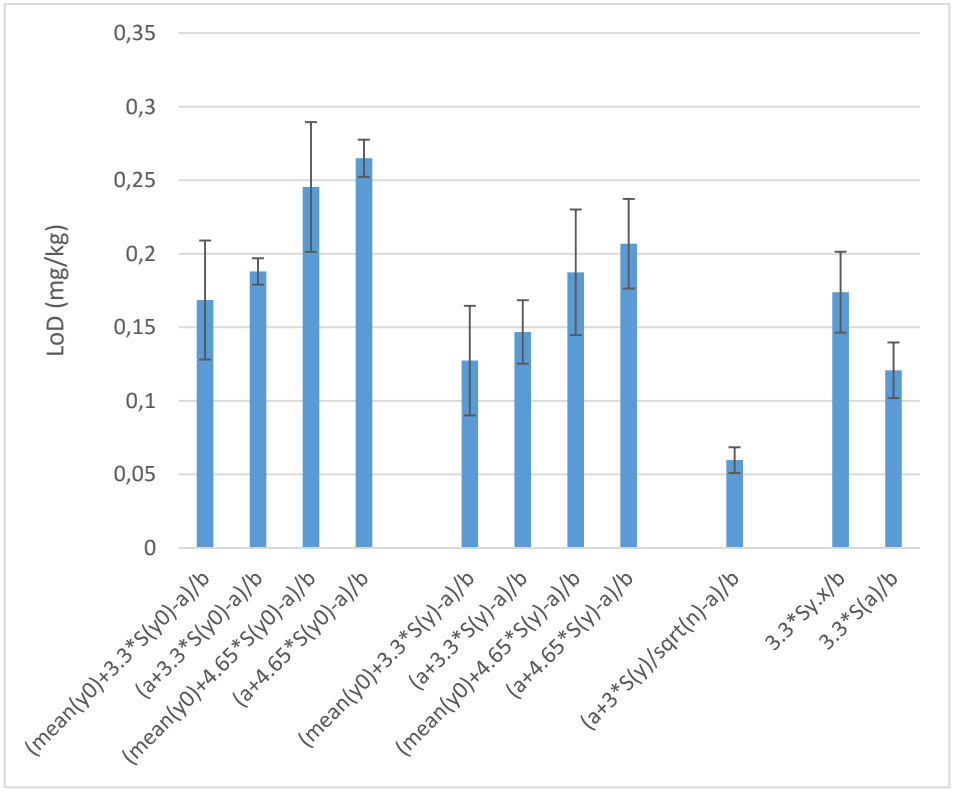


Figure 16. LoD estimates of different approaches for PSI/MS/MS method (equation for estimation are given on x-axis). The error bars represent standard deviation of the mean of 3 replicate LoD estimates obtained on different days.

5. SUMMARY

The aim of this work was to study the estimation of LoD for analytical techniques using MS. LC-ESI/MS and PSI/MS (as a screening methods) were used for this. Although the techniques used in MS can be somewhat different (e.g. using an IS addition method in case of PSI) the same general suggestions can be made for both techniques.

When estimating LoD it must be first considered whether LoD or CC_α and CC_β need to be determined for the particular analytical method. In case a certain guideline must be followed then the suggestions in that specific guideline should be used. However, guidelines often leave important decisions to the analyst. In this case the suggestions and discussion in this work can be helpful.

Experiments were made so that different LoD estimation approaches could be applied to the data and the results could be compared. The ICH approach (Table 1, group 4) that uses the $S_{y.x}$ is the approach recommended for estimating LoD in this work. This approach is suggested because it is simple to use, gives conservative results, and the results are reasonably similar to the CC_β estimates as defined by ISO 11843-2. However, the obtained LoD is strongly influenced by linearity and scedasticity of the used data and therefore before LoD estimation the linear and homoscedastic range should be found. Based on experimental results and simulations the following rule is suggested for choosing the calibration levels: at least 5 calibration levels (one measurement per calibration level) should be included in the used data of which 3 should have a well-defined signal (e.g. S/N > 3). If estimating CC_α and CC_β is required then the ISO 11843-2 approach is recommended.

For all LoD determination approaches it is useful to know the approximate LoD beforehand, so that appropriate concentrations could be chosen for the measurements. This is especially true for the approaches that use repeatability at a single concentration and are strongly influenced by the choice of the fortification level if the analytical method is heteroscedastic.

Suggestions have been made how to take into account the variability of LoD estimates between days. For this a new parameter called between-days LoD has been suggested and its use described. Moreover, for more clear comparison of different analytical methods the between-labs LoD has been suggested.

It was also shown that visual evaluation of LoD from the chromatogram, and estimation of linearity and scedasticity from residuals can be subjective (dependant on the analyst). Therefore, caution is suggested when using these approaches.

In the course of this work it has become apparent that this topic needs further and more detailed attention (e.g. taking into account *a priori* knowledge and use of tolerance interval when estimating LoD). However, the goals set for this work have been met and hopefully this work will motivate further research on this important topic.

6. SUMMARY IN ESTONIAN

Massispektromeetriliste analüüsimeetodite avastamispiiri hindamine

Käesoleva töö eesmärk oli avastamispiiri (LoD) hindamisvõimaluste uurimine MS meetodite puhul. Selle uurimuse jaoks kasutati LC-ESI/MS analüüsimeetodit ja PSI/MS põhinevat sõelmeetodit. Töö käigus tehtud üldised soovitused sobivad mõlemale MS meetodile kuigi need meetodid on mõnevõrra erinevad (näiteks sisestandardi kasutamine PSI puhul).

LoD hindamisel tuleb esimesena kaaluda, kas vastava analüütilise meetodi jaoks oleks sobivam kasutada LoD-d või otsustuspiiri (CC_α) ja avastamisvõimet (CC_β). Kui on kohustuslik järgida mõne spetsiifilise juhendi juhiseid, siis tuleb kasutada ka vastavaid spetsiifilisi juhiseid. Tihti aga jäetakse mitmed olulised otsused siiski juhendites lahtiseks ning seega tuleb need teha analüüsijal. Sellistel juhtudel sobib järgida soovitusi, mis on antud selles töös.

Selles töö käigus tehtud eksperimentide tulemusena soovitame kasutada LoD hindamise metoodikat, mida pakub ICH ja mis kasutab jääkliikmete standardhälvet LoD hindamiseks (Tabel 1, grupp 4). Seda metoodikat soovitame, kuna selle kasutamine on lihtne, see annab konservatiivseid (pigem kõrgeid kui madalaid) tulemusi ja tulemused on sarnased CC_β tulemustele, mis on hinnatud ISO 11843-2 soovituste järgi. See LoD väärtus on aga tugevasti mõjutatud kasutatud andmete lineaarsusest ja skedastilisusest ning seega tuleb enne LoD määramist hinnata, kas kasutatud andmed on lineaarses alas ning homoskedastilised. Tuginedes eksperimentaalsetele tulemustele ja simulatsioonidele anname siinkohal järgneva soovituse kalibratsiooniandmete tasemete valimiseks: andmed peaksid sisaldama vähemalt 5 kalibratsioonitaset (üks mõõtmine igal tasemel), kusjuures vähemalt kolmel tasemel peaks olema selgelt eristuv signaal (näiteks $S/N > 3$). Kui vastava analüüsimeetodi jaoks on vaja kasutada CC_α ja CC_β väärtusi, siis soovitame nende hindamiseks järgida soovitusi, mis on antud juhendis ISO 11843-2.

Olenemata selles, millist LoD määramise metoodikat kasutatakse, on kasulik enne määramist teada ligikaudset LoD väärtust. Sellisel juhul on võimalik valida LoD hindamiseks sobivad kontsentratsioonid. See on eriti oluline just metoodikate puhul, mis kasutavad LoD hindamiseks ühel kontsentratsioonil tehtud kordusmõõtmiste tulemusi, sest vastava kontsentratsiooni valik mõjutab tugevasti tulemust, kui süsteem annab heteroskedastilisi tulemusi.

Selle töö käigus uuriti ja anti soovitusi selleks, et võtta arvesse LoD hinnangu päevadevahelist varieeruvust. Selle probleemi lahendamiseks pakuti välja uus parameeter “päevadevaheline LoD”. Lisaks sellele pakuti välja uus parameeter “laboritevaheline LoD”, mis aitab erinevaid analüüsimeetodeid omavahel võrrelda.

Lisaks näidati selle töö käigus ka, et visuaalne LoD hindamine kromatogrammide järgi ja andmete lineaarsuse ning skedastilisuse hindamine jääkliik-

mete visuaalse analüüsi kaudu on subjektiivne (sõltub analüüsijast). Seetõttu tuleb neid metoodikaid kasutades olla ettevaatlik.

Töö käigus leiti, et avastamispiiri teema nõuab mõnes aspektis veelgi detailsemat tähelepanu (näiteks *a priori* teadmiste arvesse võtmine ja tolerantsintervalli kasutamine LoD arvutamisel). Selle töö eesmärgid on aga täidetud ning loodetavasti motiveerib see töö edasist uurimustööd sellel teemal.

ACKNOWLEDGEMENTS

Most importantly I would deeply like to thank my long time supervisor Anneli for helping and guiding me through all these years. I would also like to thank Ivo for the constant inspiration and great advice. I feel that it would not have been possible to ask for better supervisors.

I would also like to thank Asko, Maarja-Liisa, Jaanus, Piia, Gunnar, Mari, Karl, Tõiv, Kristjan and everyone in the Chair of Analytical Chemistry for the great atmosphere. In this regard I would also like to thank Ott, Birgit, Olga, Marie, Hedi and others.

I am also very grateful for the support from my family – this work would have not been possible without them.

I would like to thank the University of Tartu and Institute of Chemistry for providing all the possibilities for developing the useful skill and knowledge necessary for this work and for the future. Finally, for the financial support I would like to thank the Estonian Research Council (Personal research funding Project 34 and the institutional research funding IUT20-14), the Estonian Science Foundation (grant No 8572), the EU Regional Development Fund (project "Development of software for validation of chromatographic methods" with registration number 3.2.1201.13-0020 under the sub-measure "Supporting the development of R&D of info and communication technology"), Enterprise Estonia Foundation (Estonian National Research and Development Infra-structure development program of measure 2.3 "Promotion of development activities and innovation", Regulation No. 34), and Archimedes Foundation (DoRa action 6, doctoral research in foreign university/research center).

REFERENCES

[1] P. Araujo, Key aspects of analytical method validation and linearity evaluation, J. Chromatogr. B. 877 (2009) 2224–2234. doi:10.1016/j.jchromb.2008.09.030.

[2] M. Thompson, S.L.R. Ellison, R. Wood, HARMONIZED GUIDELINES FOR SINGLELABORATORY VALIDATION OF METHODS OF ANALYSIS (IUPAC Technical Report), Pure Appl Chem. 74 (2002) 835–855.

[3] L.A. Currie, Limits for qualitative detection and quantitative determination. Application to radiochemistry, Anal. Chem. 40 (1968) 586–593.

[4] ISO 11843-2:2000 Capability of detection – Part 2: Methodology in the linear calibration case, (2000).

[5] L.A. Currie, NOMENCLATURE IN EVALUATION OF ANALYTICAL METHODS INCLUDING DETECTION AND QUANTIFICATION CAPABILITIES, Pure Appl Chem. 67 (1995) 1699–1723.

[6] EUROPEAN COMMISSION, COMMISSION DECISION of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, (2002).

[7] International vocabulary of metrology – Basic and general concepts and associated terms (VIM), (2012). http://digilib.uin-suka.ac.id/10881/ (accessed July 7, 2015).

[8] L.A. Currie, Detection and quantification limits: origins and historical overview, Anal. Chim. Acta. 391 (1999) 127–134.

[9] L.V. Rajaković, D.D. Marković, V.N. Rajaković-Ognjanović, D.Z. Antanasijević, Review: The approaches for estimation of limit of detection for ICP-MS trace analysis of arsenic, Talanta. 102 (2012) 79–87. doi:10.1016/j.talanta.2012.08.016.

[10] R.S. of C. Analytical Methods Committee, What should be done with results below the detection limit? Mentioning the unmentionable, 5 (2001).

[11] K. Danzer, Analytical Chemistry Theoretical and Metrological Fundamentals, 2007.

[12] Eurachem, EURACHEM Guide. The Fitness for Purpose of Analytical Methods A Laboratory Guide to Method Validation and Related Topics, [Eurachem], [Teddington], 1998.

[13] APPENDIX B TO PART 136 – DEFINITION AND PROCEDURE FOR THE DETERMINATION OF THE METHOD DETECTION LIMIT – REVISION 1. 11, (2012).

[14] A. Hubaux, G. Vos, Decision and detection limits for calibration curves, Anal. Chem. 42 (1970) 849–855.

[15] M. Thompson, S.L.R. Ellison, Towards an uncertainty paradigm of detection capability, Anal. Methods. 5 (2013) 5857. doi:10.1039/c3ay41209a.

[16] I. Lavagnini, D. Badocco, P. Pastore, F. Magno, Theil–Sen nonparametric regression technique on univariate calibration, inverse regression and detection limits, Talanta. 87 (2011) 180–188. doi:10.1016/j.talanta.2011.09.059.

[17] A.L. Rukhin, D.V. Samarov, Limit of detection determination for censored samples, Chemom. Intell. Lab. Syst. 105 (2011) 188–194. doi:10.1016/j.chemolab. 2011. 01.001.

[18] J. Fonollosa, A. Vergara, R. Huerta, S. Marco, Estimation of the limit of detection using information theory measures, Anal. Chim. Acta. 810 (2014) 1–9. doi:10.1016/j.aca.2013.10.030.

[19] E. Desimoni, B. Brunetti, About estimating the limit of detection of heteroscedastic analytical systems, Anal. Chim. Acta. 655 (2009) 30–37. doi:10.1016/ j.aca. 2009.09.036.
[20] The Fitness for Purpose of Analytical Methods A Laboratory Guide to Method Validation and Related Topics Second Edition, (2014).
[21] Guide in Validation of Alternative Proprietary Chemical Methods, (2010).
[22] M. Valcárcel, S. Cárdenas, D. Barceló, L. Buydens, K. Heydorn, B. Karlberg, K. Klemm, B. Lendl, B. Milman, B. Neidhart, A. Rios, R. Stephany, A. Townshend, A. Zschunke, Metrology of Qualitative Chemical Analysis, (2002).
[23] N. Dyson, Chromatographic Integration Methods, Second edition, 1998.
[24] A.G. Huesgen, Reliable and Automatic Integration of Trace Compounds Using the Agilent 1200 Infi nity Series High Dynamic Range Diode Array Detector Solution, (2014).
[25] AOAC Guidelines for Single Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals, (2002).
[26] Guidance for Industry Bioanalytical Method Validation, (2001).
[27] INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL, REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN, USE, INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL, ICH HARMONISED TRIPARTITE GUIDELINE VALIDATION OF ANALYTICAL PROCEDURES: TEXT AND METHODOLOGY Q2(R1), in: Int. Conf. Harmon. Geneva, 1994: pp. 1–5. http://members3.jcom. home.ne.jp/yrq01133/ctd/quality/q2a/q2astep4.pdf (accessed October 21, 2014).
[28] J.N. Miller, J.C. Miller, Statistics and chemometrics for analytical chemistry, Fifth Edition, Pearson Prentice Hall, Harlow, England, 2005.
[29] S. De Gryze, I. Langhans, M. Vandebroek, Using the correct intervals for prediction: A tutorial on tolerance intervals for ordinary least-squares regression, Chemom. Intell. Lab. Syst. 87 (2007) 147–154. doi:10.1016/j.chemolab. 2007.03.002.
[30] R.D. Gibbons, Some statistical and conceptual issues in the detection of low-level environmental pollutants, Environ. Ecol. Stat. 2 (1995) 125–145.
[31] J. Vial, A. Jardy, Experimental Comparison of the Different Approaches To Estimate LOD and LOQ of an HPLC Method, Anal. Chem. 71 (1999) 2672–2677. doi:10.1021/ac981179n.
[32] V.R. Meyer, Weighted Linear Least-Squares Fit – A Need? Monte Carlo Simulation Gives the Answer, LC-GC- Eur. 29 (2015) 204–209.
[33] A.M. Almeida, M.M. Castel-Branco, A.C. Falcao, Linear regression for calibration lines revisited: weighting schemes for bioanalytical methods, J. Chromatogr. B. (2002) 215 – 222.
[34] E.L. Johnson, D.L. Reynolds, D.S. Wright, L.A.. Pachl, Biological Sample Preparation and Data Reduction Concepts in Pharmaceutical Analysis, J. Chromatogr. Sci. 25 (1988) 372 – 379.
[35] J. Mandel, The Statistical Analysis of Experimental Data, 1984.
[36] SANCO/12571/2013 - Guidance document on analytical quality control and validation procedures for pesticide residues analysis in food and feed., (2013).
[37] L.A. Currie, Detection: International update, and some emerging di-lemmas involving calibration, the blank, and multiple detection decision, Chemom. Intell. Lab. Syst. 37 (1997) 151–181.
[38] E. Verdon, D. Hurtaud-Pessel, P. Sanders, Evaluation of the limit of performance of an analytical method based on a statistical calculation of its critical concentrations according to ISO standard 11843: Application to routine control of

banned veterinary drug residues in food according to European Decision 657/2002/EC, Accreditation Qual. Assur. 11 (2006) 58–62. doi:10.1007/s00769-005-0055-y.
[39] E. Desimoni, B. Brunetti, R. Cattaneo, Comparing some operational approaches to the limit of detection, Ann. Chim. 94 (2004) 555–569.
[40] R. Michel, Quality assurance of nuclear analytical techniques based on Bayesian characteristic limits, J. Radioanal. Nucl. Chem. 245 (1999) 137–144.
[41] M. Woldegebriel, G. Vivo-Truyols, Probabilistic Model for Untargeted Peak Detection in LC−MS Using Bayesian Statistics, (2015).
[42] D.B. Hibbert, N. Armstrong, An introduction to Bayesian methods for analyzing chemistry data Part 2, Chemom. Intell. Lab. Syst. 97 (2009) 211–220. doi:10.1016/j.chemolab.2009.03.009.
[43] N. Armstrong, D.B. Hibbert, An introduction to Bayesian methods for analyzing chemistry data Part 1, Chemom. Intell. Lab. Syst. 97 (2009) 194–210. doi:10.1016/j.chemolab.2009.04.001.
[44] J.H. Gross, Mass Spectrometry, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. http://link.springer.com/10.1007/978-3-642-10711-5 (accessed May 9, 2016).
[45] W.A. Korfmacher, Foundation review: Principles and applications of LC-MS in new drug discovery, Drug Discov. Today. 10 (2005) 1357–1367. doi:10.1016/S1359-6446(05)03620-2.
[46] A. Kruve, K. Herodes, I. Leito, Optimization of electrospray interface and quadrupole ion trap mass spectrometer parameters in pesticide liquid chromatography/electrospray ionization mass spectrometry analysis, Rapid Commun. Mass Spectrom. 24 (2010) 919–926. doi:10.1002/rcm.4470.
[47] M. Jemal, Y.-Q. Xia, LC-MS Development strategies for quantitative bioanalysis, Curr. Drug Metab. 7 (2006) 491–502.
[48] I. Taverniers, M. De Loose, E. Van Bockstaele, Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance, TrAC Trends Anal. Chem. 23 (2004) 535–552. doi:10.1016/j.trac.2004.04.001.
[49] P.J. Taylor, Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry, Clin. Biochem. 38 (2005) 328–334. doi:10.1016/j.clinbiochem.2004.11.007.
[50] C. Singleton, Recent advances in bioanalytical sample preparation for LC-MS analysis, Bioanalysis. 4 (2012) 1123–1140. doi:10.4155/bio.12.73.
[51] A. Tan, S. Hussain, A. Musuku, R. Massé, Internal standard response variations during incurred sample analysis by LC-MS/MS: Case by case trouble-shooting, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 877 (2009) 3201–3209. doi:10.1016/j.jchromb.2009.08.019.
[52] R.J.C. Brown, R.E. Yardley, A.S. Brown, P.R. Edwards, C. Rivier, C. Yardin, Analytical Methodologies with Very Low Blank Levels: Implications for Practical and Empirical Evaluations of the Limit of Detection, Anal. Lett. 39 (2006) 1229–1241. doi:10.1080/00032710600622563.
[53] Defining lower limit of quantitation, A Discussion od Signal / Noise, Reproducibility and Detector Tehnology in Quantitative LC/MS/MS Experiments, AB Sciex. (2010).
[54] M. Belter, A. Sajnóg, D. Barałkiewicz, Over a century of detection and quantification capabilities in analytical chemistry – Historical overview and trends, Talanta. 129 (2014) 606–616. doi:10.1016/j.talanta.2014.05.018.

[55] K. Linnet, Partly Nonparametric Approach for Determining the Limit of Detection, Clin. Chem. 50 (2004) 732–740. doi:10.1373/clinchem.2003.029983.
[56] T. Delatour, P. Mottier, E. Gremaud, Limits of suspicion, recognition and confirmation as concepts that account for the confirmation transitions at the detection limit for quantification by liquid chromatography – tandem mass spectrometry, J. Chromatogr. A. 1169 (2007) 103–110. doi:10.1016/j.chroma. 2007.08.065.
[57] E. de Hoffmann, V. Stroobant, Mass spectrometry: principles and applications, 3rd ed, J. Wiley, Chichester, West Sussex, England ; Hoboken, NJ, 2007.
[58] Applied Electrospray Mass Spectrometry: Practical Spectroscopy Series Volume 32, CRC Press. (2002). https://www.crcpress.com/Applied-Electrospray-Mass-Spectrometry-Practical-Spectroscopy-Series-Volume/Pramanik-Ganguly-Gross/9780824706180 (accessed May 9, 2016).
[59] L. Konermann, E. Ahadi, A.D. Rodriguez, S. Vahidi, Unraveling the Mechanism of Electrospray Ionization, Anal. Chem. 85 (2013) 2–9. doi:10.1021/ac302789c.
[60] W.M.A. Niessen, P. Manini, R. Andreoli, Matrix effects in quantitative pesticide analysis using liquid chromatography–mass spectrometry, Mass Spectrom. Rev. 25 (2006) 881–899. doi:10.1002/mas.20097.
[61] M.-Z. Huang, S.-C. Cheng, Y.-T. Cho, J. Shiea, Ambient ionization mass spectrometry: A tutorial, Anal. Chim. Acta. 702 (2011) 1–15. doi:10.1016/ j.aca.2011. 06.017.
[62] D.T. Snyder, C.J. Pulliam, Z. Ouyang, R.G. Cooks, Miniature and Fieldable Mass Spectrometers: Recent Advances, Anal. Chem. (2015) 151021154156004. doi:10.1021/ acs.analchem.5b03070.
[63] J. Liu, H. Wang, N.E. Manicke, J.-M. Lin, R.G. Cooks, Z. Ouyang, Development, Characterization, and Application of Paper Spray Ionization, Anal. Chem. 82 (2010) 2463–2471. doi:10.1021/ac902854g.
[64] P.A. Demirev, Dried Blood Spots: Analysis and Applications, Anal. Chem. 85 (2013) 779–789. doi:10.1021/ac303205m.
[65] Q. Yang, H. Wang, J.D. Maas, W.J. Chappell, N.E. Manicke, R.G. Cooks, Z. Ouyang, Paper spray ionization devices for direct, biomedical analysis using mass spectrometry, Int. J. Mass Spectrom. 312 (2012) 201–207. doi:10.1016/ j.ijms.2011.05.013.
[66] R.D. Espy, A.R. Muliadi, Z. Ouyang, R.G. Cooks, Spray mechanism in paper spray ionization, Int. J. Mass Spectrom. 325–327 (2012) 167–171. doi:10.1016/ j.ijms.2012.06.017.
[67] R.D. Espy, N.E. Manicke, Z. Ouyang, R.G. Cooks, Rapid analysis of whole blood by paper spray mass spectrometry for point-of-care therapeutic drug monitoring, The Analyst. 137 (2012) 2344. doi:10.1039/c2an35082c.
[68] N.E. Manicke, Q. Yang, H. Wang, S. Oradu, Z. Ouyang, R.G. Cooks, Assessment of paper spray ionization for quantitation of pharmaceuticals in blood spots, Int. J. Mass Spectrom. 300 (2011) 123–129. doi:10.1016/j.ijms.2010.06.037.
[69] N.E. Manicke, P. Abu-Rabie, N. Spooner, Z. Ouyang, R.G. Cooks, Quantitative Analysis of Therapeutic Drugs in Dried Blood Spot Samples by Paper Spray Mass Spectrometry: An Avenue to Therapeutic Drug Monitoring, J. Am. Soc. Mass Spectrom. 22 (2011) 1501–1507. doi:10.1007/s13361-011-0177-x.
[70] Q. Yang, N.E. Manicke, H. Wang, C. Petucci, R.G. Cooks, Z. Ouyang, Direct and quantitative analysis of underivatized acylcarnitines in serum and whole blood using paper spray mass spectrometry, Anal. Bioanal. Chem. 404 (2012) 1389–1397. doi:10.1007/s00216-012-6211-4.

[71] H. Wang, N.E. Manicke, Q. Yang, L. Zheng, R. Shi, R.G. Cooks, Z. Ouyang, Direct Analysis of Biological Tissue by Paper Spray Mass Spectrometry, Anal. Chem. 83 (2011) 1197–1201. doi:10.1021/ac103150a.
[72] A. Li, P. Wei, H.-C. Hsu, R.G. Cooks, Direct analysis of 4-methylimidazole in foods using paper spray mass spectrometry, The Analyst. 138 (2013) 4624. doi:10.1039/c3an00888f.
[73] P.-H. Lai, P.-C. Chen, Y.-W. Liao, J.-T. Liu, C.-C. Chen, C.-H. Lin, Comparison of gampi paper and nanofibers to chromatography paper used in paper spray-mass spectrometry, Int. J. Mass Spectrom. 375 (2015) 14–17. doi:10.1016/j.ijms.2014.10.013.
[74] R.D. Espy, S.F. Teunissen, N.E. Manicke, Y. Ren, Z. Ouyang, A. van Asten, R.G. Cooks, Paper Spray and Extraction Spray Mass Spectrometry for the Direct and Simultaneous Quantification of Eight Drugs of Abuse in Whole Blood, Anal. Chem. 86 (2014) 7712–7718. doi:10.1021/ac5016408.
[75] S. Jain, A. Heiser, A.R. Venter, Spray desorption collection: an alternative to swabbing for pharmaceutical cleaning validation, The Analyst. 136 (2011) 1298. doi:10.1039/c0an00728e.
[76] L.E. Sojo, G. Lum, P. Chee, Internal standard signal suppression by co-eluting analyte in isotope dilution LC-ESI-MS, The Analyst. 128 (2003) 51–54. doi:10.1039/b209521c.
[77] M. Anastassiades, S.J. Lehotay, D. Štajnbaher, F.J. Schenck, Fast and easy multiresidue method employing acetonitrile extraction/partitioning and "dispersive solid-phase extraction" for the determination of pesticide residues in produce, J. AOAC Int. 86 (2003) 412–431.
[78] A. Kruve, I. Leito, K. Herodes, A. Laaniste, R. Lõhmus, Enhanced Nebulization Efficiency of Electrospray Mass Spectrometry: Improved Sensitivity and Detection Limit, J. Am. Soc. Mass Spectrom. 23 (2012) 2051–2054. doi:10.1007/s13361-012-0475-y.
[79] C. Vega, C. Spence, C. Zhang, B.J. Bills, N.E. Manicke, Ionization Suppression and Recovery in Direct Biofluid Analysis Using Paper Spray Mass Spectrometry, J. Am. Soc. Mass Spectrom. (2016). doi:10.1007/s13361-015-1322-8.
[80] EU Pesticide Database Regulation (EC) No. 395/2005. Retrieved from: http://ec.europa.eu/food/plant/pesticides/eu-pesticides-database-redirect/index_en.htm (9.5.2016), (n.d.).

APPENDIX 1

R script for simulation of calibration graph with known variance

R script to simulate a calibration graph with known (homoscedastic) variance and known CC_α, CC_β values is given. LOD values can be calculated from the simulated data. It is tested how the estimated LOD changes in case different calibration levels are used.

```
# Function for simulating and calculating calibration
graph parameters
# (variables similar to previous scripts)
simul <- function(Concentration, b = 5, a = 0){

        # Data is simulated:
    res = 100*rnorm(length(Concentration))
    y = b*Concentration + a + res

        # Calculations with simulated data:
    CF <- lm(y ~ Concentration)
    Nsamp <- length(Concentration)

        # Output parameters:
    Bcalc <- CF$coefficient[[2]] # estimated slope
    Sxy <- ( sum(resid(CF)^2 ) / (Nsamp - 2) )^0.5
    Acalc = CF$coefficient[[1]] # estimated intercept
    Aerror = summary(CF)$coef[1,2] # estimated
standard deviation of intercept
    LoD.resid <- 3.3*Sxy/Bcalc
    LoD.int <- 3.3*Aerror/Bcalc
    mat <- cbind(Bcalc, Syx, Acalc, Aerror, y[1],
LoD.resid, LoD.int)
    return(mat)
}

# True CCa and CCb in concentration scale at alpha
and beta 5%:
# CCa = (1.645 * 100)/5 = 32.9
# CCb = (1.645 * 100 + 1.645 * 100)/5 = 65.8

n = 1000

# Concentration levels 1: Blank + all other levels
over CCb)
C1 = c(0, 75, 100, 125, 150, 175, 200)
data1 <- matrix(nrow = n, ncol = 7)
```

```
for(i in 1:n){
    data1[i, ] <- simul(C1)
}
bind1 <- cbind(mean(data1[, 1]), sd(data1[, 1]),
mean(data1[, 2]), sd(data1[, 2]),
               mean(data1[, 3]),mean(data1[, 4]),
mean(data1[, 5]), sd(data1[, 5]),
               mean(data1[, 6]), sd(data1[, 6]),
               mean(data1[, 7]), sd(data1[, 7]))

# Concentration levels 2: Blank + all other levels
over CCa
C2 = c(0, 35, 60, 85, 110, 135, 160)
data2 <- matrix(nrow = n, ncol = 7)
for(i in 1:n){
    data2[i, ] <- simul(C2)
}
bind2 <- cbind(mean(data2[, 1]), sd(data2[, 1]),
mean(data2[, 2]), sd(data2[, 2]),
               mean(data2[, 3]),mean(data2[, 4]),
mean(data2[, 5]), sd(data2[, 5]),
               mean(data2[, 6]), sd(data2[, 6]),
               mean(data2[, 7]), sd(data2[, 7]))

# Half of Concentrations under CCb, 2 are under CCa,
1 under CCb
C3 = c(0, 5, 25, 50, 75, 100, 125)
data3 <- matrix(nrow = n, ncol = 7)
for(i in 1:n){
    data3[i, ] <- simul(C3)
}
bind3 <- cbind(mean(data3[, 1]), sd(data3[, 1]),
mean(data3[, 2]), sd(data3[, 2]),
               mean(data3[, 3]),mean(data3[, 4]),
mean(data3[, 5]), sd(data3[, 5]),
               mean(data3[, 6]), sd(data3[, 6]),
               mean(data3[, 7]), sd(data3[, 7]))


# All concentrations below CCa
C4 = c(0, 5, 10, 15, 20, 25, 30)
data4 <- matrix(nrow = n, ncol = 7)
for(i in 1:n){
    data4[i, ] <- simul(C4)
}
```

```
bind4 <- cbind(mean(data4[, 1]), sd(data4[, 1]),
mean(data4[, 2]), sd(data4[, 2]),
                mean(data4[, 3]),mean(data4[, 4]),
mean(data4[, 5]), sd(data4[, 5]),
                mean(data4[, 6]), sd(data4[, 6]),
                mean(data4[, 7]), sd(data4[, 7]))

# the results:
bind <- rbind(bind1, bind2, bind3, bind4)
df <- data.frame(meanBcalc = bind[, 1], sdBcalc =
bind[, 2],
                 meanSyx = bind[, 3], sdSyx = bind[,
4],
                 meanAcalc = bind[, 5], meanAerror =
bind[, 6],
                 meanBlank = bind[, 7], sdBlank =
bind[, 8],
                 meanLoD.res = bind[, 9], sdLoD.res =
bind[, 10],
                 meanLoD.int = bind[, 11], sdLoD.int
= bind[, 12])
df
```

R script to simulate a calibration graph with known (heteroscedastic) variance and known CC_α, CC_β values is given. LOD values can be calculated from the simulated data similarly to the case with homoscedastic data.

```
simul <- function(Concentration, b = 5, a = 0){

        # Simulation of data
    res <- NULL
    n <- length(Concentration)
    for(i in 1:n){
        res[i] = rnorm(1) * (0.5 * Concentration[i] +
100)
    }
    y = b * Concentration + a + res

        # Calculations with simulated data:
    CF <- lm(y ~ Concentration)
    Nsamp <- length(Concentration)

        # Outputs:
    Bcalc <- CF$coefficient[[2]] # estimated slope
    Sxy <- ( sum(resid(CF)^2 ) / (Nsamp - 2) )^0.5
    Acalc = CF$coefficient[[1]] # estimated intercept
```

```
    Aerror = summary(CF)$coef[1,2] # estimated
standard deviation of intercept
    LoD.resid <- 3.3*Sxy/Bcalc
    LoD.int <- 3.3*Aerror/Bcalc
    mat <- cbind(Bcalc, Syx, Acalc, Aerror, y[1],
LoD.resid, LoD.int)
    return(mat)
}

# True CCa and CCb in concentration scale at alpha
and beta 5%:
# CCa = (100 * 1.645)/5 = 32.6
# CCb = ((100 * 1.645) + 164.5) / (5 - 0.5) = 329 /
4.5 = 73.11
```

APPENDIX 2

R script for simulation of prediction interval value for blank samples

The R script to simulate prediction interval value of homoscedastic data in case of different calibration levels. Equation suggested in ISO 11843-2 [4] to estimate CC_α is used to calculate the prediction interval value at 0 concentration.

```
    # Conc = concentration values of the calibration
points;
    # b = slope of true calibration function, a =
intercept of true calibration function;
    # n = number of repeated measurements at each
calibration level;
    # m = number of sample measurements;
    # sd = standard deviation of the data.


# Simulation of prediction interval (of the mean) for
linear homoscedastic data

function.homosc <- function(Conc, b = 5, a = 5, n =
4, sd = 20, m = 5){

        # Simulation of data
    data <- matrix(nrow = length(Conc), ncol = n)
    i = 1
    while(i <= n){
        res = rnorm(length(Conc), sd = sd)
        y = b * Conc + a + res
        data[, i] = y
        i = i + 1
    }

        # Calculating the parameters of the simulated
regression
    df <- data.frame(ppm = Conc, y = data)
    Y <- c(data[, 1:n])
    C <- rep(df[, 1], n)
    OLS <- lm(Y ~ C)
    Bcalc = OLS$coef[2]
    Acalc = OLS$coef[1]
    Nsamp <- length(Conc) # Nsamp = sample size
(total number of measurements)
    Syx = ( sum(resid(OLS)^2 ) / (Nsamp - 2) )^0.5
```

```
    y <- Bcalc * Conc + Acalc

    t <- 2.35

        # Calculation of the prediction interval
values
        # (ISO 11843-2 suggested equation to estimate
CCa is used)
    PI <- t * Syx/Bcalc * sqrt( (1/m) + (1/Nsamp) +
                                (    (mean(C))^2    /
(sum((C-mean(C))^2) ) ) )

    return(PI)
}

n <- 1000

C1 = c(0, 200, 400, 600, 800, 1000)
y1 <- NULL
for(i in 1:n){

    f <- function.homosc(C1)
    y1[i] <- f[1]
}

C2 = c(0, 5, 10, 15, 20, 1000)
y2 <- NULL
for(i in 1:n){

    g <- function.homosc(C2)
    y2[i] <- g[1]

}

mean(y1)
mean(y2)
sd(y1)
sd(y2)
```

PUBLICATIONS

CURRICULUM VITAE

Name: Hanno Evard
Date of birth: August 1, 1988, Tartu, Eesti
Citizenship: Eesti
Address: Võru 100–53, Tartu, 50111
Telephone: +372 50 92 501
E-mail: evardhanno@gmail.com

Education:

2012–…	University of Tartu, Institute of Chemistry, PhD student
1.10.2013–22.12.2013	University of Urbino, Italy, (exchange student)
2010–2012	University of Tartu, Institute of Chemistry, M.Sc (Applied Measurement Science)
2007–2010	University of Tartu, B.Sc (Material science)

Professional employment:

1.3.2016 – …	Chemist (0.1), University of Tartu, Institute of Chemistry
1.9.2011–30.10.2012	Chemist (0.5), TBD-Biodiscovery
1.6.2011–30.6.2011	Trainee, Estonian Environmental Research Center, Tartu

Language skills:
Estonian – mother tongue
English – very good

Main fields of research:
Liquid chromatography, mass spectrometry, LC-ESI/MS, validation, limit of detection, ambient ionization methods, paper spray ionization

A list of publications:
a. H. Evard, A. Kruve, R. Lõhmus, I. Leito, Paper spray ionization mass spectrometry: Study of a method for fast-screening analysis of pesticides in fruits and vegetables, J. Food Compos. Anal. 41 (2015) 221–225.
b. 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. 870 (2015) 29–44
c. 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. 870 (2015) 8–28.

Research grants and scholarships:
2013 September – DoRa action 6, doctoral research in foreign university/ research center, SA Archimedes

Teaching work

Information regarding the teaching work carried out at universities:

– Analytical Chemistry I practical
– Statistics in Analytical Chemistry

ELULOOKIRJELDUS

Nimi: Hanno Evard
Sünniaeg: 1. august, 1988, Tartu, Eesti
Kodakondsus: Eesti
Aadress: Võru 100–53, Tartu, 50111
Telefon: +372 50 92 501
E-mail: evardhanno@gmail.com

Haridus:

2012– …	Tartu Ülikool, Keemia Instituut, doktorantuur
1.10.2013–22.12.2013	Urbino Ülikool, Itaalia, (vahetusüliõpilane)
2010–2012	Tartu Ülikool, Keemia Instituut, magistratuur (Rakenduslik mõõteteadus)
2007–2010	Tartu Ülikool, bakalaureus (Materjaliteadus)

Erialane teenistuskäik:

1.3.2016 – …	Keemik (0,1), Tartu Ülikool, Keemia Instituut
1.9.2011–30.10.2012	Keemik (0,5), TBD-Biodiscovery
1.6.2011–30.6.2011	Praktikant, Eesti Keskonnauuringute Keskus, Tartu labor

Keelteoskus:
Eesti keel – emakeel
Inglise keel – väga hea

Peamised uurimisvaldkonnad:
Vedelikkromatograafia, massispektromeetria, LC-ESI/MS, valideerimine, avastamispiir, atmosfäärirõhulised ionisatsiooniallikad, paberpihustus ionisatsioon

Publikatsioonide loetelu:

a. H. Evard, A. Kruve, R. Lõhmus, I. Leito, Paper spray ionization mass spectrometry: Study of a method for fast-screening analysis of pesticides in fruits and vegetables, J. Food Compos. Anal. 41 (2015) 221–225.
b. 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. 870 (2015) 29–44
c. 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. 870 (2015) 8–28.

Saadud uurimistoetused ja stipendiumid:
2013 September – DoRa tegevus 6, doktorandi uurimustöö välismaa ülikoolis/ teadusasutuses, SA Archimedes

Õppetöö
Kõrgkoolis tehtud auditoorne õppetöö:
– Analüütiline keemia I praktikum
– Statistika analüütilises keemias

DISSERTATIONES CHIMICAE UNIVERSITATIS TARTUENSIS

1. **Toomas Tamm.** Quantum-chemical simulation of solvent effects. Tartu, 1993, 110 p.
2. **Peeter Burk.** Theoretical study of gas-phase acid-base equilibria. Tartu, 1994, 96 p.
3. **Victor Lobanov.** Quantitative structure-property relationships in large descriptor spaces. Tartu, 1995, 135 p.
4. **Vahur Mäemets.** The ^{17}O and ^{1}H nuclear magnetic resonance study of H_2O in individual solvents and its charged clusters in aqueous solutions of electrolytes. Tartu, 1997, 140 p.
5. **Andrus Metsala.** Microcanonical rate constant in nonequilibrium distribution of vibrational energy and in restricted intramolecular vibrational energy redistribution on the basis of slater's theory of unimolecular reactions. Tartu, 1997, 150 p.
6. **Uko Maran.** Quantum-mechanical study of potential energy surfaces in different environments. Tartu, 1997, 137 p.
7. **Alar Jänes.** Adsorption of organic compounds on antimony, bismuth and cadmium electrodes. Tartu, 1998, 219 p.
8. **Kaido Tammeveski.** Oxygen electroreduction on thin platinum films and the electrochemical detection of superoxide anion. Tartu, 1998, 139 p.
9. **Ivo Leito.** Studies of Brønsted acid-base equilibria in water and non-aqueous media. Tartu, 1998, 101 p.
10. **Jaan Leis.** Conformational dynamics and equilibria in amides. Tartu, 1998, 131 p.
11. **Toonika Rinken.** The modelling of amperometric biosensors based on oxidoreductases. Tartu, 2000, 108 p.
12. **Dmitri Panov.** Partially solvated Grignard reagents. Tartu, 2000, 64 p.
13. **Kaja Orupõld.** Treatment and analysis of phenolic wastewater with microorganisms. Tartu, 2000, 123 p.
14. **Jüri Ivask.** Ion Chromatographic determination of major anions and cations in polar ice core. Tartu, 2000, 85 p.
15. **Lauri Vares.** Stereoselective Synthesis of Tetrahydrofuran and Tetrahydropyran Derivatives by Use of Asymmetric Horner-Wadsworth-Emmons and Ring Closure Reactions. Tartu, 2000, 184 p.
16. **Martin Lepiku.** Kinetic aspects of dopamine D_2 receptor interactions with specific ligands. Tartu, 2000, 81 p.
17. **Katrin Sak.** Some aspects of ligand specificity of P2Y receptors. Tartu, 2000, 106 p.
18. **Vello Pällin.** The role of solvation in the formation of iotsitch complexes. Tartu, 2001, 95 p.
19. **Katrin Kollist.** Interactions between polycyclic aromatic compounds and humic substances. Tartu, 2001, 93 p.

20. **Ivar Koppel.** Quantum chemical study of acidity of strong and superstrong Brønsted acids. Tartu, 2001, 104 p.
21. **Viljar Pihl.** The study of the substituent and solvent effects on the acidity of OH and CH acids. Tartu, 2001, 132 p.
22. **Natalia Palm.** Specification of the minimum, sufficient and significant set of descriptors for general description of solvent effects. Tartu, 2001, 134 p.
23. **Sulev Sild.** QSPR/QSAR approaches for complex molecular systems. Tartu, 2001, 134 p.
24. **Ruslan Petrukhin.** Industrial applications of the quantitative structure-property relationships. Tartu, 2001, 162 p.
25. **Boris V. Rogovoy.** Synthesis of (benzotriazolyl)carboximidamides and their application in relations with *N*- and *S*-nucleophyles. Tartu, 2002, 84 p.
26. **Koit Herodes.** Solvent effects on UV-vis absorption spectra of some solvatochromic substances in binary solvent mixtures: the preferential solvation model. Tartu, 2002, 102 p.
27. **Anti Perkson.** Synthesis and characterisation of nanostructured carbon. Tartu, 2002, 152 p.
28. **Ivari Kaljurand.** Self-consistent acidity scales of neutral and cationic Brønsted acids in acetonitrile and tetrahydrofuran. Tartu, 2003, 108 p.
29. **Karmen Lust.** Adsorption of anions on bismuth single crystal electrodes. Tartu, 2003, 128 p.
30. **Mare Piirsalu.** Substituent, temperature and solvent effects on the alkaline hydrolysis of substituted phenyl and alkyl esters of benzoic acid. Tartu, 2003, 156 p.
31. **Meeri Sassian.** Reactions of partially solvated Grignard reagents. Tartu, 2003, 78 p.
32. **Tarmo Tamm.** Quantum chemical modelling of polypyrrole. Tartu, 2003. 100 p.
33. **Erik Teinemaa.** The environmental fate of the particulate matter and organic pollutants from an oil shale power plant. Tartu, 2003. 102 p.
34. **Jaana Tammiku-Taul.** Quantum chemical study of the properties of Grignard reagents. Tartu, 2003. 120 p.
35. **Andre Lomaka.** Biomedical applications of predictive computational chemistry. Tartu, 2003. 132 p.
36. **Kostyantyn Kirichenko.** Benzotriazole – Mediated Carbon–Carbon Bond Formation. Tartu, 2003. 132 p.
37. **Gunnar Nurk.** Adsorption kinetics of some organic compounds on bismuth single crystal electrodes. Tartu, 2003, 170 p.
38. **Mati Arulepp.** Electrochemical characteristics of porous carbon materials and electrical double layer capacitors. Tartu, 2003, 196 p.
39. **Dan Cornel Fara.** QSPR modeling of complexation and distribution of organic compounds. Tartu, 2004, 126 p.
40. **Riina Mahlapuu.** Signalling of galanin and amyloid precursor protein through adenylate cyclase. Tartu, 2004, 124 p.

41. **Mihkel Kerikmäe.** Some luminescent materials for dosimetric applications and physical research. Tartu, 2004, 143 p.
42. **Jaanus Kruusma.** Determination of some important trace metal ions in human blood. Tartu, 2004, 115 p.
43. **Urmas Johanson.** Investigations of the electrochemical properties of polypyrrole modified electrodes. Tartu, 2004, 91 p.
44. **Kaido Sillar.** Computational study of the acid sites in zeolite ZSM-5. Tartu, 2004, 80 p.
45. **Aldo Oras.** Kinetic aspects of dATPαS interaction with $P2Y_1$ receptor. Tartu, 2004, 75 p.
46. **Erik Mölder.** Measurement of the oxygen mass transfer through the air-water interface. Tartu, 2005, 73 p.
47. **Thomas Thomberg.** The kinetics of electroreduction of peroxodisulfate anion on cadmium (0001) single crystal electrode. Tartu, 2005, 95 p.
48. **Olavi Loog.** Aspects of condensations of carbonyl compounds and their imine analogues. Tartu, 2005, 83 p.
49. **Siim Salmar.** Effect of ultrasound on ester hydrolysis in aqueous ethanol. Tartu, 2006, 73 p.
50. **Ain Uustare.** Modulation of signal transduction of heptahelical receptors by other receptors and G proteins. Tartu, 2006, 121 p.
51. **Sergei Yurchenko.** Determination of some carcinogenic contaminants in food. Tartu, 2006, 143 p.
52. **Kaido Tämm.** QSPR modeling of some properties of organic compounds. Tartu, 2006, 67 p.
53. **Olga Tšubrik.** New methods in the synthesis of multisubstituted hydrazines. Tartu. 2006, 183 p.
54. **Lilli Sooväli.** Spectrophotometric measurements and their uncertainty in chemical analysis and dissociation constant measurements. Tartu, 2006, 125 p.
55. **Eve Koort.** Uncertainty estimation of potentiometrically measured ph and pK_a values. Tartu, 2006, 139 p.
56. **Sergei Kopanchuk.** Regulation of ligand binding to melanocortin receptor subtypes. Tartu, 2006, 119 p.
57. **Silvar Kallip.** Surface structure of some bismuth and antimony single crystal electrodes. Tartu, 2006, 107 p.
58. **Kristjan Saal.** Surface silanization and its application in biomolecule coupling. Tartu, 2006, 77 p.
59. **Tanel Tätte.** High viscosity $Sn(OBu)_4$ oligomeric concentrates and their applications in technology. Tartu, 2006, 91 p.
60. **Dimitar Atanasov Dobchev**. Robust QSAR methods for the prediction of properties from molecular structure. Tartu, 2006, 118 p.
61. **Hannes Hagu**. Impact of ultrasound on hydrophobic interactions in solutions. Tartu, 2007, 81 p.

62. **Rutha Jäger.** Electroreduction of peroxodisulfate anion on bismuth electrodes. Tartu, 2007, 142 p.
63. **Kaido Viht.** Immobilizable bisubstrate-analogue inhibitors of basophilic protein kinases: development and application in biosensors. Tartu, 2007, 88 p.
64. **Eva-Ingrid Rõõm.** Acid-base equilibria in nonpolar media. Tartu, 2007, 156 p.
65. **Sven Tamp.** DFT study of the cesium cation containing complexes relevant to the cesium cation binding by the humic acids. Tartu, 2007, 102 p.
66. **Jaak Nerut.** Electroreduction of hexacyanoferrate(III) anion on Cadmium (0001) single crystal electrode. Tartu, 2007, 180 p.
67. **Lauri Jalukse.** Measurement uncertainty estimation in amperometric dissolved oxygen concentration measurement. Tartu, 2007, 112 p.
68. **Aime Lust.** Charge state of dopants and ordered clusters formation in CaF_2:Mn and CaF_2:Eu luminophors. Tartu, 2007, 100 p.
69. **Iiris Kahn**. Quantitative Structure-Activity Relationships of environmentally relevant properties. Tartu, 2007, 98 p.
70. **Mari Reinik.** Nitrates, nitrites, N-nitrosamines and polycyclic aromatic hydrocarbons in food: analytical methods, occurrence and dietary intake. Tartu, 2007, 172 p.
71. **Heili Kasuk.** Thermodynamic parameters and adsorption kinetics of organic compounds forming the compact adsorption layer at Bi single crystal electrodes. Tartu, 2007, 212 p.
72. **Erki Enkvist.** Synthesis of adenosine-peptide conjugates for biological applications. Tartu, 2007, 114 p.
73. **Svetoslav Hristov Slavov**. Biomedical applications of the QSAR approach. Tartu, 2007, 146 p.
74. **Eneli Härk.** Electroreduction of complex cations on electrochemically polished Bi(*hkl*) single crystal electrodes. Tartu, 2008, 158 p.
75. **Priit Möller.** Electrochemical characteristics of some cathodes for medium temperature solid oxide fuel cells, synthesized by solid state reaction technique. Tartu, 2008, 90 p.
76. **Signe Viggor.** Impact of biochemical parameters of genetically different pseudomonads at the degradation of phenolic compounds. Tartu, 2008, 122 p.
77. **Ave Sarapuu.** Electrochemical reduction of oxygen on quinone-modified carbon electrodes and on thin films of platinum and gold. Tartu, 2008, 134 p.
78. **Agnes Kütt.** Studies of acid-base equilibria in non-aqueous media. Tartu, 2008, 198 p.
79. **Rouvim Kadis.** Evaluation of measurement uncertainty in analytical chemistry: related concepts and some points of misinterpretation. Tartu, 2008, 118 p.
80. **Valter Reedo.** Elaboration of IVB group metal oxide structures and their possible applications. Tartu, 2008, 98 p.

81. **Aleksei Kuznetsov.** Allosteric effects in reactions catalyzed by the cAMP-dependent protein kinase catalytic subunit. Tartu, 2009, 133 p.
82. **Aleksei Bredihhin.** Use of mono- and polyanions in the synthesis of multisubstituted hydrazine derivatives. Tartu, 2009, 105 p.
83. **Anu Ploom.** Quantitative structure-reactivity analysis in organosilicon chemistry. Tartu, 2009, 99 p.
84. **Argo Vonk.** Determination of adenosine A_{2A}- and dopamine D_1 receptor-specific modulation of adenylate cyclase activity in rat striatum. Tartu, 2009, 129 p.
85. **Indrek Kivi.** Synthesis and electrochemical characterization of porous cathode materials for intermediate temperature solid oxide fuel cells. Tartu, 2009, 177 p.
86. **Jaanus Eskusson.** Synthesis and characterisation of diamond-like carbon thin films prepared by pulsed laser deposition method. Tartu, 2009, 117 p.
87. **Marko Lätt.** Carbide derived microporous carbon and electrical double layer capacitors. Tartu, 2009, 107 p.
88. **Vladimir Stepanov.** Slow conformational changes in dopamine transporter interaction with its ligands. Tartu, 2009, 103 p.
89. **Aleksander Trummal.** Computational Study of Structural and Solvent Effects on Acidities of Some Brønsted Acids. Tartu, 2009, 103 p.
90. **Eerold Vellemäe.** Applications of mischmetal in organic synthesis. Tartu, 2009, 93 p.
91. **Sven Parkel.** Ligand binding to 5-HT_{1A} receptors and its regulation by Mg^{2+} and Mn^{2+}. Tartu, 2010, 99 p.
92. **Signe Vahur.** Expanding the possibilities of ATR-FT-IR spectroscopy in determination of inorganic pigments. Tartu, 2010, 184 p.
93. **Tavo Romann**. Preparation and surface modification of bismuth thin film, porous, and microelectrodes. Tartu, 2010, 155 p.
94. **Nadežda Aleksejeva.** Electrocatalytic reduction of oxygen on carbon nanotube-based nanocomposite materials. Tartu, 2010, 147 p.
95. **Marko Kullapere.** Electrochemical properties of glassy carbon, nickel and gold electrodes modified with aryl groups. Tartu, 2010, 233 p.
96. **Liis Siinor.** Adsorption kinetics of ions at Bi single crystal planes from aqueous electrolyte solutions and room-temperature ionic liquids. Tartu, 2010, 101 p.
97. **Angela Vaasa.** Development of fluorescence-based kinetic and binding assays for characterization of protein kinases and their inhibitors. Tartu 2010, 101 p.
98. **Indrek Tulp.** Multivariate analysis of chemical and biological properties. Tartu 2010, 105 p.
99. **Aare Selberg.** Evaluation of environmental quality in Northern Estonia by the analysis of leachate. Tartu 2010, 117 p.
100. **Darja Lavõgina.** Development of protein kinase inhibitors based on adenosine analogue-oligoarginine conjugates. Tartu 2010, 248 p.

101. **Laura Herm.** Biochemistry of dopamine D_2 receptors and its association with motivated behaviour. Tartu 2010, 156 p.
102. **Terje Raudsepp.** Influence of dopant anions on the electrochemical properties of polypyrrole films. Tartu 2010, 112 p.
103. **Margus Marandi.** Electroformation of Polypyrrole Films: *In-situ* AFM and STM Study. Tartu 2011, 116 p.
104. **Kairi Kivirand.** Diamine oxidase-based biosensors: construction and working principles. Tartu, 2011, 140 p.
105. **Anneli Kruve.** Matrix effects in liquid-chromatography electrospray mass-spectrometry. Tartu, 2011, 156 p.
106. **Gary Urb.** Assessment of environmental impact of oil shale fly ash from PF and CFB combustion. Tartu, 2011, 108 p.
107. **Nikita Oskolkov.** A novel strategy for peptide-mediated cellular delivery and induction of endosomal escape. Tartu, 2011, 106 p.
108. **Dana Martin.** The QSPR/QSAR approach for the prediction of properties of fullerene derivatives. Tartu, 2011, 98 p.
109. **Säde Viirlaid.** Novel glutathione analogues and their antioxidant activity. Tartu, 2011, 106 p.
110. **Ülis Sõukand.** Simultaneous adsorption of Cd^{2+}, Ni^{2+}, and Pb^{2+} on peat. Tartu, 2011, 124 p.
111. **Lauri Lipping.** The acidity of strong and superstrong Brønsted acids, an outreach for the "limits of growth": a quantum chemical study. Tartu, 2011, 124 p.
112. **Heisi Kurig.** Electrical double-layer capacitors based on ionic liquids as electrolytes. Tartu, 2011, 146 p.
113. **Marje Kasari.** Bisubstrate luminescent probes, optical sensors and affinity adsorbents for measurement of active protein kinases in biological samples. Tartu, 2012, 126 p.
114. **Kalev Takkis.** Virtual screening of chemical databases for bioactive molecules. Tartu, 2012, 122 p.
115. **Ksenija Kisseljova**. Synthesis of aza-β^3-amino acid containing peptides and kinetic study of their phosphorylation by protein kinase A. Tartu, 2012, 104 p.
116. **Riin Rebane.** Advanced method development strategy for derivatization LC/ESI/MS. Tartu, 2012, 184 p.
117. **Vladislav Ivaništšev.** Double layer structure and adsorption kinetics of ions at metal electrodes in room temperature ionic liquids. Tartu, 2012, 128 p.
118. **Irja Helm.** High accuracy gravimetric Winkler method for determination of dissolved oxygen. Tartu, 2012, 139 p.
119. **Karin Kipper.** Fluoroalcohols as Components of LC-ESI-MS Eluents: Usage and Applications. Tartu, 2012, 164 p.
120. **Arno Ratas.** Energy storage and transfer in dosimetric luminescent materials. Tartu, 2012, 163 p.

121. **Reet Reinart-Okugbeni**. Assay systems for characterisation of subtype-selective binding and functional activity of ligands on dopamine receptors. Tartu, 2012, 159 p.
122. **Lauri Sikk.** Computational study of the Sonogashira cross-coupling reaction. Tartu, 2012, 81 p.
123. **Karita Raudkivi.** Neurochemical studies on inter-individual differences in affect-related behaviour of the laboratory rat. Tartu, 2012, 161 p.
124. **Indrek Saar.** Design of GalR2 subtype specific ligands: their role in depression-like behavior and feeding regulation. Tartu, 2013, 126 p.
125. **Ann Laheäär.** Electrochemical characterization of alkali metal salt based non-aqueous electrolytes for supercapacitors. Tartu, 2013, 127 p.
126. **Kerli Tõnurist.** Influence of electrospun separator materials properties on electrochemical performance of electrical double-layer capacitors. Tartu, 2013, 147 p.
127. **Kaija Põhako-Esko.** Novel organic and inorganic ionogels: preparation and characterization. Tartu, 2013, 124 p.
128. **Ivar Kruusenberg.** Electroreduction of oxygen on carbon nanomaterial-based catalysts. Tartu, 2013, 191 p.
129. **Sander Piiskop.** Kinetic effects of ultrasound in aqueous acetonitrile solutions. Tartu, 2013, 95 p.
130. **Ilona Faustova**. Regulatory role of L-type pyruvate kinase N-terminal domain. Tartu, 2013, 109 p.
131. **Kadi Tamm.** Synthesis and characterization of the micro-mesoporous anode materials and testing of the medium temperature solid oxide fuel cell single cells. Tartu, 2013, 138 p.
132. **Iva Bozhidarova Stoyanova-Slavova.** Validation of QSAR/QSPR for regulatory purposes. Tartu, 2013, 109 p.
133. **Vitali Grozovski**. Adsorption of organic molecules at single crystal electrodes studied by *in situ* STM method. Tartu, 2014, 146 p.
134. **Santa Veikšina**. Development of assay systems for characterisation of ligand binding properties to melanocortin 4 receptors. Tartu, 2014, 151 p.
135. **Jüri Liiv.** PVDF (polyvinylidene difluoride) as material for active element of twisting-ball displays. Tartu, 2014, 111 p.
136. **Kersti Vaarmets.** Electrochemical and physical characterization of pristine and activated molybdenum carbide-derived carbon electrodes for the oxygen electroreduction reaction. Tartu, 2014, 131 p.
137. **Lauri Tõntson**. Regulation of G-protein subtypes by receptors, guanine nucleotides and Mn^{2+}. Tartu, 2014, 105 p.
138. **Aiko Adamson.** Properties of amine-boranes and phosphorus analogues in the gas phase. Tartu, 2014, 78 p.
139. **Elo Kibena**. Electrochemical grafting of glassy carbon, gold, highly oriented pyrolytic graphite and chemical vapour deposition-grown graphene electrodes by diazonium reduction method. Tartu, 2014, 184 p.

140. **Teemu Näykki.** Novel Tools for Water Quality Monitoring – From Field to Laboratory. Tartu, 2014, 202 p.
141. **Karl Kaupmees.** Acidity and basicity in non-aqueous media: importance of solvent properties and purity. Tartu, 2014, 128 p.
142. **Oleg Lebedev**. Hydrazine polyanions: different strategies in the synthesis of heterocycles. Tartu, 2015, 118 p.
143. **Geven Piir.** Environmental risk assessment of chemicals using QSAR methods. Tartu, 2015, 123 p.
144. **Olga Mazina.** Development and application of the biosensor assay for measurements of cyclic adenosine monophosphate in studies of G protein-coupled receptor signalinga. Tartu, 2015, 116 p.
145. **Sandip Ashokrao Kadam.** Anion receptors: synthesis and accurate binding measurements. Tartu, 2015, 116 p.
146. **Indrek Tallo.** Synthesis and characterization of new micro-mesoporous carbide derived carbon materials for high energy and power density electrical double layer capacitors. Tartu, 2015, 148 p.
147. **Heiki Erikson.** Electrochemical reduction of oxygen on nanostructured palladium and gold catalysts. Tartu, 2015, 204 p.
148. **Erik Anderson.** *In situ* Scanning Tunnelling Microscopy studies of the interfacial structure between Bi(111) electrode and a room temperature ionic liquid. Tartu, 2015, 118 p.
149. **Girinath G. Pillai.** Computational Modelling of Diverse Chemical, Bio-chemical and Biomedical Properties. Tartu, 2015, 140 p.
150. **Piret Pikma.** Interfacial structure and adsorption of organic compounds at Cd(0001) and Sb(111) electrodes from ionic liquid and aqueous electrolytes: an *in situ* STM study. Tartu, 2015, 126 p.
151. **Ganesh babu Manoharan.** Combining chemical and genetic approaches for photoluminescence assays of protein kinases. Tartu, 2016, 126 p.
152. **Carolin Siimenson**. Electrochemical characterization of halide ion adsorption from liquid mixtures at Bi(111) and pyrolytic graphite electrode surface. Tartu, 2016, 110 p.
153. **Asko Laaniste.** Comparison and optimisation of novel mass spectrometry ionisation sources. Tartu, 2016, 156 p.