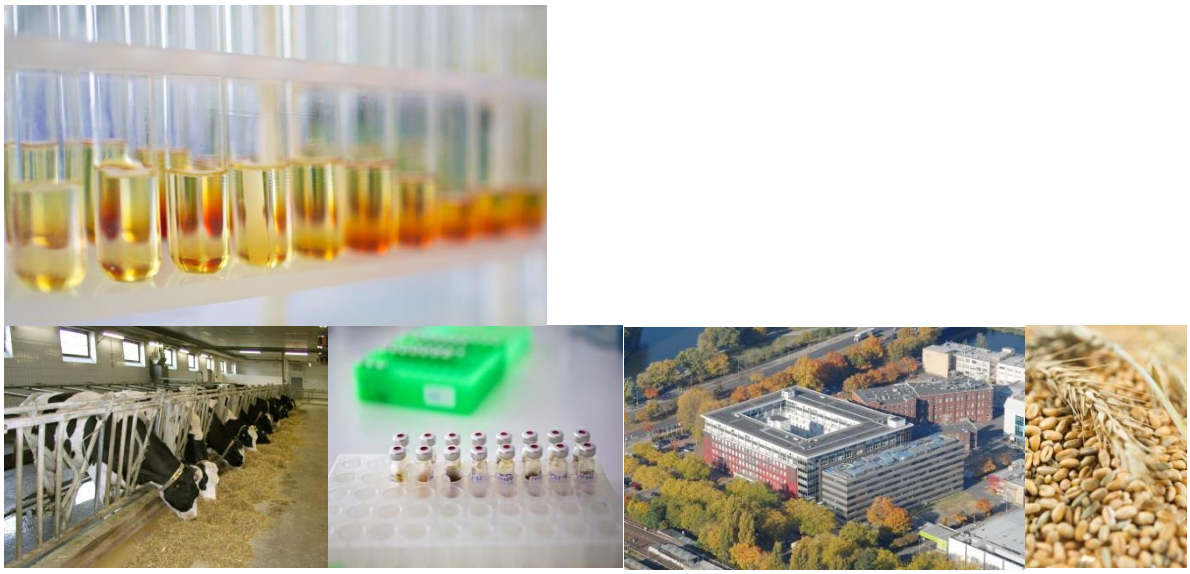
Bundesinstitut für Risikobewertung

Report on the 2022 Proficiency Test of the German Reference Laboratory for Mycotoxins and Plant Toxins

Determination of pyrrolizidine alkaloids and tropane alkaloids in herbs and spices



Impressum

BfR Wissenschaft

BfR-Autoren:
Anja These, Emanuele Pydde, Michael Weiß

Report on the 2022 Proficiency Test of the German Reference Laboratory for Mycotoxins and Plant Toxins. Determination of pyrrolizidine alkaloids and tropane alkaloids in herbs and spices

Herausgeber:
Bundesinstitut für Risikobewertung
Anstalt des öffentlichen Rechts
Vertreten durch den Präsidenten Professor Dr. Dr. Andreas Hensel
Max-Dohrn-Straße 8–10
10589 Berlin

Telefon: 030 18412 - 0
Telefax: 030 18412 – 99099

E-Mail: pressestelle@bfr.bund.de

Aufsichtsbehörde: Bundesministerium für Ernährung und Landwirtschaft
Ust.-IdNr. des BfR: DE 16589344

V.i.S.d.P: Dr. Suzan Fiack

Berlin 2023 (BfR-Wissenschaft 02/2023)
50 Seiten, 12 Abbildungen, 12 Tabellen

ISBN 978-3-948484-62-0
ISSN 1614-3795 (Print) 1614-3841 (Online)

DOI 10.17590/20230317-124743-0
Download als kostenfreies PDF unter www.bfr.bund.de

Content

1 Summary 5

2 Introduction 7

3 Scope and Study design 9

4 Test Material 11

4.1 Preparation procedure 11
4.2 Homogeneity and stability 12

5 Statistical evaluation 13

5.1 Pre-evaluation of data in combination of compliance test according to Mandel 13
5.2 Procedure for statistical evaluation 13

6 Proficiency test results 15

6.1 General overview 15
6.2 Evaluation of methods and performance 20
6.3 Consideration of the influence of homogeneity of samples on RSD_R and estimation of the influence of the sample size used for analysis on RSD_R in inhomogeneous samples 24
6.4 Pilot project on standard stability of PA and TA 26

7 Conclusions 38

8 References 40

9 Acknowledgement 40

10 Annex 41

10.1 Results of homogeneity testing 41
10.2 Results of statistical evaluation 41
10.3 Reported limits of quantification 43

11 List of Figures 44

12 List of Tables 45

1 Summary

The National Reference Laboratory for Mycotoxins and Plant Toxins organized a proficiency test (PT) for the determination of pyrrolizidine (PA) and tropane alkaloids (TA) in herbs and spices. Twenty-two laboratories from Germany participated, with a high level of expertise in PA/TA analysis. Three test materials (cumin, oregano, and parsley) and one standard solution were prepared to cover the entire analytical scope proposed to control maximum levels. The assigned values were calculated as robust mean values of laboratory results, and the performance of the laboratories was evaluated using z-scores, with a target standard deviation of 25 % (a 25 % deviation of a laboratory from the assigned value results in a |z| score of 1).

Across all proficiency test materials and laboratories, a total of 1082 z-score values were evaluated for PAs and 130 for TAs, of which 92 % for PAs and 95 % for TAs were satisfactory ($|z| \leq 2$). The relative reproducibility standard deviation (RSD_R) for total PA contents ranged from 9.3 to 13.0 % and was fully compliant with the RSD_R requirements specified in the draft regulation laying down methods for sampling and analysis for the official control of plant toxins (SANTE/11494R2/2021). Results of this PT indicate that the proficiency of the laboratories is satisfactory and the methods in use are fit for purpose to monitor maximum levels of PAs and TAs in herbs and spices.

In this PT, a procedure was proposed to evaluate the performance of a laboratory or method with respect to precision criteria under reproducibility conditions stipulated in the draft regulation. For this purpose, a pragmatic approach was used by calculating a mean z-score from the 95th percentile of absolute z-scores obtained by a laboratory among all analyte-matrix-combinations tested. This mean z-score value describes the deviation of the laboratory from the assigned values and is a meaningful measure for any bias and precision. Since the target standard deviation of 25 % for calculating z-scores reflects the requirement for the precision under reproducibility conditions specified in the draft regulation, a laboratory that achieves a mean z-score of $|z| \leq 1$ demonstrates that it meets the criteria required in the draft regulation (regardless of whether the method has been validated in a collaborative study or in-house). This was the case for 18 of 22 laboratories, while none of the laboratories exceeded a mean z-score of two, indicating that the 50 % value for the expanded measurement uncertainty generally would cover the inter-laboratory variability among participating laboratories.

Results of previous inter-laboratory tests have shown that despite thorough homogenization of samples to a particle size of approx. 500 µm, the (dry) plant test materials are still not sufficiently homogeneous. The objectives of this study included determining the extent to which heterogeneity impairs the reproducibility precision and adds to the analytical measurement uncertainty, and the extent to which increasing the sample amount used for analysis reduces the spread of analytical results.

One material (oregano) was shipped containing four analytes present due to natural contamination (heterogeneous) and 13 spiked analytes (homogeneous). In this way, the performance of the laboratories on the same material could be compared once for heterogeneous and once for homogeneous PA distributions. A rough estimate showed that the influence of inhomogeneity contributes on average to a 15 % higher deviation for inhomogeneous compared to homogeneous analyte-matrix-combinations. Furthermore, a naturally contaminated cumin sample (inhomogeneous) was sent. The laboratories were asked to analyze this sample once with a sample weight of 2 g and once with 10 g to evaluate the influence of sample size on reproducibility. Based on these inter-laboratory data, it was estimated that a fivefold increase in sample size (2 g and 10 g in this case) resulted in an absolute decrease of 8 % (relative decrease 21 %) in RSD_R.

In addition, this PT started a pilot study to test the storage stability of PA and TA standard solutions beyond the expiration date specified by the manufacturer. This standard will be

stored long-term and shipped regularly as part of the NRL's future PT program. Based on the expected assigned values obtained from the current and subsequent PTs, a trend for storage stability over the next few years can be derived.

2 Introduction

Plant toxins like 1,2-unsaturated pyrrolizidine alkaloids (PA) may act as genotoxic carcinogens in humans and could potentially present a risk of both acute and chronic effects in the consumer. To protect consumers and limit the exposure to pyrrolizidine alkaloids, maximum levels of pyrrolizidine alkaloids in certain foodstuffs such as dried herbs and seed spices entered into force on 1st of July 2022 (Commission Regulation (EU) 2020/2040 amending Regulation (EC) 1881/2006) [1]. The maximum levels refer to the lowerbound sum of 21 pyrrolizidine alkaloids and additionally their 14 naturally occurring isomers. For determination, these isomers can either be analysed as sum (in case of (partially) co-elution) or separated chromatographically, quantified individually and subsequently summed. The complete analytical scope for monitoring maximum levels is shown in Table 1.

Table 1: Analytical scope for monitoring the PA maximum levels in food. Natural occurring isomers can be summarized as group. The maximum level refers to the sum of the given PA and/or PA groups [1].

PA or PA group [abbreviation]	Ester form	Necine base	Natural isomers
Echimidine group [Em-G]	open chained diester	retronecine	Echimidine [Em], Heliosupine [Hs]
Echimidine-*N*-oxide group [EmN_G]	open chained diester	retronecine	Echimidine-*N*-oxide [EmN], Heliosupine-*N*-oxide [HsN]
Europine [Eu]	monoester	heliotridine	
Europine-*N*-oxide [EuN]	monoester	heliotridine	
Heliotrine [He]	monoester	heliotridine	
Heliotrine-*N*-oxide [HeN]	monoester	heliotridine	
Intermedine group [Im-G]	monoester	retronecine	Intermedine [Im] Lycopsamine [Ly] Indicine [Id] Echinatine [En] Rinderine [Rn]
Intermedine-*N*-oxide group [ImN-G]	monoester	retronecine	Intermedine-*N*-oxide [ImN] Lycopsamine-*N*-oxide [LyN] Indicine-*N*-oxide [IdN] Echinatine-*N*-oxide [EnN] Rinderine-*N*-oxide [RnN]
Lasiocarpine [Lc]	open chained diester	heliotridine	
Lasiocarpine-*N*-oxide [LcN]	open chained diester	heliotridine	
Retrorsine group [Re-G]	cyclic diester	retronecine	Retrorsine [Re] Usaramine [Us]
Retrorsine-*N*-oxide group [ReN-G]	cyclic diester	retronecine	Retrorsine-*N*-oxide [ReN] Usaramine-*N*-oxide (UsN]
Senecionine group [Sc-G]	cyclic diester	retronecine	Senecionine [Sc] Senecivernine [Sv] Integerrimine [Ig]
Senecionine-*N*-oxide group [ScN-G]	cyclic diester	retronecine	Senecionine-*N*-oxide [ScN] Senecivernine-*N*-oxide [SvN] Integerrimine-*N*-oxide [IgN]
Seneciphylline group [Sp-G]	cyclic diester	retronecine	Seneciphylline [Sp] Spartioidine [St]
Seneciphylline-*N*-oxide group [SpN-G]	cyclic diester	retronecine	Seneciphylline-*N*-oxide [SpN] Spartioidine-*N*-oxide [StN]
Senkirkine [Sk]	cyclic diester	otonecine	

A new regulation is to be established laying down methods for sampling and analysis for the official control of plant toxins. A draft is currently under discussion (SANTE/11494R2/2021 repealing Regulation (EU) No. 2015/705) [2]). The new regulation will also establish performance criteria for the methods to be used. With regard to the Limit of Quantification (LOQ) specific requirements are requested for the determination of PAs in dried plant-based products and methods shall have an LOQ of ≤ 10 µg/kg per individual PA. With regard to precision, a distinction is made between precision data that must be provided for methods within laboratories and in ring tests. Within-laboratory precision includes the repeatability relative standard

deviation (RSD_r) and the within-laboratory reproducibility relative standard deviation (RSD_{wR}) that both shall not exceed ≤ 20 %. A new feature of the regulation is the restriction of the relative reproducibility standard deviation (RSD_R). The RSD_R can be derived from either method validation studies or proficiency testing and should not exceed 25 %. Data obtained under reproducibility conditions such as the analysis of the same sample by different laboratories (inter-laboratory precision) indicate the spread of results for instance in the case of official cross-checks. A precondition for the evaluation of those performance characteristics is the supply of the same, sufficiently homogeneous, test material to each participating laboratory. Previous ring trails have demonstrated that in dried naturally contaminated plant samples PAs are still inhomogeneously distributed although the samples were carefully homogenized. Consequently, the RSD_R values obtained from those test materials were significantly higher than those from artificially contaminated or liquid test materials [3]. Several changes in sample preparation may reduce the spread of results. The most common ones are: increasing the amount of sample used for analysis or reducing the particle size by finer grinding. In order to estimate to what extent an increase in the sample weight can improve the RSD_R, data from inter-laboratory comparisons are helpful as a valid data basis. Therefore, participating laboratories were asked to analyse an inhomogeneous cumin sample with two different sample weights.

In addition, a pilot project was started in this PT, which is described in more detail in section 0, and which aims to investigate the stability of PAs during storage beyond the shelf-life guaranteed by the supplier.

To assess the **performance of laboratories and the precision of methods** for the control of plant toxins (draft regulation) for each laboratory a mean z-score was calculated (for more information please refer to section 6.2). The intent of the draft regulation, in terms of specifying performance criteria of methods and precision to be achieved, is to ensure that monitoring of maximum levels is based on reliable measurement data and that enforcement action is taken only when the exceedance of maximum level is beyond reasonable doubt. This is the case when the laboratory value minus the measurement uncertainty exceeds the maximum level. The prerequisite is that the (expanded) measurement uncertainty (MU) actually reflects the range of uncertainty, i.e. whether a result can be reproduced by another laboratory in the case of a reanalysis. In the field of residue analysis, a default MU of at least 50 % is required of a laboratory, and compliance can be demonstrated by a practical approach using results from proficiency testing. This is also discussed to apply in the field of mycotoxins and plant toxins and could be established by the following procedure.

The z-score represents the participant's deviation from the assigned value, and the target standard deviation was set at 25 % in this PT. The relative standard deviation of reproducibility (RSD_R) should be 25 % or better according to the draft regulation. Thus, a laboratory can demonstrate compliance with this criterion if the z-scores for the total PA content are ≤|1| and if a mean z-score of ≤|1| is achieved across all analyte-matrix-combinations.

No maximum levels are set for tropane alkaloids (TA) in herbs and spices [4]. But since PAs and TAs share the same way of contamination and thus similar analyte-matrix-combinations are affected and to be analysed, TAs were included in this PT.

3 Scope and Study design

This PT was organized to evaluate the proficiency of the laboratories and fitness for purpose of methods in use to determine PAs and TAs in herbs, spices and standard solutions. 22 laboratories active in food control, either as official laboratories of the federal states of Germany or as contract laboratories, participated and reported results (Table 2).

Three materials (parsley, oregano, and cumin) and a standard solution were sent with the documents on 05/01/2022, and the deadline for submitting the results was 07/31/2022. The standard solution contained individual PAs and TAs within a concentration range from 8 to 38 ng/mL and the mass fractions of individual toxins in matrix test materials ranged from 9 to 7000 µg/kg.
This PT did not aim at determining the influence of co-occurring natural isomers and their determination as a sum/group on the relative standard deviation (RSD_R). Therefore, only one isomer per isomeric group (Table 1) was spiked per material and standard. The detailed description of the samples and the spiking profile is given in section 0.

The cumin sample was naturally contaminated and therefore not sufficient homogeneous to fulfil criteria of a ring test material. This material was sent to estimate the influence of the sample size on the reproducibility of analytical results in the case of naturally contaminated samples. The laboratories were asked to analyse this cumin sample once at a sample weight of 2 g and once at 10 g. It was requested that the extraction volume had to be adjusted accordingly.

Finally, the objective of this PT was:

- to assess the fitness for purpose of the methods to control maximum level of PAs in herbs and spices and to comply with the required $RSD_R \leq 25\,\%$
- to estimate whether the increase of sample amounts for analysis decreases the spread of results for naturally contaminated (inhomogeneous) samples
- to start a long-term project to assess the storage stability of PAs and TAs

The laboratories that participated in this proficiency test, alphabetically listed in Table 2 below, are sincerely acknowledged.

Table 2: Participating laboratories

Bayrisches Landesamt für Gesundheit und Lebensmittelsicherheit
chelab Dr. V. Ara GmbH & Co. KG
Chemisches und Veterinäruntersuchungsamt Münsterland-Emscher-Lippe (CVUA-MEL) – AöR
Eurofins Dr. Specht International GmbH
Eurofins SOFIA GmbH
Eurofins WEJ Contaminants
GBA – Gesellschaft für Bioanalytik mbH
Institut Kirchhoff Berlin GmbH
Intertek Food Services GmbH
Labor Friedle GmbH
Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei Mecklenburg-Vorpommern
Landesamt für Verbraucherschutz Sachsen-Anhalt, Fachbereich Lebensmittelsicherheit
Landesbetrieb Hessisches Landeslabor
Landeslabor Berlin-Brandenburg
Landesuntersuchungsanstalt für das Gesundheits- und Veterinärwesen Sachsen
Landesuntersuchungsamt Rheinland-Pfalz
Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit
Nationales Referenzlabor für Mykotoxine und Pflanzentoxine in Lebens- und Futtermitteln
PhytoLab GmbH & Co. KG
Quality Services International GmbH, QSI
SGS Germany GmbH
Teekanne GmbH & Co. KG

4 Test Material

The participating laboratories received the following samples:

- Sample 1: oregano – appr. 25 g
- Sample 2: parsley – appr. 25 g
- Sample 3: cumin – appr. 30 g.
  This sample had to be divided by the participants in order to analyse the sample once with a sample amount of 2 g and once with a sample amount of 10 g.
- Standard solution in 5 % MeOH: appr. 1.0 ml
-

For the contamination profile, refer to Table 3

4.1 Preparation procedure

Homogenisation: The parsley and oregano materials were ground to a particle size of 500 µm using a centrifugal mill. The cumin sample was ground with liquid nitrogen (Grindomix 200, 2 x 10 seconds) and dried to constant mass. Particles above 1 mm were removed by sieving and the sample was shaken in a drum hoof mixer for two hours.

Pre-tests: The tested parsley material was shown to be free of analytes while the oregano was naturally contaminated with europine, lasiocarpine and their respective *N*-oxides. Both materials were spiked to obtain the profile as given in Table 3 using the procedure described below. The cumin sample contained a PA-profile typical for a contamination with *Heliotropium spp.* and was not spiked with any further analytes.

Spiking procedure: Test materials were spiked according to a protocol applied for the PT for "The determination of tropane alkaloids in herbal tea and herbal infusion" [5]. The parsley and oregano samples were spiked in the laboratory with a mixture of individual PAs as well as atropine and scopolamine (Table 3). For the spiking procedure, multi-analyte mixtures with target concentrations for each individual substance were prepared. For this purpose, defined volumes of stock solution of each substance were pipetted into an Erlenmeyer flask and diluted with tert-butylmethylether (TBME) for spiking procedure. Volumes of TBME required for complete wetting of materials were individually tested for parsley and oregano. Samples were divided into portions of approximately 25 g, mixed with the TBME multi-analyte solution by manual stirring under a fume hood, and then spread out in a flat aluminium tray to dry. Samples and TBME multi-analyte solution were effectively mixed by manual stirring in the fume cupboard and spread in a flat aluminium tray. After the solvent had evaporated, all subsamples were combined and mixed for about 12 hours in a Rhoenrad-mixer to obtain homogeneous test material. Table 3 shows the contamination profile of samples.

Aliquots of 25 g for parsley and oregano as well as 30 g for cumin were packed in 50 ml plastic tubes, closed with screw caps and additionally sealed with parafilm. The storage up to dispatch took place at room temperature. Standard solution (5 % methanol) was filled into 1.5 ml glass vials. The solutions were stored in the refrigerator at 5 °C until dispatch. The laboratories were asked to store the samples in a similar way until analysis.

Table 3: Contamination profile of proficiency test samples. Analytes were either spiked (+) or naturally contaminated (x)

Analyte/analyte-group	Standard solution	Oregano	Parsley	Cumin
Atropine	At (+)	At (+)	At (+)	-
Scopolamine	Sco (+)	Sco (+)	Sco (+)	-
Echimidine-group	Em (+)	Em (+)	Em (+)	-
Echimidine-*N*-oxide group	EmN (+)	EmN (+)	EmN (+)	HsN (x)*
Europine	Eu (+)	Eu (x)	Eu (+)	Eu (x)
Europine-*N*-oxide	EuN (+)	EuN (x)	EuN (+)	EuN (x)
Heliotrine	He (+)	He (+)	He (+)	He (x)
Heliotrine-*N*-oxide	HeN (+)	HeN (+)	HeN (+)	HeN (x)
Intermedine-group	Im (+)	Im (+)	Im (+)	Ec, Rn (x)*
Intermedine-*N*-oxide group	ImN (+)	ImN (+)	ImN (+)	EcN, RdN (x)
Lasiocarpine	Lc (+)	Lc (x)	Lc (+)	Lc (x)
Lasiocarpine-*N*-oxide group	LcN (+)	LcN (x)	LcN (+)	LcN (x)
Retrorsine-group	Re (+)	Re (+)	Re (+)	-
Retrorsine-*N*-oxide group	ReN (+)	ReN (+)	ReN (+)	-
Senecionine group	Sc (+)	Sc (+)	Sc (+)	-
Senecionine-*N*-oxide group	ScN (+)	ScN (+)	ScN (+)	-
Seneciphylline group	Sp (+)	Sp (+)	Sp (+)	-
Seneciphylline-*N*-oxide group	SpN (+)	SpN (+)	SpN (+)	-
Senkirkine	Sk (+)	Sk (+)	Sk (+)	-
concentration range	**8–38 ng/ml**	**55–445 µg/kg**	**23–1016 µg/kg**	**9–7000 µg/kg**

(*) trace amounts

4.2 Homogeneity and stability

The homogeneity of samples was determined according to ISO 13528 [6]. For this purpose, 10 units per test sample each were randomly selected and examined in duplicate analyses under repeatability conditions. The analyte distribution was considered as sufficiently homogeneous if at least the extended condition for homogeneity according to ISO 13528/point B.2.3 is fulfilled:

$$s_s = \sqrt{c} \qquad \text{(Equation 1)}$$

$$c = F_1 * (0.3 * \sigma_{PT})^2 + F_2 * s_w^2 \qquad \text{(Equation 2)}$$

s_s: *Standard deviation between samples*
σ_{PT}: *Target standard deviation*
s_w: *Standard deviation within the sample (duplicate analysis)*
F_1 und F_2: *from standard statistical tables, see ISO 13528*

The distribution of all analytes in parsley were found to be sufficiently homogenous according to the criteria of DIN 13528 (B.2.3). Oregano was naturally contaminated with europine, lasiocarpine and their corresponding *N*-oxides. As expected, the comminution to 500 µm did not result in a sufficiently homogenous material and consequently laboratory performance was not evaluated for these analyte-matrix-combinations. The cumin samples were only analysed for the scientific purpose to assess if an increase of sample amount reduces the spread of results between laboratories.

Results of homogeneity testing are summarised in Table 10 and Table 11 in the appendix. The stability was not tested, as previous ring trials did not provide any indication of instability.

5 Statistical evaluation

Twenty-two laboratories submitted results for the requested analytical scope of PAs and TAs (Table 4 to Table 8). Laboratory L-19 did not provide data for the standard solution and the cumin sample using 2 g sample weight. Laboratory L-12 and L-14 did not send results for cumin using 10 g of sample.

5.1 Pre-evaluation of data in combination of compliance test according to Mandel

A pre-evaluation of data was carried out to identify outlying laboratories. The identification criterion here was a systematic deviation of the laboratory from the other laboratories for the majority of analytes and test materials.

In addition to this visual evaluation, the Mandel's h-statistics was used. This test also evaluates deviations of the mean values of a single laboratory compared to the mean values of the other laboratories. The graphical reports of Mandel's h-statistics can be found in Figure 10 to Figure 12 in the appendix. Here the critical values for the significance level of 5 % are shown as a yellow line and for the significance level of 1 % as a red line. Statistically deviating values of the laboratories are accordingly marked as yellow bars (significance level 5 %) or red bars (significance level 1 %). Laboratory L-14 and L-15 had significant deviations in the majority of sample-analyte-combinations or analytes in all samples. Laboratory L-16 deviated for the majority of analytes in the standard solution. Data from laboratories L-14 and L-15 were excluded from the statistical evaluation to determine the assigned value and precision for all test materials and laboratory L-16 for the standard solution.

5.2 Procedure for statistical evaluation

The reported results of the participants were used for determining **the assigned value (consensus value of participants)**. This procedure was also used for the standard solution, i.e. the concentrations spiked by the organizer were only used as a comparison value. To minimize the effects of potential outliers on the assigned value as well as on precision data the evaluation was carried out by robust statistics according to ISO 13528 [6]. Robust statistics is recommended for data that are essentially normally distributed with a small proportion of strong outliers. In previous PTs the robust mean value and reproducibility standard deviation was determined according to the Q-Hampel method using the ProLabPlus software (version 2019.1.23.0). After intensive review of all previous results, it has become apparent that the calculation of the reproducibility standard deviation according to the Q-Hampel method can lead to results that are difficult to reproduce, since in individual cases the method tends to underestimate the actual scatter of the laboratory results. This could be due to the specified breaking point included in this method, which leads to further exclusion of laboratory results ("the rejection of outliers" is higher than with Huber). In order to gain more experience, this proficiency test was evaluated comparatively with two algorithms of robust statistics described and recommended in ISO 13528 [6]. In addition to the Q-Hampel method, the assigned value and reproducibility standard deviation were evaluated with the algorithm A described in ISO 13528. This method of robust statistics is also referred to as Huber estimator [7] and was calculated using the "hubers" function of the MASS package [8] in R [9].

Finally, the robust mean values as well as the robust reproducibility standard deviation according to Hampel and Huber were calculated.

For the **calculation of z-scores** a target standard deviation (σ_{PT}) of 25 % was applied in accordance to the drafted European regulation laying down the methods of sampling and analysis for the official control of the levels of plant toxins in foodstuffs [2].

The z-scores used for the assessment of laboratory performance were calculated according to equation (3)

$$z = \frac{(x_i - x_{PT})}{\sigma_{PT}} \qquad \text{(Equation 3)}$$

x_i: measurement result reported by the participant [µg/kg or ng/ml]
x_{PT}: assigned value [µg/kg or ng/ml]
σ_{PT}: target standard deviation [µg/kg or ng/ml]

For some analyte-matrix-combinations with an increased uncertainty of the calculated robust mean ($0.3 < u_x/\sigma_{PT}$) a z'-score according to ISO 13528 was determined using equation (4).

$$z' = \frac{(x_i - x_{PT})}{\sigma'_{PT}} \qquad \text{(Equation 4)}$$

x_i: measurement result reported by the participant [µg/kg or ng/ml]
x_{PT}: assigned value [µg/kg or ng/ml]
σ'_{PT} : expanded target deviation [µg/kg or ng/ml]

The z-score or z'-score (in cases $u_x/\sigma_{PT} > 0.3$) compares the participant's deviation from the reference value with the target standard deviation accepted for the proficiency test, σ_{PT}/σ'_{PT}.

The z-score/z'-score is interpreted as follows:

$\|z\| \leq 2$	result is considered to be acceptable
$2 < \|z\| < 3$	result is considered to be questionable (or warning signal)
$\|z\| \geq 3$	result is considered to be unacceptable (or action signal)

6 Proficiency test results

The statistical evaluation of the results was performed as described in section 5.2. Laboratory results and statistical characteristics for each sample are given in Table 4 to Table 8. A graphical overview on z-score results is shown in Figure 1 to Figure 2.

6.1 General overview

For none of the samples a certified content was available. The robust mean values were calculated from the laboratory results and used as assigned values for respective samples. A target standard deviation of 25 % was set and the performance of the laboratories was assessed using z-scores. A z-score of z equal to or less than |2| reflects a deviation of 50 % or less from the assigned value and indicates a satisfactory result for a laboratory.

Three test materials (cumin, oregano and parsley) and one standard solution were sent for analysis. The cumin sample was naturally contaminated and not homogeneous. This test material only was shipped to collect inter-laboratory data in order to evaluate, to which extent an increased amount of sample used for analysis can reduce RSD_R in inhomogeneous materials. No PA-free oregano material was available at the time of samples preparation, and the oregano used for this PT contained europine and lasiocarpine and their corresponding *N*-oxides. These analytes did not meet the homogeneity criteria for proficiency testing materials and were not used for laboratory evaluation. All other analytes in oregano were spiked and thus homogeneously distributed. Therefore, the presence of homogeneously and inhomogeneously distributed analytes in one sample analysed by the same group of laboratories was used to estimate the additive contribution of inhomogeneity towards the analytical measurement uncertainty of laboratories and therefore the spread of test results. The parsley sample was shown to be a PA-free material, all analytes were spiked and shown to be homogeneously distributed. The contamination profile for all analytes and materials is summarized in Table 3.

A **standard solution** was sent containing all PAs (only one representative per isomer group) and TAs regulated to monitor maximum levels in food (Table 3). For each laboratory, 18 analyte-matrix-combinations could be evaluated for PA analysis and two for TA analysis (Table 4). 96 % of z-scores for PAs and 84 % of z-scores for TAs fell in the acceptable range ($|z| \leq 2$). The relative reproducibility standard deviation (RSD_R) for individual PAs ranged from 7 % for heliotrine-*N*-oxide to 28 % for echimidine and was 11 % for the total PA concentration. A higher spread of results was obtained for atropine (tropane alkaloid) for which a RSD_R of 28 % was determined, while for scopolamine 14 % RSD_R was obtained (refer to Table 4).

With regard to the PA concentrations spiked by the organizer, the obtained assigned values represent recoveries between 85 and 100 % (mean: 93 %). This sufficient recovery was interpreted as an indication that the assigned values, which were derived as the consensus of participants, were more or less unbiased. For a more detailed evaluation regarding the storage stability and potential influence of supplier, please refer to section 6.4.

Table 4: Analytical results and statistical characteristics for the standard solution

	Eu	EuN	Hn	HnN	Lc	LcN	Sk	Em-G	EmN-G	Im-G	ImN-G	Re-G	ReN-G	Sc-G	ScN-G	Sp-G	SpN-G	PA-sum	At	Sco
	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml	ng/ml
L-01	19.4	34.4	30.5	17.2	20.1	13.1	8.8	7.1	14.5	34.1	10.3	30.5	16.6	33.6	14.5	27.3	17.2	348.9	13.7	27.1
L-02	22.2	37.5	38.7	17.5	24.9	13.9	9.1	9.5	16.1	39.5	9.8	33.4	20.5	43.7	14.6	29.8	19.6	400.4	4.0	8.3
L-03	22.3	34.7	25.7	15.8	19.7	12.2	8.4	7.1	13.7	32.8	9.4	26.4	15.7	29.7	13.4	22.2	16.9	326.0	10.9	26.8
L-04	24.1	46.6	34.4	18.6	29.3	14.8	10.4	11.1	17.0	52.1	11.9	36.5	20.8	42.2	19.3	33.5	22.4	444.9	19.5	39.2
L-05	17.2	36.9	30.2	16.2	18.3	13.1	8.8	7.8	15.1	35.3	9.7	31.2	19.9	31.4	13.7	30.1	17.3	352.3	12.8	28.2
L-06	21.5	40.8	34.6	16.6	23.8	13.3	10.2	12.2	14.3	36.1	5.7	34.0	23.4	34.0	14.4	31.9	18.5	385.0	27.4	32.2
L-07	22.7	42.0	33.9	15.6	26.1	14.2	10.1	11.0	15.8	37.0	7.7	39.9	18.7	41.8	13.8	42.1	17.8	410.2	18.9	42.9
L-08	18.5	37.1	32.3	17.0	23.3	14.0	9.4	7.0	14.5	35.2	8.4	32.4	18.6	38.1	16.5	27.9	18.6	368.7	18.3	32.3
L-09	20.9	35.3	28.5	14.9	20.7	13.7	8.6	7.5	13.8	33.6	9.6	28.3	17.1	36.3	14.9	25.5	16.7	345.8	19.5	32.8
L-10	18.1	35.0	28.1	17.9	16.1	11.9	8.7	6.8	13.9	33.9	9.3	29.0	17.5	32.6	13.0	26.5	17.0	335.1	16.3	34.2
L-11	23.9	34.5	28.2	15.2	23.8	13.8	9.0	12.3	14.1	36.0	9.7	29.9	21.0	29.0	15.5	24.2	19.9	360.0	15.4	33.3
L-12	18.9	37.1	36.6	16.5	23.4	13.7	9.9	9.6	15.6	40.1	10.8	34.6	19.9	49.8	15.5	31.2	22.2	405.3	25.2	37.9
L-13	29.7	49.8	40.0	20.1	30.3	15.7	11.6	11.9	26.2	50.1	9.9	43.7	< 2	45.3	18.5	37.2	24.5	464.5	10.0	26.8
L-14	25.8	42.1	46.1	18.3	66.1	23.7	16.3	5.6	16.5	15.3	3.4	18.0	14.5	17.6	8.4	19.5	11.0	368.1	7.6	13.8
L-15	13.9	19.9	26.2	10.5	12.3	5.4	6.6	7.2	13.6	7.9	1.4	19.6	11.2	23.6	6.7	22.3	11.2	219.5	12.5	25.1
L-16	11.3	20.9	17.3	9.4	12.0	7.0	4.8	4.4	10.5	14.4	6.6	17.8	8.8	18.8	8.0	15.4	10.4	197.5	11.0	17.4
L-17	17.0	30.6	32.7	17.1	21.8	13.5	9.1	5.9	17.4	36.7	10.1	33.2	16.3	73.5	14.3	37.4	18.6	405.2	17.5	32.4
L-18	18.7	33.6	28.8	15.3	19.3	12.2	8.3	6.9	13.1	33.9	8.8	29.4	18.5	30.0	13.4	11.7	18.4	320.3	13.5	33.7
L-20	24.2	39.6	29.5	15.7	22.4	13.8	8.7	7.1	14.5	34.5	9.4	31.9	16.9	33.1	14.1	27.6	16.8	359.8	0.8	1.8
L-21	19.8	38.5	30.6	16.2	20.8	13.1	8.6	7.1	13.6	36.9	4.4	17.7	21.2	35.1	7.5	15.2	9.1	315.5	15.8	35.3
L-22	25.4	40.4	31.0	16.6	21.8	16.1	9.4	9.6	15.8	37.9	11.6	37.5	22.2	42.9	16.1	32.5	21.5	408.1	21.5	34.7
No. of labs	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
No. of evaluated results	18	18	18	18	18	18	18	18	18	18	18	18	17	18	18	18	18	18	16	16
rel. target std. dev [%]	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
assigned value Hampel [ng/ml]	21.2	37.6	31.8	16.6	22.4	13.6	9.2	8.5	14.8	35.8	9.6	32.4	19.1	36.9	14.8	28.9	18.9	373.8	17.1	32.9
rel. reprod. std. dev. Hampel [%]	**15.7**	**12.5**	**14.0**	**7.4**	**16.0**	**7.18**	**7.8**	**18.1**	**9.9**	**8.4**	**12.4**	**16.2**	**13.8**	**20.1**	**11.6**	**26.5**	**12.8**	**12.0**	**32.9**	**10.5**
assigned value Huber [ng/ml]	21.2	37.6	31.7	16.5	22.3	13.6	9.2	8.7	15	36.4	9.5	32.3	19.1	37.6	14.7	29	18.7	373.8	17	32.9
rel. reprod. std. dev. Huber [%]	**15.2**	**10.4**	**12.4**	**7.1**	**14.7**	**7.6**	**8.6**	**27.9**	**9.8**	**7.8**	**14.4**	**14.1**	**12.9**	**19.7**	**10.5**	**22.7**	**13.4**	**12.0**	**28.7**	**14.3**

In **oregano**, for PA analysis 14 analyte-matrix-combinations could be evaluated per laboratory and two for TA analysis. 89 % of z-scores for PA and 82 % of z-scores for TA fell in the acceptable range ($|z| \leq 2$) (refer to Figure 1 or Table 5). The success rate of laboratories (percentage of z-score values $|z| \leq 2$) varied from 0 to 100 % for TAs and from 50 to 100 % for PAs. RSD_R values across all individual PAs and the concentration range tested spanned from 12 % for senkirkine to 30 % for seneciphylline and were 13 % for total PA content. Higher RSD_R values were obtained for the naturally present analytes such as europine, lasiocarpine and their respective *N*-oxides, ranging from 34 % for europine to 71 %for lasiocarpine-*N*-oxide.

In **parsley**, for PA analysis 18 analyte-matrix-combinations could be evaluated per laboratory and two for TA analysis. 92 % of z-scores for PA and 89 % of z-scores for TA fell in the acceptable range ($|z| \leq 2$) (refer to Figure 2 or Table 6). The success rate of laboratories varied from 0 to 100 % for TAs and from 50 to 100 % for PAs. The RSD_R values ranged for the tested contents from 13 % for senkirkine to 30 % for echimidine and was 13 % for the total PA content.

Table 5: Analytical results and statistical characteristics for oregano

	Eu+	EuN+	Hn	HnN	Lc+	LcN+	Sk	Em-G	EmN-G	Im-G	ImN-G	Re-G	ReN-G	Sc-G	ScN-G	Sp-G	SpN-G	PA-sum	At	Sco
	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg
L-01	87.4	426.9	207.7	327.0	32.3	111.1	212.1	135.7	1112.0	168.0	287.5	26.3	38.2	302.4	552.9	46.8	75.4	3492.0	160.9	434.4
L-02	74.0	240.0	231.0	280.0	22.0	43.0	223.0	170.0	931.0	164.0	245.0	19.0	34.0	301.0	569.0	41.0	60.0	3268.0	207	383
L-03	81.4	196.5	214.7	297.0	17.0	13.4	208.4	180.6	990.3	190.7	237.4	43.8	39.5	443.2	505.1	79.6	86.6	3516.8	220.8	466.8
L-04	118.0	468.0	259.0	372.0	34.0	54.0	324.0	185.0	1352.0	213.0	342.0	33.0	37.0	492.0	750.0	87.0	100.0	4546.0	252	662
L-05	118.0	356.9	202.5	197.7	57.5	60.3	192.9	143.2	894.3	158.4	83.7	26.8	27.6	233.2	391.9	52.1	58.3	2662.4	137.8	365.6
L-06	70.2	232.7	213.1	285.0	17.8	17.8	252.1	283.3	1075.2	178.5	175.8	26.0	53.8	284.1	706.0	56.0	96.0	3684.9	395.7	406.8
L-07	72.3	121.9	208.8	240.0	24.6	13.4	219.3	187.3	965.0	167.2	163.5	26.8	46.7	288.5	394.2	75.5	55.4	3038.1	148.6	218.0
L-08	191.0	479.0	257.0	305.0	81.4	38.3	269.0	182.0	928.0	174.0	239.0	39.1	46.7	342.0	522.0	50.2	81.4	3435.4	495	1010
L-09	104.2	225.6	180.2	258.3	16.8	19.3	201.3	137.9	994.4	146.1	258.8	25.4	32.0	300.7	576.9	42.8	78.1	3232.9	246.4	585.3
L-10	53.8	167.9	215.6	321.3	18.4	15.6	245.3	152.8	871.8	156.2	265.3	29.2	39.4	392.0	560.5	58.2	68.9	3376.5	228.4	534.3
L-11	44.3	190.0	195.0	246.3	16.3	19.3	233.8	201.3	785.0	177.5	240.0	24.1	47.4	286.3	621.3	40.5	75.3	3173.6	270	505
L-12	76.4	317.3	269.4	384.8	34.1	44.4	221.6	216.1	1005.6	197.5	264.3	23.6	34.4	302.9	587.6	50.7	65.7	3624.2	212.0	452.88
L-13	52.1	113.5	196.5	221.0	8.2	8.5	252.2	123.1	1121.0	127.2	177.8	28.4	< 5	206.2	456.5	45.1	72.2	3027.2	252.1	406.3
L-14	47.0	144.0	187.0	275.0	16.0	12.0	210.0	60.0	687.0	39.0	63.0	17.0	23.0	104.0	161.0	18.0	36.0	1880.0	204	398
L-15	94.2	145.6	155.7	169.9	15.9	7.4	139.7	138.0	769.0	41.4	49.8	20.1	21.1	122.0	136.9	32.0	41.5	1837.0	113.8	233.5
L-16	74.8	315.7	263.3	274.8	37.5	34.3	239.8	180.0	1545.9	124.2	287.8	52.4	54.5	546.8	642.6	81.8	96.0	4389.9	354.2	483.1
L-17	73.6	280.0	219.2	311.8	16.8	20.0	249.6	108.1	1375.6	166.9	266.2	30.0	35.1	284.1	259.5	74.1	70.9	3451.0	307.7	604.02
L-18	80.6	242.8	217.5	332.2	34.9	18.9	241.6	143.0	1152.8	175.8	266.1	22.3	29.3	307.2	576.7	22.8	70.7	3557.9	190.5	423.5
L-19	79.0	200.0	210.0	280.0	15.0	13.0	220.0	140.0	920.0	180.0	250.0	<20	<20	150.0	270.0	62.0	73.0	2755.0	230	450
L-20	50.5	174.5	188.3	270.5	11.7	10.2	205.2	137.0	931.5	147.8	244.6	26.5	38.6	290.3	537.5	49.1	70.4	3137.3	299.9	633.4
L-21	127.4	533.5	273.3	371.8	16.6	22.4	275.0	183.0	773.1	201.3	510.3	32.2	48.8	358.3	799.8	67.8	84.9	3979.5	26.41	57.91
L-22	59.8	259.9	237.7	363.0	< 30.0	65.2	233.0	233.3	1112.8	101.4	377.1	27.2	48.1	339.2	606.5	62.0	108.0	3849.3	297.1	515.2
No. of labs	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
No. of evaluated results	20	20	20	20	20	20	20	20	20	20	20	20	19	20	20	20	20	20	20	20
rel. target std. dev [%]	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
assigned value Hampel [µg/kg]	76.8	270	222	297.6	22.8	25.7	233.4	167.7	1016.2	167	248.6	27.7	40.6	304.6	550.4	57.5	76.8	3430.9	244.4	469.7
rel. reprod. std. dev. Hampel [%]	**26.6**	**43.4**	**11.8**	**19.6**	**46.8**	**51.4**	**12.3**	**22.4**	**16.3**	**15.7**	**19.2**	**19.3**	**23.9**	**16.8**	**22.1**	**29.1**	**18.4**	**14.1**	**38.0**	**28.2**
assigned value Huber [µg/kg]	80.8	272.2	222.8	297.6	24.1	29.3	233.4	168	1020.2	166.9	253.2	28.3	40.6	315.2	550.3	57.5	76.9	3427.8	243.6	477.7
rel. reprod. std. dev. Huber [%]	**34.4**	**46.1**	**14.1**	**19.2**	**50.3**	**70.7**	**11.7**	**23.0**	**17.0**	**16.1**	**24.7**	**18.8**	**22.6**	**23.8**	**24.4**	**30.0**	**19.7**	**13.0**	**34.4**	**26.5**

+ due to inhomogeneity this analyte was not used for laboratory evaluation

Table 6: Analytical results and statistical characteristics for parsley

	Eu	EuN	Hn	HnN	Lc	LcN	Sk	Em-G	EmN-G	Im-G	ImN-G	Re-G	ReN-G	Sc-G	ScN-G	Sp-G	SpN-G	PA-sum	At	Sco
	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg
L-01	238.4	168.3	171.0	76.2	122.8	219.1	283.3	41.5	207.7	477.7	161.4	284.2	80.1	104.9	190.3	194.7	222.3	3243.9	3433.1	76.5
L-02	227.0	142.0	174.0	58.0	130.0	135.0	277.0	58.0	177.0	451.0	121.0	246.0	74.0	126.0	127.0	190.0	182.0	2895.0	3023.0	63.0
L-03	277.3	181.8	150.1	75.5	123.4	150.6	301.5	42.4	198.3	467.0	160.7	276.9	82.8	121.9	176.0	192.8	194.3	3173.1	3173.0	89.7
L-04	327.0	264.0	205.0	87.0	160.0	218.0	429.0	63.0	246.0	582.0	200.0	339.0	80.0	191.0	260.0	333.0	286.0	4270.0	4270.0	109.0
L-05	141.1	79.4	134.8	31.0	99.1	69.6	236.2	50.4	119.7	317.0	64.0	243.6	42.5	144.2	90.2	179.2	112.4	2153.9	2153.9	73.4
L-06	244.2	139.9	169.7	59.5	147.8	124.0	312.1	178.0	155.1	442.8	70.5	257.0	104.1	115.4	117.8	236.7	161.4	3036.0	3276.3	157.1
L-07	216.5	148.0	158.9	59.3	128.4	166.7	291.6	51.2	215.8	424.4	102.0	316.3	76.9	120.2	145.5	255.8	191.1	3068.4	3260.3	87.7
L-08	246.0	179.0	160.0	76.8	93.8	156.0	309.0	43.6	210.0	411.0	149.0	310.0	99.4	109.0	277.0	211.0	267.0	3307.6	3470.0	52.6
L-09	354.9	164.7	152.9	56.6	102.1	150.4	274.0	68.3	197.5	407.9	132.4	287.2	62.1	155.9	172.9	179.5	181.7	3101.0	3047.9	98.6
L-10	201.3	154.8	142.2	74.9	135.1	124.3	266.8	43.1	171.4	415.4	153.0	264.8	81.3	103.3	181.8	184.5	217.8	2915.8	2915.8	85.2
L-11	252.5	177.5	180.0	67.3	107.3	155.0	312.5	66.3	217.5	527.5	150.0	272.5	115.0	115.0	200.0	157.5	265.0	3338.3	3338.3	99.5
L-12	243.4	179.2	210.1	79.6	123.0	120.0	276.5	81.9	145.0	492.9	150.9	273.3	85.1	159.9	131.4	264.8	179.9	3196.9	3196.9	107.2
L-13	170.3	153.3	131.7	49.6	90.2	122.5	207.6	40.6	254.2	395.8	108.7	225.8	< 5	82.8	131.4	183.5	171.1	2519.1	2519.0	113.5
L-14	188.0	116.0	174.0	34.0	191.0	25.0	340.0	67.0	48.0	110.0	17.0	180.0	11.0	61.0	7.0	144.0	15.0	1728.0	1728.0	81.0
L-15	150.0	90.6	117.2	53.8	66.0	77.2	196.2	45.5	198.0	88.1	34.6	147.0	62.9	67.0	63.9	126.6	126.4	1711.0	1130.9	42.1
L-16	226.2	160.5	164.0	65.7	131.9	139.8	283.1	61.6	286.8	311.0	147.6	213.4	80.2	117.7	148.4	223.2	188.2	2949.3	3124.3	95.9
L-17	200.3	147.9	197.6	83.2	116.0	153.6	360.1	33.6	259.4	508.6	162.5	305.0	89.1	228.2	180.5	290.8	260.3	3576.7	3576.7	132.6
L-18	229.4	171.5	170.9	68.2	126.5	161.3	331.3	52.3	217.7	478.7	144.0	266.3	96.2	104.7	171.7	91.8	256.5	3138.8	3292.2	89.2
L-19	260.0	140.0	170.0	52.0	190.0	54.0	260.0	130.0	100.0	510.0	89.0	350.0	24.0	190.0	46.0	320.0	59.0	2944.0	2855.0	86.0
L-20	239.6	163.7	140.4	63.1	98.1	182.8	284.1	38.0	217.8	395.8	145.3	226.8	92.4	82.2	213.6	134.5	264.5	2982.7	2982.7	135.9
L-21	286.3	207.6	188.1	75.1	118.3	162.8	343.1	50.4	228.6	609.7	250.3	341.1	100.9	107.2	224.8	248.3	277.4	3820.1	3820.1	9.2
L-22	312.9	197.2	207.3	64.8	126.5	156.6	316.7	91.0	190.0	259.9	82.5	296.5	109.5	146.8	156.8	256.6	229.5	3201.0	3201.0	138.4
No. of labs	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
No. Of evaluated results	20	20	20	20	20	20	20	20	20	20	20	20	19	20	20	20	20	20	20	20
rel. target std. dev [%]	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
assigned value Hampel [µg/kg]	243.8	164.5	168.6	66.8	120.3	147.7	294.5	54.6	202.4	445.4	134.1	279.6	85.7	125.4	167.3	216.4	213.3	3117.3	96.3	99.3
rel. reprod. std. dev. Hampel [%]	**19.7**	**17.0**	**14.3**	**23.2**	**15.6**	**20.7**	**13.7**	**35.8**	**22.5**	**20.5**	**24.7**	**16.0**	**22.1**	**23.1**	**33.7**	**29.2**	**24.4**	**9.8**	**30.9**	**22.7**
assigned value Huber [µg/kg]	243.8	165	168.9	67.0	121.3	147.6	295.2	57.3	202.4	445.2	135	279.6	85.4	127.7	166.9	216.1	212.8	3128.6	96.3	98
rel. reprod. std. dev. Huber [%]	**19.7**	**14.4**	**15.9**	**19.0**	**17.3**	**19.9**	**13.0**	**34.2**	**22.6**	**19.8**	**31.3**	**15.7**	**21.3**	**26.9**	**30.6**	**27.9**	**25.9**	**9.3**	**31.7**	**24.7**

Table 7: Analytical results and statistical characteristics for cumin applying a sample amount of 2 g for analysis

	Eu[a]	EuN[a]	Hn[a]	HnN[a]	Lc[a]	LcN[a]	EmN-G[a]	Im-G[a]	ImN-G[a]	PA-sum
	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg
L-01	138.7	1616.0	764.5	5842.0	183.6	1801.0	12.2	11.1	66.5	10437.4
L-02	156.0	1572.0	1018.0	5743.0	225.0	1230.0	9.0	7.0	72.0	10032.0
L-03	194.9	2806.1	904.2	10775.8	286.5	2809.0	< 10	< 10	147.3	17923.9
L-04	209.0	2780.0	971.0	7290.0	242.0	1927.0	11.0	11.0	64.0	13505.0
L-05	99.9	768.5	810.0	4305.8	177.9	1103.3	< 10	< 10	39.3	7304.6
L-06	192.3	2305.0	1005.7	7323.8	237.9	1488.5	12.8	10.1	82.2	12658.4
L-07	195.9	2584.8	1095.6	9227.4	266.4	2111.3				15481.4
L-08	116.0	1150.0	675.0	4660.0	122.0	721.0	15.9	7.4	85.8	7550.0
L-09	126.0	1090.6	503.0	3129.9	114.3	841.8	13.0	< 1	48.8	5867.4
L-10	164.3	1839.8	914.4	7550.1	207.1	1353.3	14.8	10.1	74.4	12128.3
L-11	120.0	1450.0	595.0	4700.0	160.0	1050.0	7.3	6.3	50.0	8138.6
L-12	170.4	1593.0	1103.5	5419.7	202.6	1127.6	< 2	13.2	90.9	10270.6
L-13	90.3	2113.0	611.0	7711.0	62.7	1118.0	< 5	< 5	48.8	11987.0
L-14	144.0	1705.0	790.0	5631.0	262.0	3551.0	2.0	1.0	10.0	12104.0
L-15	84.8	953.0	580.0	4425.0	109.5	527.0	< 0.5	< 0.5	< 0.50	6679.3
L-16	161.6	2202.8	968.9	6422.5	155.9	765.9	12.0		84.0	10773.6
L-17	161.6	2898.7	1318.9	15048.1	204.6	2126.8				21758.7
L-18	149.6	2254.0	916.3	7451.4	196.1	1263.6	12.6	8.7	93.6	12346.0
L-19										
L-20	129.6	1787.3	689.6	7795.8	210.4	1548.8	9.2	< 6	57.5	12228.1
L-21	178.8	1437.6	669.4	5636.5	207.0	1177.6	9.6	7.0	53.4	9376.9
L-22	350.5	3529.6	1357.0	9715.5	289.4	3139.1	< 30	< 30	< 30	18381.1
No. of labs	21	21	21	21	21	21	21	21	21	21
No. of evaluated results	19	19	19	19	19	19	12	10	16	19
rel. target std. dev [%]	25	25	25	25	25	25	25	25	25	25
assigned value Hampel [µg/kg]	154.1	1973.3	878.8	6906.5	199.8	1447.5	11.6	9.1	68.8	11798.3
rel. reprod. std. dev. Hampel [%]	**31.0**	**42.9**	**26.5**	**37.8**	**30.6**	**42.6**	**26.7**	**25.9**	**32.2**	**35.5**
assigned value Huber [µg/kg]	156.7	1969.5	880.2	6891.5	200	1441.8	11.6	9.2	69.5	11761.8
rel. reprod. std. dev. Huber [%]	**26.2**	**38.0**	**28.1**	**33.6**	**29.8**	**40.6**	**24.2**	**27.2**	**30.7**	**34.2**

[a] this sample was inhomogeneous and was analysed as research sample but not for laboratory evaluation

Table 8: Analytical results and statistical characteristics for cumin applying a sample amount of 10 g for analysis

	Eu [a]	EuN [a]	Hn [a]	HnN [a]	Lc [a]	LcN [a]	EmN-G [a]	Im-G [a]	ImN-G [a]	PA-sum [a]
	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg	µg/kg
L-01	193.4	2557.0	880.0	9090.0	228.4	2877.0	20.9	13.7	99.5	15962
L-02	213.0	2679.0	1212.0	8291.0	268.0	1685.0	19.0	11.0	92.0	14470
L-03	185.0	2213.3	818.8	7756.8	211.8	1468.4	< 10	< 10	131.1	12793
L-04	201.0	2778.0	917.0	7256.0	224.0	2032.0	11.0	13.0	77.0	13509
L-05	102.1	704.1	769.3	3740.9	170.1	921.7	< 10	< 10	42.0	6450
L-06	193.3	1891.8	1039.2	6659.9	235.0	1285.1	10.6	< 8.0	73.2	11388
L-07	211.4	2072.7	984.9	7202.4	272.2	2003.9				12747
L-08	139.0	1720.0	796.0	5940.0	32.1	1240.0	23.6	7.9	80.2	9980
L-09	171.2	1164.7	553.6	3981.6	133.6	1021.4	10.8	< 2.5	42.6	7079
L-10	118.8	1510.2	757.5	5874.8	176.1	994.3	< 10	< 10	65.4	9497
L-11	480.0	2300.0	845.0	7300.0	200.0	1550.0	30.0	9.6	92.0	12807
L-12										
L-13	105.7	2147.0	672.0	6842.0	64.9	1177.0	< 5	< 5.00	40.9	11322
L-14										
L-15	189.0	1691.0	1309.0	7308.0	152.0	836.0	< 0.5	< 0.5	< 0.5	11485
L-16	201.3	2273.9	1128.8	7731.8	192.8	1282.2	14.4	8.1	88.6	12925
L-17	156.9	2265.1	1311.2	11970.9	225.0	2117.9				18047
L-18	133.1	2157.6	844.6	8832.9	189.1	1416.2	7.9	5.3	56.7	13643
L-19	170.0	1800.0	800.0	5300.0	160.0	900.0	< 10	< 10	< 10	9130
L-20	141.4	1665.6	726.6	7919.5	188.4	1439.8	9.4	< 6	54.0	12145
L-21	151.1	1311.6	576.6	4670.0	178.1	871.9	8.8	8.2	69.3	7846
L-22	361.8	3759.8	1338.9	9453.7	252.3	2625.8	< 30	< 30	< 30	17792
No. of labs	20	20	20	20	20	20	20	20	20	20
No. of evaluated results	19	19	19	19	19	19	11	8	15	19
rel. target std. dev [%]	25	25	25	25	25	25	25	25	25	25
assigned value Hampel [µg/kg]	165.8	2029.8	879.6	7064.8	196.9	1475.6	14.5	9.6	72.4	12016
rel. reprod. std. dev. Hampel [%]	**30.5**	**35.8**	**24.2**	**29.5**	**27.3**	**37.9**	**41.3**	**36.7**	**36.5**	**27.9**
assigned value Huber [µg/kg]	172.6	2029.8	884	7073.8	197.1	1470.9	14.7	9.6	72.2	12034
rel. reprod. std. dev. Huber [%]	**28.2**	**29.2**	**26.6**	**27.7**	**26.0**	**36.2**	**49.3**	**33.4**	**34.2**	**29.4**

[a] this sample was inhomogeneous and was analysed as research sample but not for laboratory evaluation

z-scores of all participants

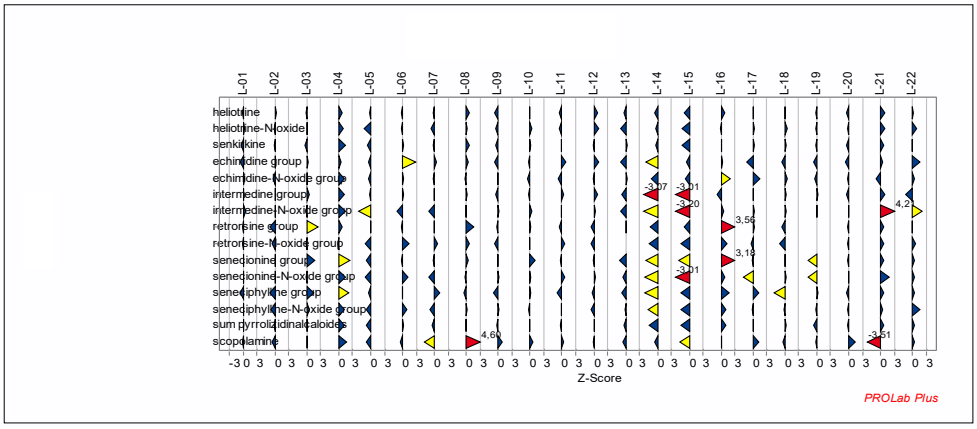


Figure 1: z-score results for oregano, rel. target standard deviation is 25 % and corresponds to |z| score = 1 (blue triangle: |z| score ≤ 2, yellow triangle: 2 < |z| score < |3|, red triangle: |z| score ≥ 3)

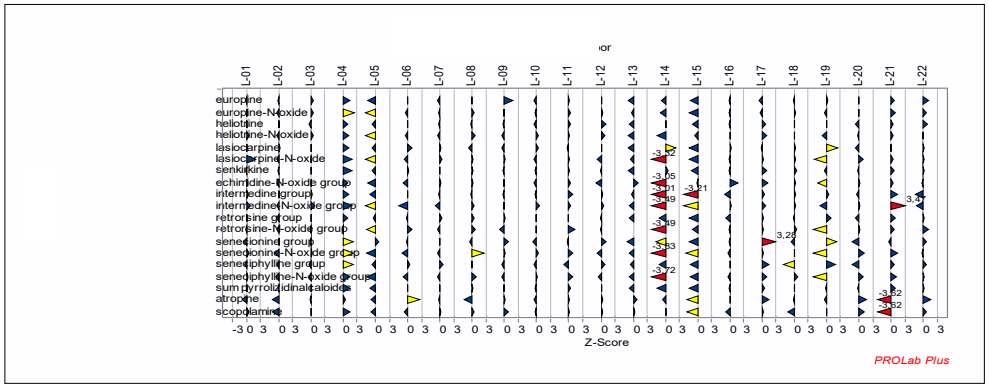


Figure 2: z-score results for parsley, rel. target standard deviation is 25 % and corresponds to |z| score = 1 (blue triangle: |z| score ≤ 2, yellow triangle: 2 < |z| score < |3|, red triangle: |z| score ≥ 3)

Across **all proficiency test materials and laboratories**, a total of 1082 z-score values were evaluated for PAs and 130 for TA. Among them, 92 % of PA z-scores were satisfactory and 95 % for TAs. Nine of the 47 individual PA-matrix-combination values exceeded the RSD_R value of 25 %. To determine if, within the analytical scope, some PAs tend to result in higher RSD_R values than others, the mean RSD_R values were calculated across the standard solution, parsley and oregano (excluding inhomogeneous PAs).

As shown in Figure 3, there was a range among PAs, from 11 % for senkirkine to 28 % for echimidine. Although it should be noted that the number of materials tested was limited, there was some tendency for PAs from isomeric groups to have higher RSD_R values than those without isomers. Since the PT samples contained only one isomer per group, these group representatives have long been practiced in PA analysis, and no additional isomer was added, reasons other than “analytical challenges due to differential isomer response” appear to contribute to this observation. For example, all samples contained echimidine but no heliosupine (Table 1 and Table 3). This means that the higher RSD_R value of echimidine (Em in Figure 3) cannot be explained by the different MS response of the (partially) co-eluting isomer heliosupine, which could lead to an incorrect quantification. Rather, it suggests that many laboratories attempt to analyse the isomers separately to avoid these potential errors, which in turn can lead to other challenges. Potential pitfalls such as double identification of a single isomer should be mentioned here. Regarding the example of the echimidine group, there is a possibility that echimidine can be identified once as echimidine and additionally as its closely eluting isomer heliosupine resulting in a highly overestimated echimidine-group content.

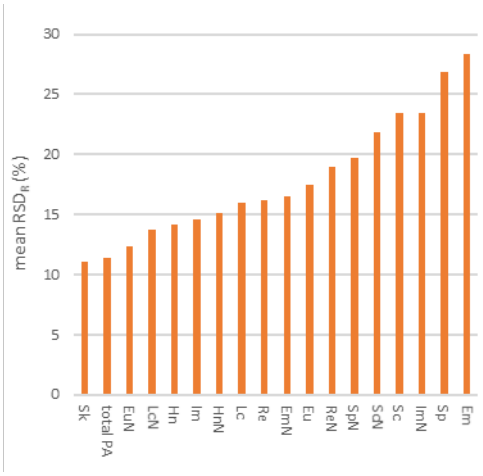


Figure 3: Mean RSD_R [%] for each analyte and the total PA content in all materials (only homogeneous analyte-matrix-combinations)

In Figure 4 the RSD_R obtained for all tested analyte-matrix-combinations is shown in relation to the respective concentration (assigned value). These results confirm the outcome in previous PTs and indicate that there is no concentration dependence in RSD_R and there is no evidence that higher RSD_R values are obtained at lower concentrations. Further, there is no indication for a matrix specific increase or decrease for RSD_R although slightly higher values were obtained for cumin and some analytes in oregano. This is explained by the inhomogeneity and will be considered in more detail in section 6.3.

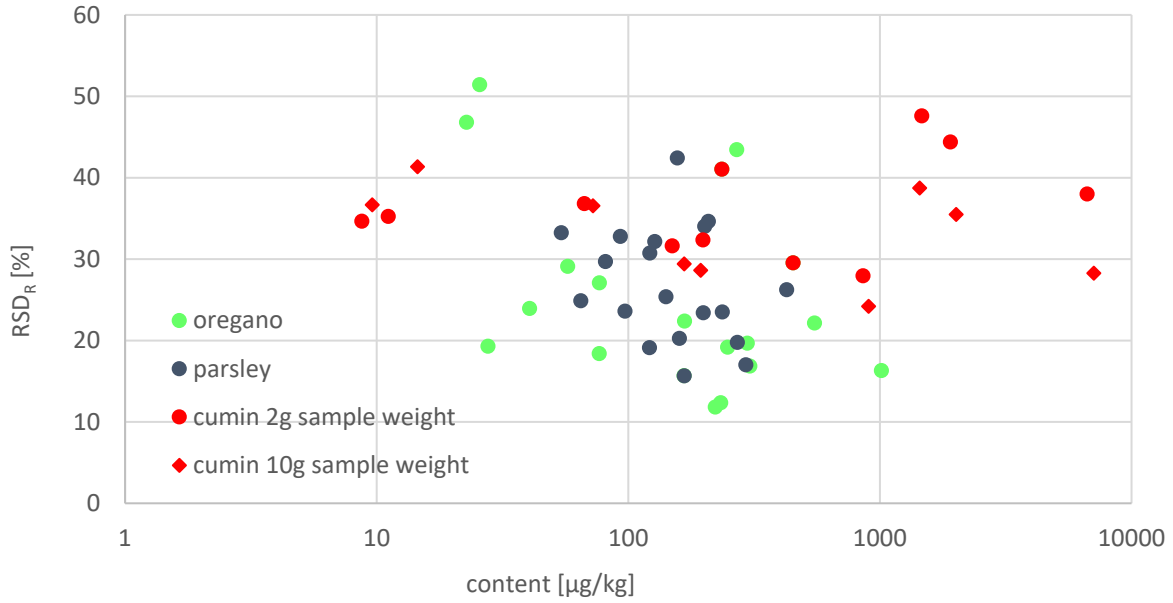


Figure 4: Graphical presentation of the relative reproducibility standard deviation (RSD_R) for all analyte-matrix-combinations as a function of the content (µg/kg). Analyte-matrix-combinations belonging to the same material are presented in the same color.

6.2 Evaluation of methods and performance

To assess the **performance of laboratories and the precision of methods** for the control of plant toxins (draft regulation) a mean z-score was calculated. The intent of the draft regulation, in terms of specifying performance criteria of methods and precision to be achieved, is to ensure monitoring of maximum levels is based on reliable measurement data and that enforcement action is taken only when the exceedance of the maximum level is beyond reasonable doubt. This is the case when the laboratory value minus the measurement uncertainty exceeds the maximum level. The prerequisite is that the (expanded) measurement uncertainty (MU) actually reflects the laboratories range of uncertainty, i.e. whether the laboratories result can be reproduced by another laboratory in the case of a reanalysis. In the field of residue analysis, a default MU of max. 50 % is required of a laboratory, and compliance can be demonstrated by a practical approach using results from proficiency testing. This PT was intended to suggest a procedure that could also be discussed to be applied in the field of mycotoxins and plant toxins.

The z-score represents the participant's deviation from the assigned value, and the target standard deviation was set at 25 % in this PT. The relative standard deviation of reproducibility (RSD_R) should be 25 % or better according to the draft regulation. Thus, a laboratory demonstrates compliance with this criterion if the z-scores for total PA content are $|z| \leq 1$ and if a mean z-score of $|z| \leq 1$is achieved across all analyte-matrix-combinations. A pragmatic approach was taken to calculate the mean z-score by using the mean of 95 % of all PT z-scores. The omission of 5 % of the values is intended to prevent a rare but strongly deviating laboratory result from distorting the overall performance of a laboratory. This omission can also be justified by the fact that individual values with a strong deviation are in most cases not errors inherent to the method (in this case, all or at least several analytes would have a bias). Typical causes of such errors are, for example, the use of an incorrectly prepared or decomposed standard solution. These errors will be corrected by a laboratory as soon as the results are received, and then this source of error no longer contributes to a laboratory's measurement uncertainty. In Figure 5 the mean z-score value of each laboratory is shown that was achieved for the total PA content and in Figure 6 among all analyte-matrix-combinations tested. Almost all laboratories

achieved a mean z-score value of 1 or better. Those laboratories could demonstrate that under reproducibility conditions the sum of their uncertainty components for bias and precision fulfil performance criteria required in the draft regulation. None of the laboratories were found to exceed a mean z-score of 2, indicating that the 50 % value for the expanded measurement uncertainty generally would cover the inter-laboratory variability among participating laboratories.

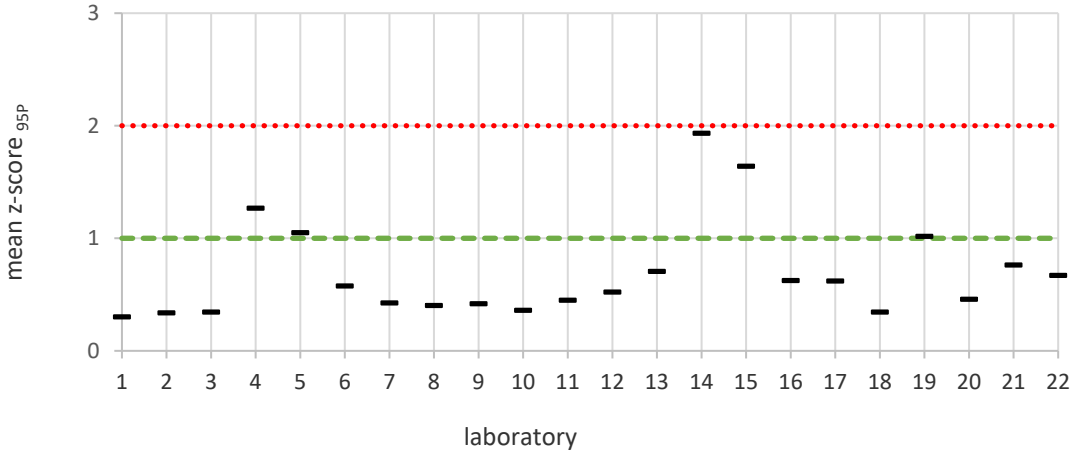


Figure 5: Mean z-score for each laboratory calculated from the 95th percentile of the absolute z-scores a laboratory achieved among all analyte-matrix-combinations tested (only homogeneous analytes; n = 30). Values below 1 demonstrate that the average deviation of the reported results of a laboratory is less than 25 % of the assigned value (green dotted line). The red line indicates 50 % deviation from the assigned value.

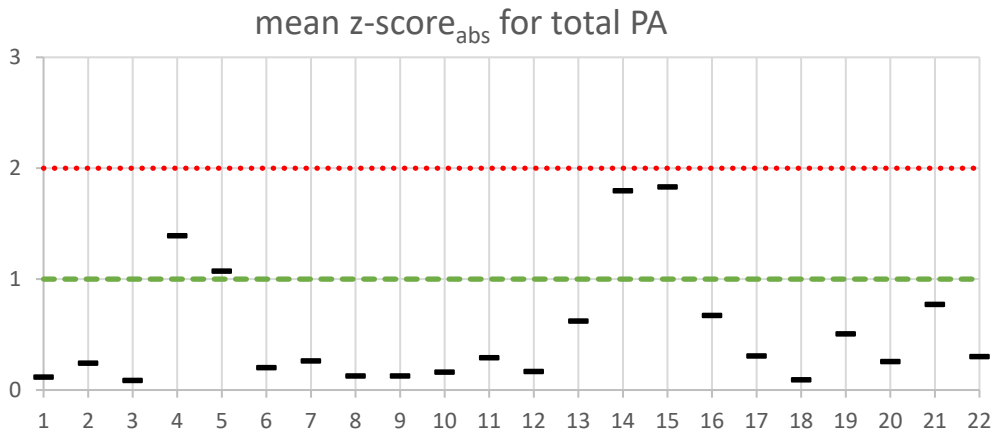


Figure 6: Mean z-score for each laboratory calculated of the absolute z-scores of the total PA-content in tested materials (parsley and oregano). Values below 1 demonstrate that the average deviation of determined total-PA content of a laboratory is less than 25 % of the assigned value (green dotted line). The red line indicates 50 % deviation from the assigned value.

Laboratories were asked to give information on sample preparation and detection with special focus on calibration. All laboratories analyzed the samples using liquid chromatography in combination with mass spectrometry (mostly tandem mass spectrometry in MRM mode; one laboratory applied high resolution mass spectrometry). About 50 % of laboratories performed direct analysis of extracts (“dilute and shoot”). Solid phase extraction (SPE) was applied by 36 % of the participants either in combination with matrix-matched calibration or with subse-

quent dilution of SPE eluates in combination with a calibration with external standards in solvent. A few laboratories applied more specialized procedures, such as liquid-liquid extraction or modified versions of QuEChERS.

To obtain indications of whether specific steps in the sample preparation of the methods used might correlate with any bias, the methods (laboratories) were arranged according to their respective performance in this PT. For this purpose, the mean z-score values obtained by the methods (laboratories) for oregano and parsley were calculated and ordered from the lowest to the highest value (Table 9).

Table 9: Summary of method information ordered by the mean absolute z-score values (95th percentile) obtained from the laboratories/methods for oregano and parsley

Mean z-scores^{+}	Percentage [%] of z-scores below |2|$^{+}$	Type of calibration	Number of cal. level	Accuracy tolerance of Standards [%]	Force through origin	Weighed calibration	Information on extraction	Sample preparation	TA analysis integrated in PAs
0.30	100	st.add.	4	15	no	no	2 x using 0.05 M H_2SO_4	SPE	yes
0.34	100	ext. solvent	5 or more	20	yes	1/x	acidic water-methanol-mixture	dilute and shoot	no
0.35	94	st.add.	1		no	no	2 x using 0.05 M H_2SO_4	dilute and shoot	yes
0.35	97	ext. MMS	6	20	no	1/x	2 x using 0.05 M H_2SO_4	SPE/dilute	no
0.36	100	ext. solvent	10	20	yes	no	2 x using 0.05 M H_2SO_4	SPE/dilute	no
0.40	92	ext. solvent	5	R^2 ≥0.99	no	1/x	1 x using 2 % HCOOH	dilute and shoot	yes
0.42	100	st.add.	6	10	no	1/x	H_2O/ACN /MeOH 1:1:1	dilute and shoot	yes
0.43	97	ext. solvent	7	20	no	1/x	H_2O+AcOH/ MeOH	dilute and shoot	no
0.45	100	ext. solvent	8	20	yes	1/x	2 x using 0.05 M H_2SO_4	dilute and shoot	no
0.46	100	ext. solvent	3	10	yes	no	n.a.	dilute and shoot	yes
0.52	100	ext. MMS	8		no	1/x	2 x using 0.05 H_2SO_4	SPE/dilute	yes
0.58	89	ext. MMS	10	20	yes	no	2 x using 0.05 H_2SO_4	SPE	yes
0.62	94	ext. solvent	8		no	no	1 x using H_2O+ HCOOH	dilute and shoot	yes
0.63	92	ext. MMS	10		yes	no	2 x using 0.05 M H_2SO_4	SPE	yes
0.67	94	ext. solvent	8	20	no	1/x	acidic water	dilute and shoot	yes

Continuation Table 9: Summary of method information ordered by the mean absolute z-score values (95th percentile) obtained from the laboratories/methods for oregano and parsley

Mean z-scores+	Percentage [%] of z-scores below |2|+	Type of calibration	Number of cal. level	Accuracy tolerance of Standards [%]	Force through origin	Weighed calibration	Information on extraction	Sample preparation	TA analysis integrated in PAs
0.71	100	ext. MMS	5	25	no	1/x	2 x using 0.05 M H_2SO_4	SPE	no
0.76	83	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	mod. QuEChERS	yes
1.02	72	st.add.	2	R^2 ≥0.95	yes	1/x	2 x using 0.05 M H_2SO_4	enrichment by evaporation	yes
1.05	89	st.add.	3	20	no	no	1 x using water-alcohol mixture	SPE	yes
1.27	86	st.add.	n.a.	n.a.	n.a.	n.a.	1 x using H2O + AcOH/ MeOH 1/2 (v/v)	dilute and shoot	yes
1.64	69	ext. solvent	5	20	yes	no	1 x using 2 % HCOOH	dilute and shoot	no
1.92	56	st.add.	1		no	no	n.a.	LLE	no

+ only for homogeneous analyte-matrix-combinations from oregano and parsley [n = 30]

A low mean z-score reflects a low deviation from the assigned values over all analytes in both matrices and indicates a good accuracy of the applied methods. As visible from Table 9 there is no correlation with the asked information on sample procedure/method and the respective performance (mean z-score). These results show that satisfactory performance does not depend on a particular procedure and that different procedures lead to satisfactory results. Rather, applying more or less the same procedure can either lead to very low mean z-scores implicating a high precision and accuracy as well to certain degree of questionable results (Table 9).

This result suggests that in LC-MS-based multi-analyte methods, such as the analysis of PAs/TAs, the combination of many factors affects the results. These include, for example experience (new staff or new equipment in the laboratory), the quality of the standard solution used or even errors in calculating the final content (e.g. failure to take dilution steps into account). The challenge in analytics is generally to establish tools to detect these potential errors. Possible options here include the use of some kind of reference material that allows, for example, the detection of erroneous calculations and/or also frequent participation in proficiency tests.

Acknowledging that different methods for the determination of PAs and TAs led to comparable results, it can be concluded that the definition of performance criteria as foreseen in the draft Commission Regulation is a good option to allow laboratories to adapt their methodology to their specific conditions and to take into account new developments such as the implementation of multi-methods. The performance of each lab using its individual methods can be easily evaluated by for instance calculating absolute mean z-scores (see Figure 6).

6.3 Consideration of the influence of homogeneity of samples on RSD_R and estimation of the influence of the sample size used for analysis on RSD_R in inhomogeneous samples

Results of previous inter-laboratory tests have shown that despite thorough homogenization of samples to a particle size of approx. 500 µm, the plant-based test materials are still not sufficiently homogeneous. Using a sample size of 2 g for analysis, which is the common procedure in PA analysis, resulted in higher RSD_R values for (dry) naturally contaminated materials than for spiked or liquid materials [3]. This can be explained by the fact that sufficient precision for the analysis of duplicate samples can only be achieved if the same number of contaminating PA plant particles and the same part of the PA plants are contained in 2 g sample amount, since, e.g. flowers have a different PA content than leaves. Since this can hardly be achieved in practice, the replicate analysis will result in an impaired reproducibility, since the influence of the inhomogeneity adds to the analytical measurement uncertainty.

Therefore, one objective of this study was:

- to determine the extent to which the heterogeneity impairs the reproducibility precision and adds to the analytical measurement uncertainty and
- to what extent the increase of the sample amount used for analysis reduces the spread of analytical results.

The contribution of heterogeneity of samples to the measurement uncertainty was assessed by comparing homogeneous analyte-matrix-combinations with inhomogeneous ones in terms of precision data obtained for analytes and additionally in terms of the deviation of the laboratories from the assigned values (z-score).

As shown in Figure 7, the mean RSD_R values differed between inhomogeneous and homogeneous analyte-matrix-combinations. In this PT, the highest values were determined for europine, lasiocarpine and their respective *N*-oxides in oregano, with a mean RSD_R of 42 % (all naturally contaminated). For the naturally contaminated cumin sample (2 g analysis) a mean RSD_R of 33 % was calculated across all PAs. The mean RSD_R for the spiked PAs in oregano and parsley was only 21 % (Figure 7). When the data of the inhomogeneous analyte-matrix-combinations in oregano and cumin are averaged, a 16 % increase of RSD_R was observed when naturally contaminated PAs were investigated compared to spiked PAs.

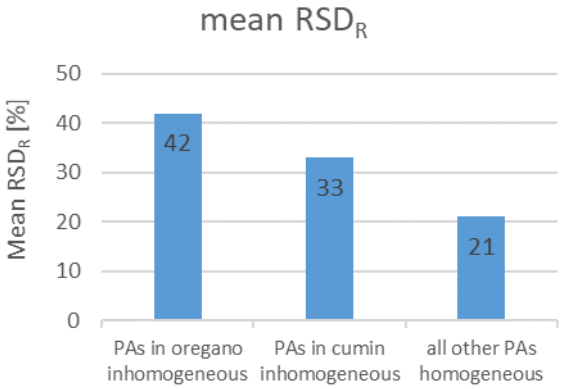


Figure 7: Mean RSD_R values for analytes obtained for naturally contaminated PAs in oregano [n = 4], in cumin [n = 9] and for spiked and therefore homogeneous PAs [n = 30].

This observation was confirmed by comparing the mean z-scores obtained by a laboratory for homogeneous versus inhomogeneous analyte-matrix-combinations. As shown in Figure 8 the laboratories obtained much higher deviations from the assigned value for heterogeneous analyte-matrix-combinations than for homogeneous ones. This spread of results was greater for oregano than for cumin, but on average across all laboratories, a 15 % higher deviation was observed for inhomogeneous compared to homogeneous analyte-matrix-combinations. This observation is consistent with previous results evaluating 484 analyte-matrix-combinations in terms of RSD_R obtained in ring test for either spiked or naturally contaminated (dry) samples [3].

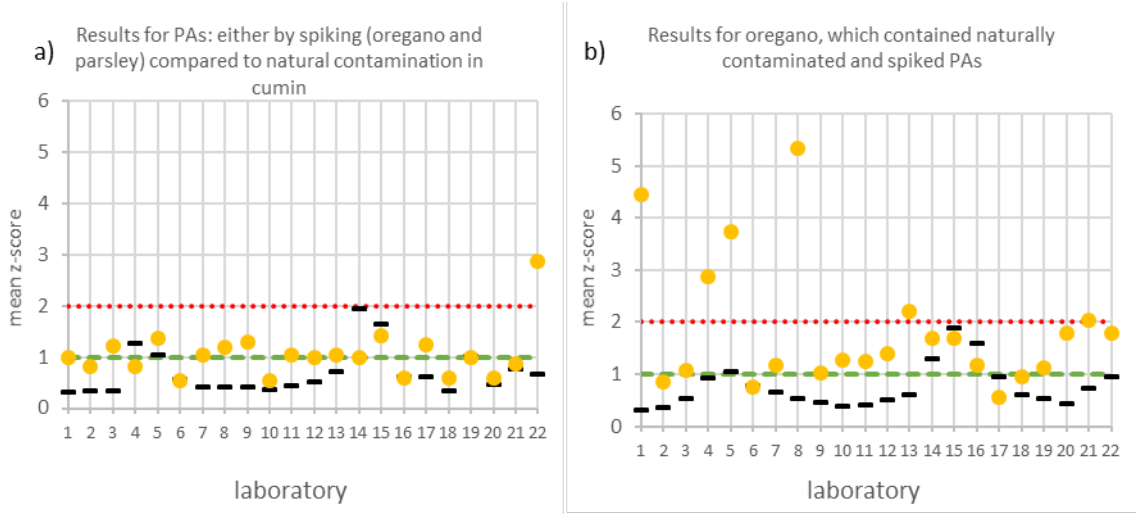


Figure 8: Comparison of mean z-scores achieved by laboratories for analyte-matrix-combinations that were heterogeneous (orange dot) or homogeneous (black dash), with a) oregano with four naturally contained analytes (mean, orange dot) and 13 spiked analytes (mean, black dash) and b) cumin with nine naturally contained PAs (mean, orange dot) compared the mean z-score calculated from the 95th percentile of the absolute z-scores a laboratory achieved among all homogeneous analyte-matrix-combinations tested (black, dash).

To collect inter-laboratory data and assess the extent to which increasing sample size can reduce the spread of analytical results, each participant was asked to analyze the naturally contaminated cumin sample tested as inhomogeneous once at 2 g and once at 10 g sample weight (extraction volume was adjusted accordingly).

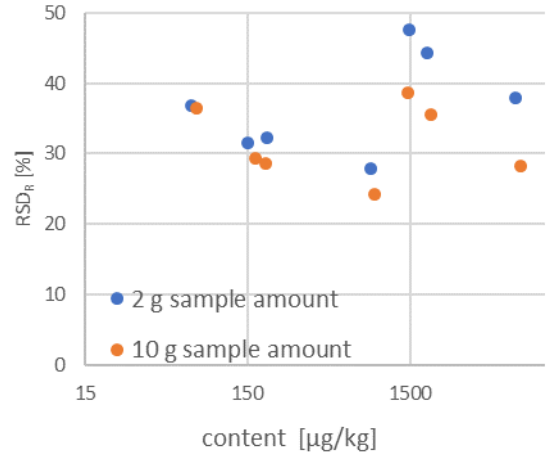


Figure 9: RSD$_R$ values for all analytes in cumin in relation to the concentration determined when either a sample size of 10 g (orange dots) or 2 g (blue dots) is used for analysis.

As expected, and demonstrated in Figure 9, increasing the sample size reduces the spread of results, leading to an improvement in precision. Interestingly, the reduction was higher for analytes with a high content in the cumin material. Since the PAs with high content are the main contributors to the total PA content, the RSD$_R$ reduction in the total PA content was closely related to PAs that had high content/assigned value (refer to Table 7 and Table 8). Therefore, different results may be obtained for other materials with different contamination profiles.

For the cumin material investigated, it was found that a fivefold increase of sample size led to an absolute RSD$_R$ reduction of 8 %.

6.4 Pilot project on standard stability of PA and TA

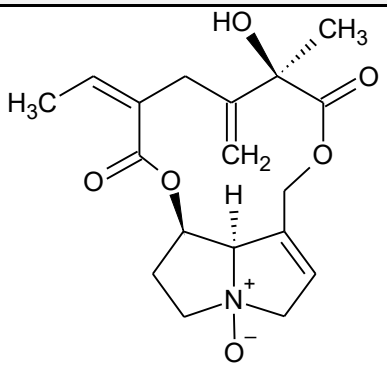
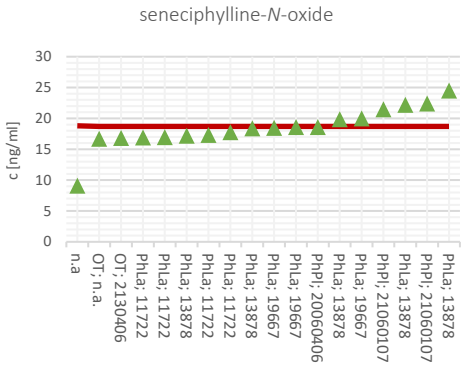
This PT was used to start a pilot study to test the storage stability of PA and TA standard solutions beyond the expiry date specified by the manufacturer. The multi-toxin solution tested in the current PT will be stored in freezer and shipped regularly as part of the NRL's future PT program. Participants will analyze unknown dilutions (prepared by the organizer) of this stored solution using their "freshly prepared" calibration standards. Based on the expected assigned values obtained from the current and subsequent PTs, a trend for storage stability over the next few years can be derived.

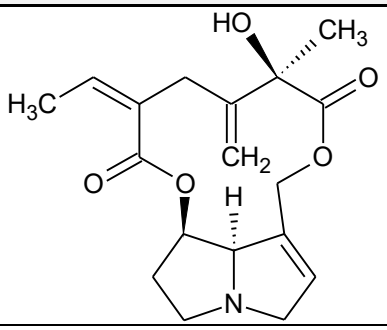
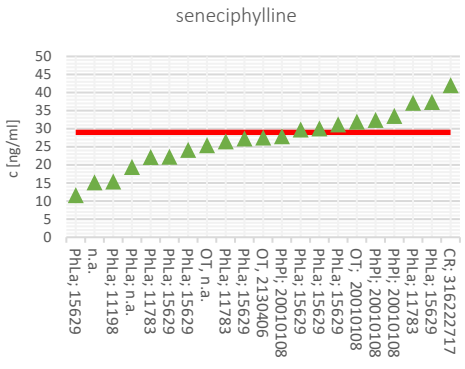
In addition, this standard solution can be used as a reference solution, since the assigned values of the individual PAs and TAs were derived from inter-laboratory tests. In addition, such reference solutions offer each (participating) laboratory the possibility to check the quality of its own standard and thus to perform a quality assurance for the calibration used in its own laboratory. Deviating from the analysis of the samples (cumin, oregano and parsley), the laboratories were asked to analyze the standard solution by a fivefold injection. Further, the respective relative repeatability standard deviation (RSD_R) as well as the information on supplier/batch used for calibration should be reported. The RSD_R values of the five replicate measurements reported by laboratories were on average below 5 % and the RSD_R between laboratories ranged from 7.1 to 28.7 % (Table 4). The assigned values derived from the robust statistic (Hampel or Huber) differed little, indicating that the estimation of the mean as a location parameter was reliable because of the sufficient number of participants (at least 17). The number of individual suppliers and/or unique batches used among laboratories per analyte as calibrant varied between five and eleven. This means that there were participants among the laboratories that used different suppliers and individual batches, but there were also "subunits of laboratories" that used the same supplier and the same batch of the standard. After statistical analysis of the data, no indication was found that a biased result from a laboratory correlated with the use of a standard from a particular supplier or batch. The results for the standard solution are summarized in the "Certificate of Analysis" below.

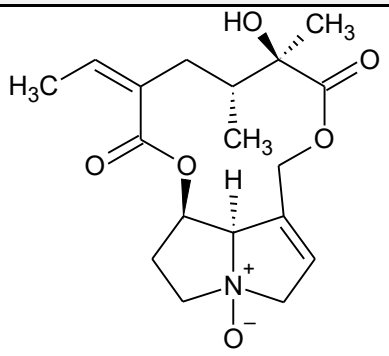
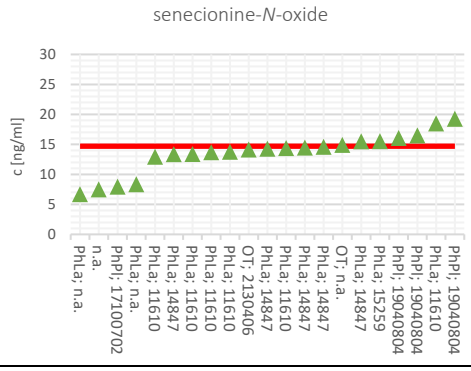
Summary of the results for the analysis of the standard solution

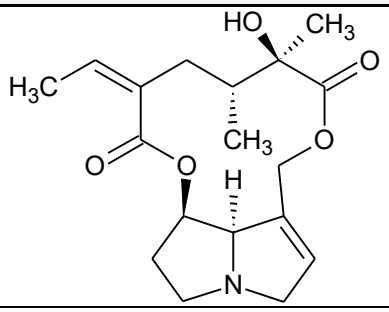
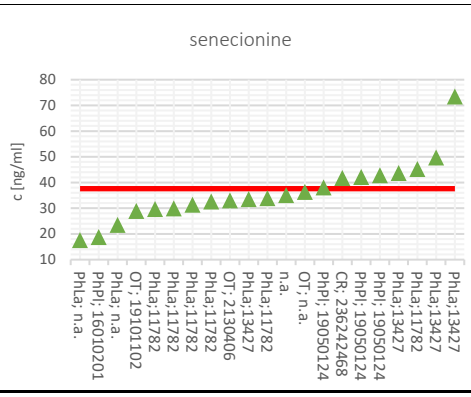
Analyte	Assigned value [ng/ml]	RSD_R [%]	No. laboratory results	No. of individual supplier and/or unique charges
europine	21.2	15.2	18	6
europine-*N*-oxide	37.6	10.4	18	6
heliotrine	31.7	12.4	18	8
heliotrine-*N*-oxide	16.5	7.1	18	8
lasiocarpine	22.3	14.7	18	8
lasiocarpine-*N*-oxide	13.6	7.6	18	9
echimidine	9.2	27.9	18	6
echimidine-*N*-oxide	8.7	9.8	18	8
intermedine	15.0	7.8	18	8
intermedine-*N*-oxide	36.4	14.4	18	10
retrorsine	9.5	17.5	18	7
retrorsine-*N*-oxide	32.3	11.7	17	5
senecionine	19.1	19.7	18	6
senecionine-*N*-oxide	37.6	10.5	18	6
seneciphylline	14.7	22.7	18	5
seneciphylline-*N*-oxide	29.0	13.4	18	6
senkirkine	18.7	8.6	18	9
atropine	17.0	28.7	16	9
scopolamine	32.9	14.3	16	11

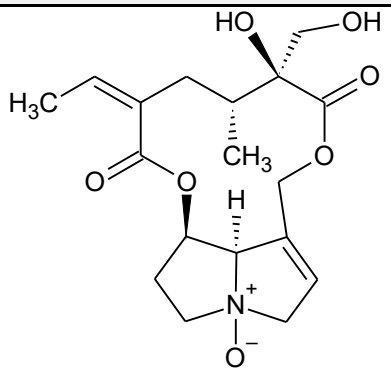
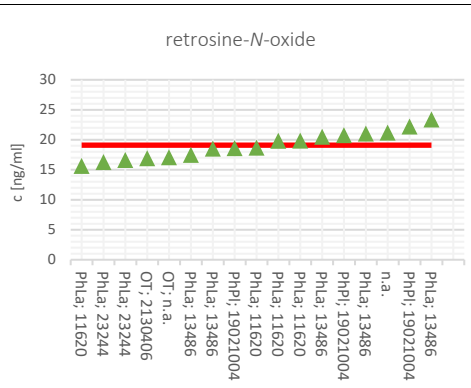
Abbreviations used for standard supplier in alphabetically order	
Cer	Ceriliant
HPC	HPC standards GmbH
LGC	LGC standards
OT	Cfm Oskar Tropitzsch GmbH
PhLa	PhytoLab -phytoproof
PhPl	PHYTOPLAN
S	Sigma-Aldrich
TRC	Toronto Research Chemicals

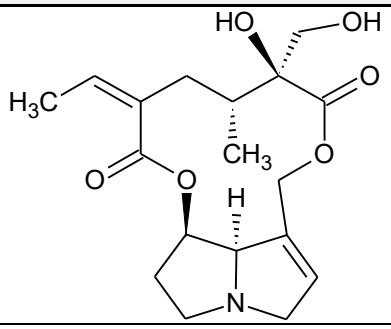
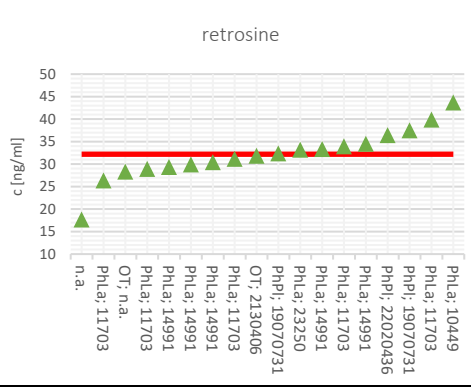
Certificate Of Proficiency Test Analysis		
compound	seneciphylline-*N*-oxide $C_{18}H_{23}NO_6$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	18.7	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	6	
relative standard deviation between laboratory results [%]	13.4	

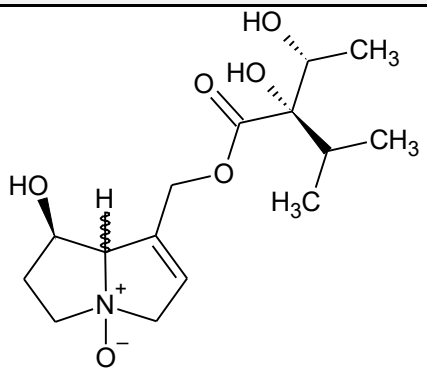
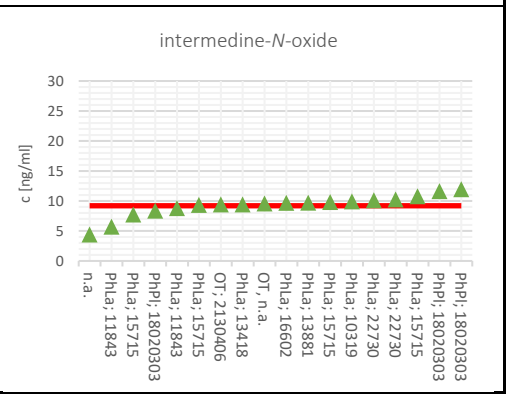
Certificate Of Proficiency Test Analysis		
compound	seneciphylline $C_{18}H_{23}NO_5$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	29.0	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	5	
relative standard deviation between laboratory results [%]	22.7	

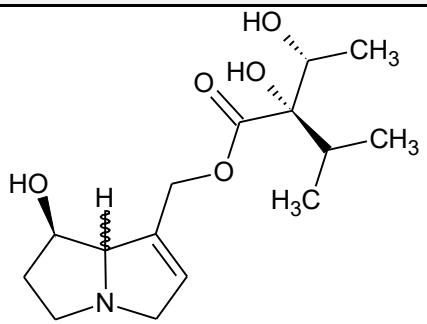
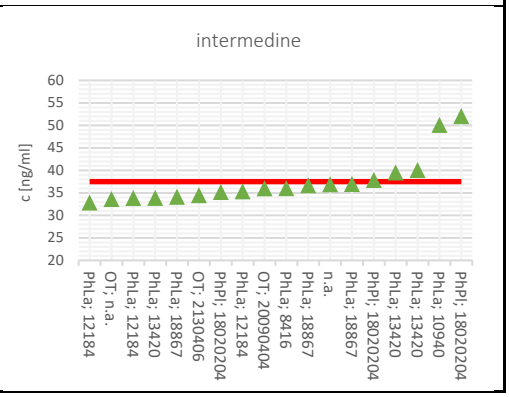
Certificate Of Proficiency Test Analysis		
compound	senecionine-*N*-oxide $C_{18}H_{25}NO_6$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	14.7	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	6	
relative standard deviation between laboratory results [%]	10.5	

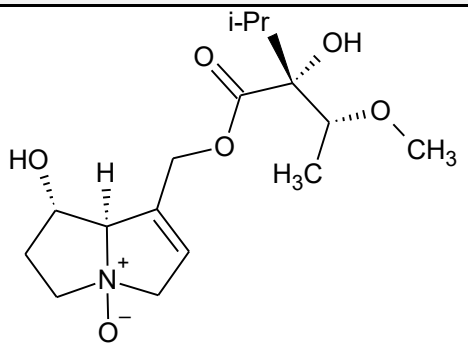
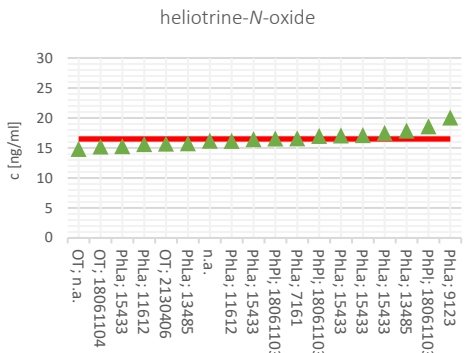
Certificate Of Proficiency Test Analysis		
compound	senecionine $C_{18}H_{25}NO_5$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	37.6	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	6	
relative standard deviation between laboratory results [%]	19.7	

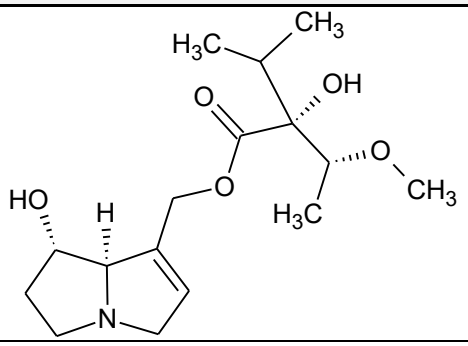
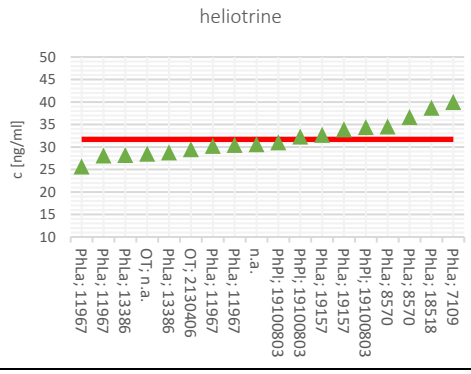
Certificate Of Proficiency Test Analysis		
compound	retrorsine-*N*-oxide $C_{18}H_{25}NO_7$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	19.1	
number of individual laboratory results	17	
number of individual supplier and/or unique charges	5	
relative standard deviation between laboratory results [%]	11.7	

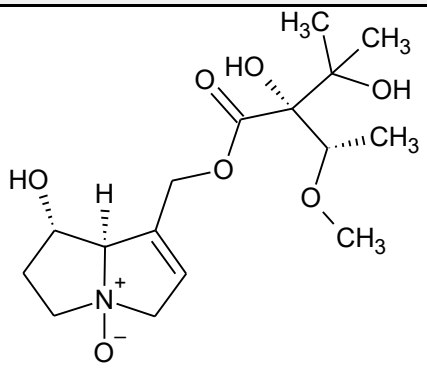
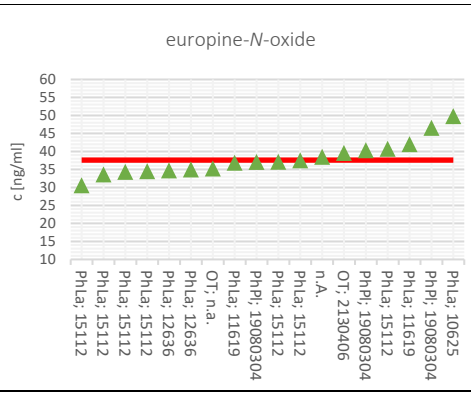
Certificate Of Proficiency Test Analysis		
compound	retrorsine $C_{18}H_{25}NO_6$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	32.2	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	7	
relative standard deviation between laboratory results [%]	17.5	

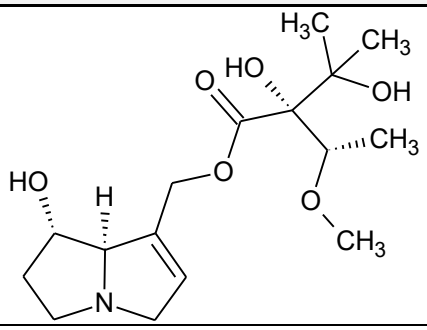
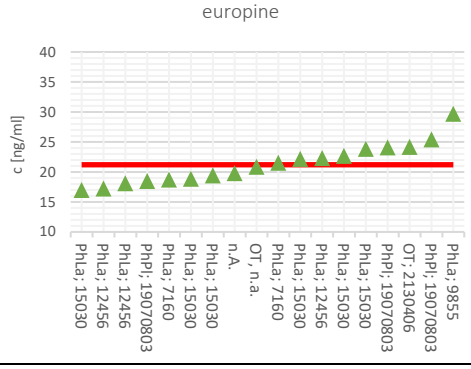
Certificate Of Proficiency Test Analysis		
compound	intermedine-*N*-oxide $C_{15}H_{25}NO_6$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	9.5	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	10	
relative standard deviation between laboratory results [%]	14.4	

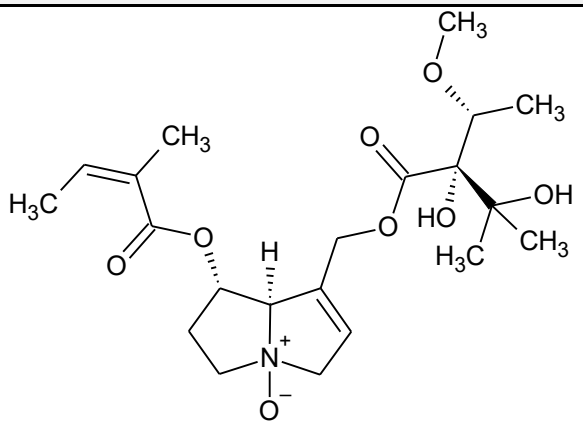
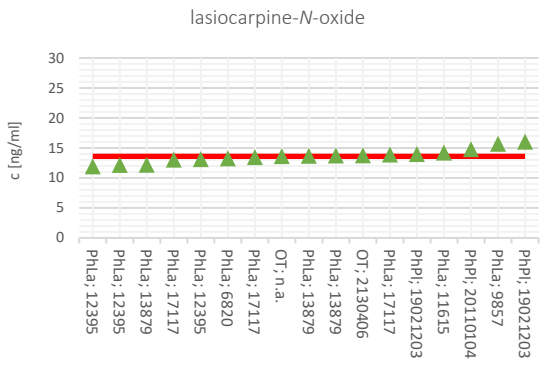
Certificate Of Proficiency Test Analysis		
compound	intermedine $C_{15}H_{25}NO_5$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	36.4	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	8	
relative standard deviation between laboratory results [%]	7.8	

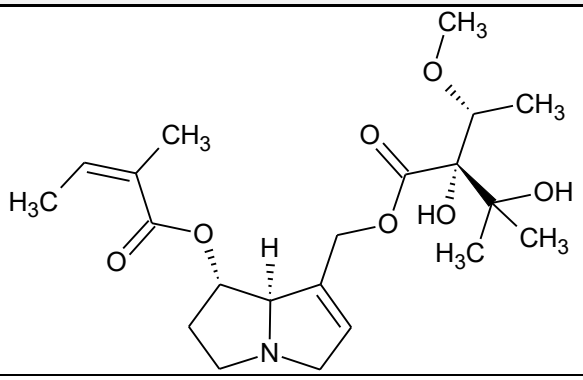
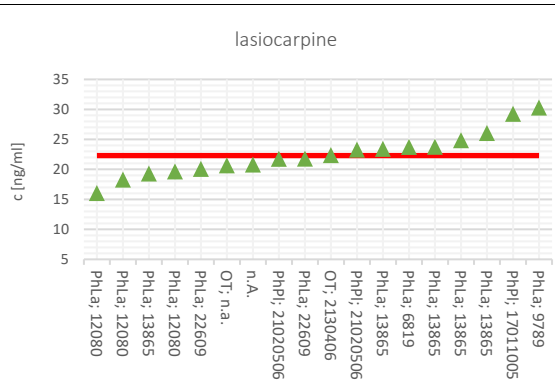
Certificate Of Proficiency Test Analysis		
compound	heliotrine-*N*-oxide $C_{16}H_{27}NO_6$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	16.5	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	8	
relative standard deviation between laboratory results [%]	7.1	

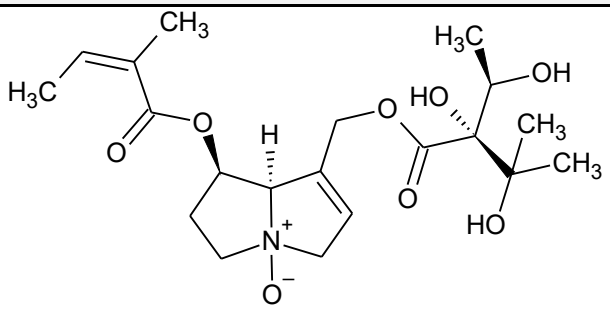
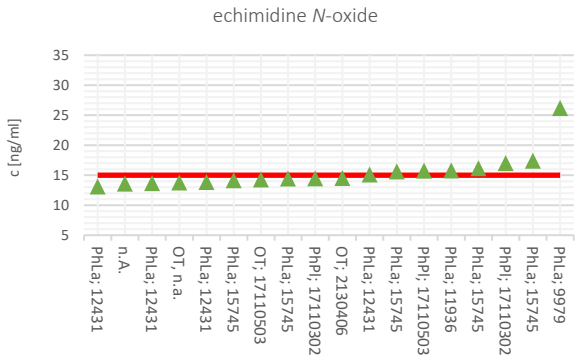
Certificate Of Proficiency Test Analysis		
compound	heliotrine $C_{16}H_{27}NO_5$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	31.7	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	8	
relative standard deviation between laboratory results [%]	12.4	

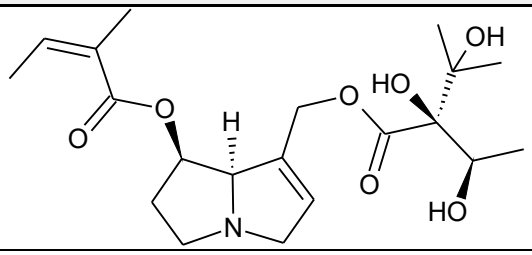
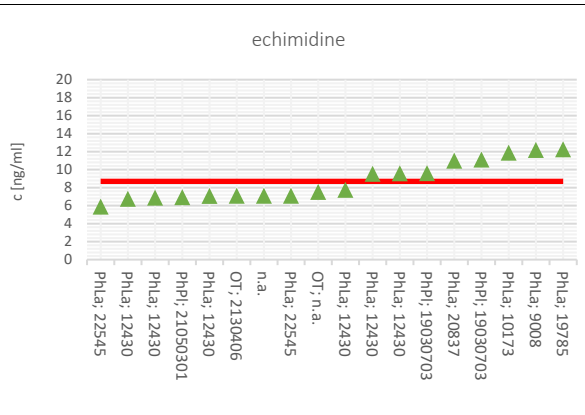
Certificate Of Proficiency Test Analysis		
compound	europine-*N*-oxide $C_{16}H_{27}NO_7$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	37.6	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	6	
relative standard deviation between laboratory results [%]	10.4	

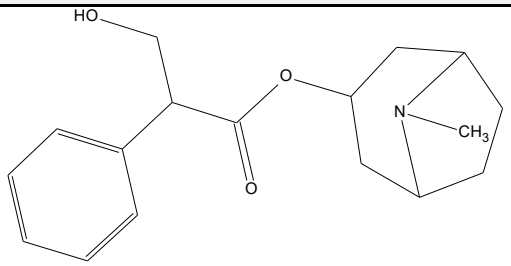
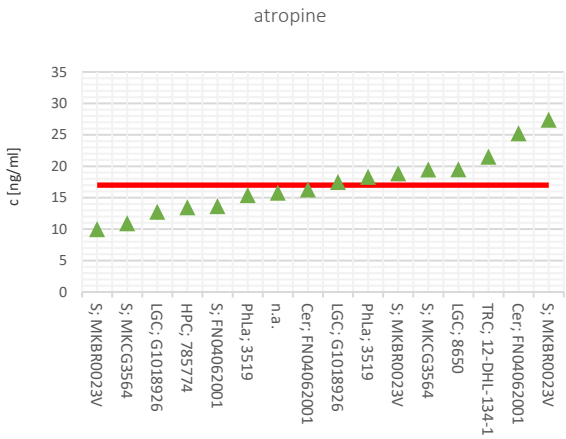
Certificate Of Proficiency Test Analysis		
compound	europine $C_{16}H_{27}NO_8$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	21.2	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	6	
relative standard deviation between laboratory results [%]	15.2	

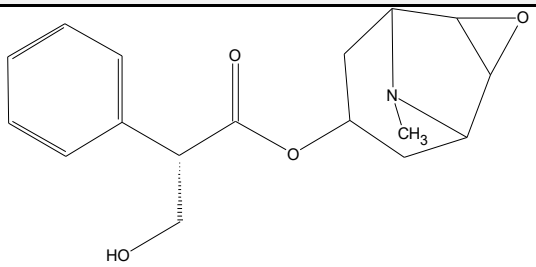
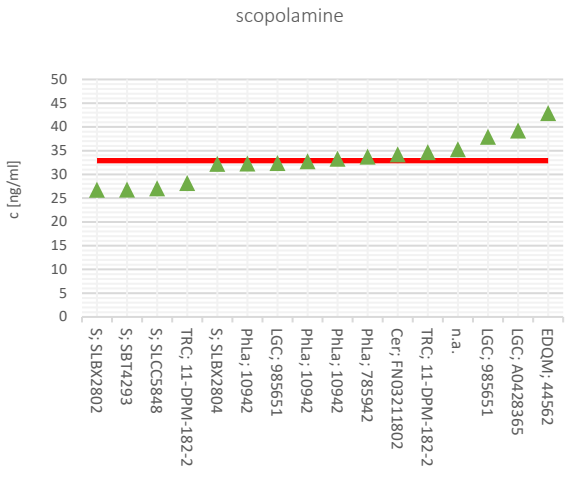
Certificate Of Proficiency Test Analysis		
compound	lasiocarpine-*N*-oxide $C_{21}H_{33}NO_8$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	13.6	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	9	
relative standard deviation between laboratory results [%]	7.6	

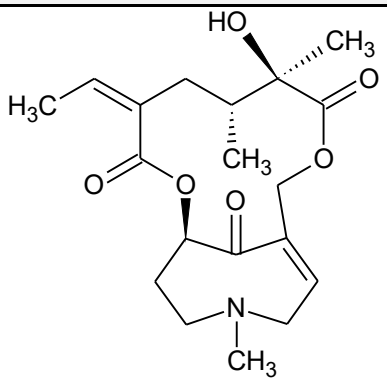
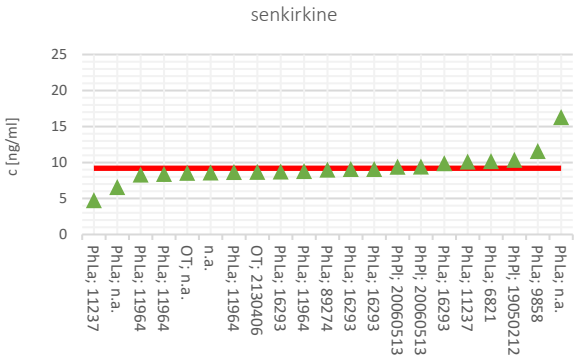
Certificate Of Proficiency Test Analysis		
compound	lasiocarpine $C_{21}H_{33}NO_7$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	22.3	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	8	
relative standard deviation between laboratory results [%]	14.7	

Certificate Of Proficiency Test Analysis		
compound	echimidine-*N*-oxide $C_{20}H_{31}NO_8$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	15.0	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	8	
relative standard deviation between laboratory results [%]	9.8	

Certificate Of Proficiency Test Analysis		
compound	echimidine $C_{20}H_{31}NO_7$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	8.7	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	6	
relative standard deviation between laboratory results [%]	27.9	

Certificate Of Proficiency Test Analysis		
compound	atropine $C_{17}H_{23}NO_3$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	17.0	
number of individual laboratory results	16	
number of individual supplier and/or unique charges	9	
relative standard deviation between laboratory results [%]	28.7	

Certificate Of Proficiency Test Analysis		
compound	scopolamine $C_{17}H_{21}NO_4$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	32.9	
number of individual laboratory results	16	
number of individual supplier and/or unique charges	11	
relative standard deviation between laboratory results [%]	14.3	

Certificate Of Proficiency Test Analysis		
compound	senkirkine $C_{19}H_{27}NO_6$	
start of ring testing	2022/06	
results from previous test rounds	none	
assigned value [ng/ml]	9.2	
number of individual laboratory results	18	
number of individual supplier and/or unique charges	9	
relative standard deviation between laboratory results [%]	8.6	

7 Conclusions

The NRL for Mycotoxins and Plant Toxins organized a PT aiming at assessing the measurement capability of laboratories regarding the determination of pyrrolizidine and tropane alkaloids in herbs and spices as well as in standard solutions. 22 laboratories from Germany participated. These are active in food control, either as official laboratories of the federal states or as contract laboratories and had a high level of expertise in PA analysis.

Three test materials (cumin, oregano and parsley) and one standard solution were sent in the concentration range from 7 to 7065 µg/kg and 8.5 to 37.6 ng/ml, respectively. The materials were prepared to cover the entire analytical scope proposed to control maximum levels. This PT did not focus on the analysis of co-occurring isomers and materials were only spiked with one isomer per isomer group (Table 1, Table 3).

For each material, the assigned values were calculated as robust mean values from the laboratory results (consensus values) and the performance of the laboratories was evaluated using z-scores. As stipulated in the draft Commission Regulation and as already applied in previous PTs, a target standard deviation of 25 % was set. That means that a 25 % deviation of a laboratory result from the consensus value results in a z-score of 1.

To evaluate laboratory performance and method precision, the mean z-score was calculated from the 95th percentile of absolute z-scores obtained by the laboratory among all (homogeneous) analyte-matrix-combinations tested [n = 30]. The mean z-score value describes the deviation of the laboratory over all tested assigned values, which in turn is a meaningful parameter for evaluating the measurement uncertainty of the laboratory. Since the target standard deviation of 25 % reflects the requirement for the precision under reproducibility conditions specified in the draft regulation, a mean z-score of $|z| \leq 1$ indicates that the required criterion is met. Almost all laboratories achieved a mean z-score of one or better and none of the laboratories exceeded a mean z-score of two (Figure 5 and Figure 6).

A total of 1082 values for analyte-matrix-combinations were evaluated for PAs and 130 for TAs, of which 92 % of PA and 95 % for TAs z-scores were satisfactory ($|z| \leq 2$). For homogeneous materials, the relative reproducibility standard deviation (RSD_R) for individual PAs ranged from 7.1 to 34.2 % and 9 out of 47 individual PA-matrix-combinations exceeded the RSD_R value of 25 %. The RSD_R for total PA-contents ranged from 9.3 to 13.0 % and fully met the RSD_R requirements (Tables 4–6). In summary, these values indicate that the proficiency of the laboratories is satisfactory and the methods in use are fit for purpose to monitor maximum levels of PAs and TAs in herbs and spices.

Results of previous inter-laboratory tests have shown that despite thorough homogenization of samples to a particle size of approx. 500 µm, the (dry) plant test materials are still not sufficiently homogeneous. The use of a sample size of 2 g for analysis resulted in about 15 % higher RSD_R values for (dry) naturally contaminated materials than for spiked or liquid materials [3]. In the present PT, one material (oregano) contained four analytes that were present due to natural contamination (heterogeneous) and 13 that were spiked (homogeneous). Comparing the performance of the laboratories achieved in the same material for heterogeneous compared to homogeneous PAs it becomes clear that the dispersion of the particles of the contaminating PA plants is very large and even laboratories with low mean z-score values can have very high deviations from the assigned value (Figure 8). Different calculations were done to give a rough estimation on to what extent the influence of the inhomogeneity adds to the analytical measurement uncertainty. On average across all laboratories a 15 % higher deviation was observed for inhomogeneous compared to homogeneous analyte-matrix-combinations what is consistent with previous results (Figure 9).

In addition, each participant was asked to analyze the naturally contaminated cumin sample tested as inhomogeneous once at 2 g and once at 10 g sample weight (extraction volume was adjusted accordingly). The objective was to estimate, on the basis of data obtained in an inter-laboratory test, the extent to which an increasing sample size can reduce the spread of the analytical results. Based on the one material investigated (cumin), it was estimated that a five-fold increase of sample size (here 2 g and 10 g) led to an 8 % reduction of RSD_R.

8 References

1 *Commission Regulation (EU) 2040/2020 of 11 December 2020 amending regulation (EC) 1881/2006 as regards maximum levels of pyrrolizidine alkaloids in certain foodstuffs.* 2020. Official Journal of the European Union, L 420/1

2 European Commission, Draft Commission implementing regulation *laying down the methods of sampling and analysis for the control of the levels of plant toxins in food and repealing Regulation (EU) 2015/705.* 2021. SANTE/11494R2/2021

3 These, A., Pydde, E., *Report on the 2020 Proficiency Test of the German Reference Laboratory for Mycotoxins and Plant Toxins. Determination of pyrrolizidine and tropane alkaloids in herbal tea and solutions. 2020. BfR Wissenschaft 03/2020*. Available online:https://www.bfr.bund.de/cm/350/report-on-the-2020-proficiency-test-of-the-german-reference-laboratory-for-mycotoxins-and-plant-toxins.pdf

4 *Commission Regulation (EU) 2021/1408 of 27 August 2021 amending Regulation (EC) No 1881/2006 as regards maximum levels of tropane alkaloids in certain foodstuffs*. 2021. Official Journal of the European Union, L 304/1

5 European Commission, Joint Research Centre, Cubero-Leon, E., Oliveira Gonçalves, C., Tamosiunas, V., et al., *Report on the 2016 proficiency test of the European Union Reference Laboratory for Mycotoxins, for the Network of National Reference Laboratories : determination of tropane alkaloids in tea and herbal infusions*, Publications Office, 2017, https://data.europa.eu/doi/10.2760/90137

6 ISO 13528:2015: *Statistical methods for use in proficiency testing by interlaboratory comparison.* 2015. International Organization for Standardization, Geneva, Switzerland.

7 Huber, P.J., *Robust statistics*. 1981: Wiley.

8 Venables, W.N.a.R., B. D., *Modern Applied Statistics with S*. Fourth edition ed. 2002: Springer, New York.

9 Team, R.C., *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. 2022, Vienna, Austria.

9 Acknowledgement

We want to thank the participating laboratories (see Table 2) for their efforts in reporting the results and filling in the questionnaire on the methods used.

10 Annex

10.1 Results of homogeneity testing

Table 10: Results of homogeneity testing for oregano

	Em_G	EmN_G	Eu	EuN	Ht	HtN	Im_G	ImN_G	Lc	LcN	Re	ReN	Sc_G	ScN_G	Sk	Sp	SpN	At	So
F-Test; homogenous, if $F < F_{krit}$	yes	no	no	no	yes	yes	yes	yes	no	no	yes	yes	yes	no	yes	yes	yes	yes	yes
(DIN 13582 B2.2); homogenous, if $s_s \leq 0{,}3*s_{Soll}$	no	yes	no	no	yes	yes	yes	yes	no	no	yes	yes	yes	no	yes	yes	yes	yes	yes
(DIN 13582 B2.3);homogenous, if $s_s < \sqrt{c}$	yes	yes	no	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes
used for laboratories evaluation	yes	yes	no	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes

Table 11: Results of homogeneity testing for parsley

	Em_G	EmN_G	Eu	EuN	Ht	HtN	Im_G	ImN_G	Lc	LcN	Re	ReN	Sc_G	ScN_G	Sk	Sp	SpN	At	So
F-Test; homogenous, if $F < F_{krit}$	yes	yes	no	yes	no	yes	yes	no	yes	yes	yes	yes	yes	yes	no	yes	yes	no	no
(DIN 13582 B2.2); homogenous, if $s_s \leq 0{,}3*s_{Soll}$	yes	yes	yes	yes	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
(DIN 13582 B2.3);homogenous, if $s_s < \sqrt{c}$	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
used for laboratories evaluation	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

10.2 Results of statistical evaluation

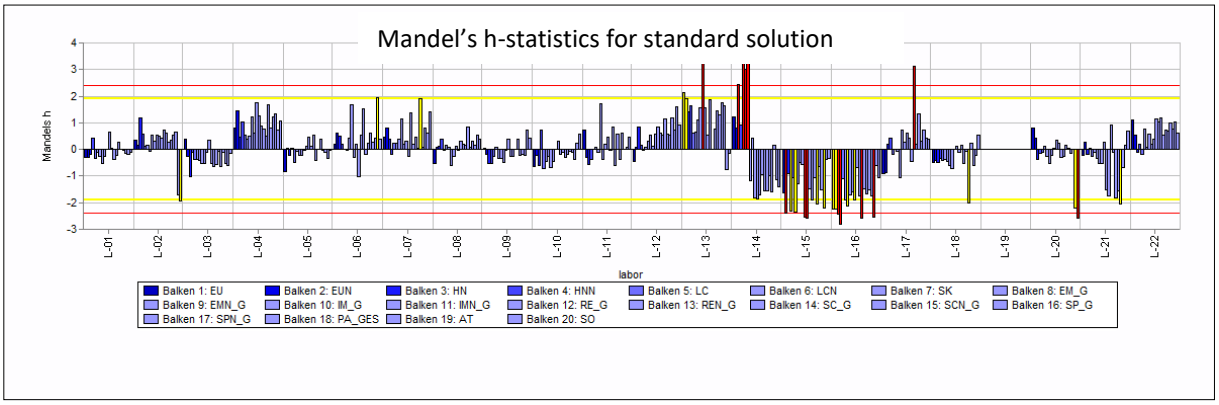


Figure 10: Mandel's h-statistics for standard solution (critical values for the 5 % and 1 % significance level are shown as yellow and red lines respectively. Laboratories deviations are marked with the same colour)

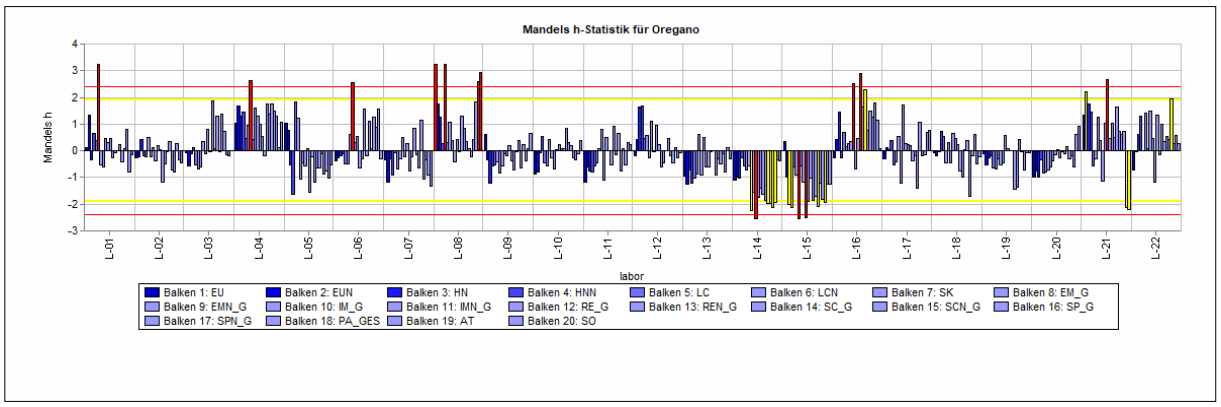


Figure 11: Mandel's h-statistics for oregano (critical values for the 5 % and 1 % significance level are shown as yellow and red lines respectively. Laboratories deviations are marked with the same colour)

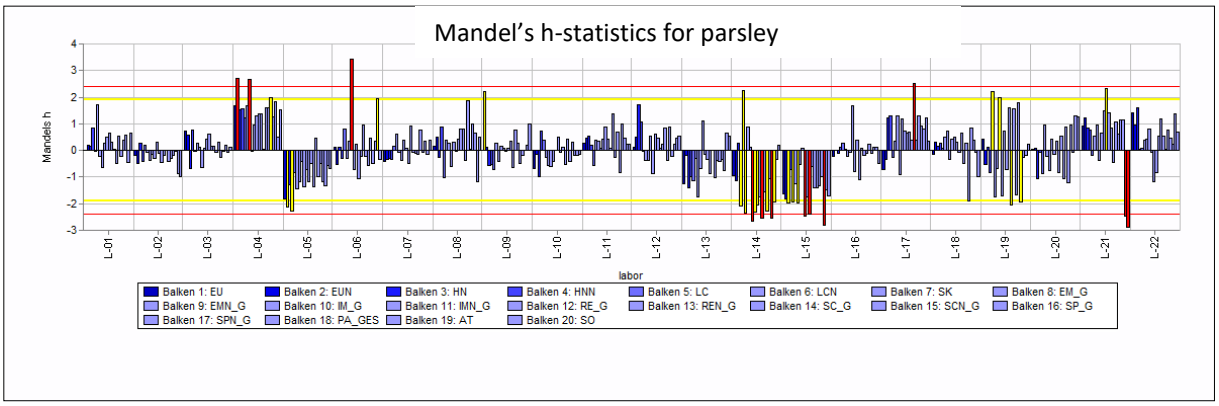


Figure 12: Mandel's h-statistics for parsley (critical values for the 5 % and 1 % significance level are shown as yellow and red lines respectively. Laboratories deviations are marked with the same colour)

10.3 Reported limits of quantification

Table 12: Information from the laboratories on LOQ in the standard solution [ng/ml] or in the test materials [µg/kg]

		Eu	EuN	Hn	HnN	Lc	LcN	Em-G	EmN-G	Im-G	ImN-G	Re-G	ReN-G	Sc-G	ScN-G	Sp-G	SpN-G	Sk	At	Sco
L-01	ng/ml	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025
	µg/kg	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
L-02	ng/ml	0.016	0.016	0.016	0.016	0.016	0.04	0.016	0.04	0.016	0.04	0.04	0.08	0.04	0.04	0.08	0.08	0.04	0.0025	0.0025
	µg/kg	2	2	2	2	2	5	2	5	2	5	5	10	5	5	10	10	5	5	5
L-03	ng/ml	0.02	0.02	0.02	0.02	0.02	0.05	0.02	0.02	0.05	0.02	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.02	0.02
	µg/kg	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
L-04	ng/ml	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.
	µg/kg	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
L-05	ng/ml	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	µg/kg	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
L-06	ng/ml	0.1	0.07	0.09	0.08	0.06	0.08	0.08	0.07	0.1	0.07	0.1	0.08	0.07	0.06	0.06	0.1	0.07	0.07	0.07
	µg/kg	10	7	9	8	6	8	8	7	10	7	10	8	7	6	6	10	7	7	7
L-07	ng/ml	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.1	0.1
	µg/kg	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	1	1
L-08	ng/ml	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025
	µg/kg	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
L-09	ng/ml	0.125	0.25	0.125	0.125	0.125	0.125	0.05	0.125	0.05	0.05	0.125	0.25	0.25	0.125	0.125	0.25	0.125	0.05	0.05
	µg/kg	2.5	5	2.5	2.5	2.5	2.5	1	2.5	1	1	2.5	5	5	2.5	2.5	5	2.5	1	1
L-10	ng/ml	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
	µg/kg	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	2	2
L-11	ng/ml	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	µg/kg	5	5	5	5	5	5	5	5	5	5	5	10	10	5	5	5	5	5	5
L-12	ng/ml	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25
	µg/kg	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
L-13	ng/ml	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.01	0.01
	µg/kg	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	1	1
L-14	ng/ml	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.
	µg/kg	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.
L-15	ng/ml	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.01	0.01
	µg/kg	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	1
L-16	ng/ml	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	µg/kg	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
L-17	ng/ml	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.05	0.05	0.02	0.05	0.05	0.02	0.02	0.02	0.02
	µg/kg	10	10	10	10	10	10	10	10	10	10	20	20	10	20	20	10	10	10	10
L-18	ng/ml	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.2	0.2
	µg/kg	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	1	1
L-19	ng/ml	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.
	µg/kg	10	10	10	10	10	10	10	10	10	10	20	20	20	10	20	10	10	10	10
L-20	ng/ml	0.1	0.1	0.05	0.05	0.05	0.1	0.4	0.5	0.05	0.05	0.3	0.5	0.3	0.3	0.1	0.1	0.05	0.025	0.025
	µg/kg	1	3	1	1	1	2	6	7	1	1	3	5	4	4	4	4	1	0.5	0.5
L-21	ng/ml	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.
	µg/kg	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.	n. a.
L-22	ng/ml	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	µg/kg	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30

11 List of Figures

Figure 1: z-score results for oregano, rel. target standard deviation is 25 % and corresponds to |z| score = 1 (blue triangle: |z| score ≤ 2, yellow triangle: 2 < |z| score < |3|, red triangle: |z| score ≥ 3) 18

Figure 2: z-score results for parsley, rel. target standard deviation is 25 % and corresponds to |z| score = 1 (blue triangle: |z| score ≤ 2, yellow triangle: 2 < |z| score < |3|, red triangle: |z| score ≥ 3) 19

Figure 3: Mean RSD_R [%] for each analyte and the total PA content in all materials (only homogeneous analyte-matrix-combinations) 19

Figure 4: Graphical presentation of the relative reproducibility standard deviation (RSD_R) for all analyte-matrix-combinations as a function of the content (µg/kg). Analyte-matrix-combinations belonging to the same material are presented in the same color. 20

Figure 5: Mean z-score for each laboratory calculated from the 95th percentile of the absolute z-scores a laboratory achieved among all analyte-matrix-combinations tested (only homogeneous analytes; n = 30). Values below 1 demonstrate that the average deviation of the reported results of a laboratory is less than 25 % of the assigned value (green dotted line). The red line indicates 50 % deviation from the assigned value. 21

Figure 6: Mean z-score for each laboratory calculated of the absolute z-scores of the total PA-content in tested materials (parsley and oregano). Values below 1 demonstrate that the average deviation of determined total-PA content of a laboratory is less than 25 % of the assigned value (green dotted line). The red line indicates 50 % deviation from the assigned value. 21

Figure 7: Mean RSD_R values for analytes obtained for naturally contaminated PAs in oregano [n = 4], in cumin [n = 9] and for spiked and therefore homogeneous PAs [n = 30]. 24

Figure 8: Comparison of mean z-scores achieved by laboratories for analyte-matrix-combinations that were heterogeneous (orange dot) or homogeneous (black dash), with a) oregano with four naturally contained analytes (mean, orange dot) and 13 spiked analytes (mean, black dash) and b) cumin with nine naturally contained PAs (mean, orange dot) compared the mean z-score calculated from the 95th percentile of the absolute z-scores a laboratory achieved among all homogeneous analyte-matrix-combinations tested (black, dash). 25

Figure 9: RSD_R values for all analytes in cumin in relation to the concentration determined when either a sample size of 10 g (orange dots) or 2 g (blue dots) is used for analysis. 25

Figure 10: Mandel's h-statistics for standard solution (critical values for the 5 % and 1 % significance level are shown as yellow and red lines respectively. Laboratories deviations are marked with the same colour) 41

Figure 11: Mandel's h-statistics for oregano (critical values for the 5 % and 1 % significance level are shown as yellow and red lines respectively. Laboratories deviations are marked with the same colour) 42

Figure 12: Mandel's h-statistics for parsley (critical values for the 5 % and 1 % significance level are shown as yellow and red lines respectively. Laboratories deviations are marked with the same colour) 42

12 List of Tables

Table 1: Analytical scope for monitoring the PA maximum levels in food. Natural occurring isomers can be summarized as group. The maximum level refers to the sum of the given PA and/or PA groups [1]. 7

Table 2: Participating laboratories 10

Table 3: Contamination profile of proficiency test samples. Analytes were either spiked (+) or naturally contaminated (x) 12

Table 4: Analytical results and statistical characteristics for the standard solution 16

Table 5: Analytical results and statistical characteristics for oregano 16

Table 6: Analytical results and statistical characteristics for parsley 17

Table 7: Analytical results and statistical characteristics for cumin applying a sample amount of 2 g for analysis 17

Table 8: Analytical results and statistical characteristics for cumin applying a sample amount of 10 g for analysis 18

Table 9: Summary of method information ordered by the mean absolute z-score values (95^{th} percentile) obtained from the laboratories/methods for oregano and parsley 22

Table 10: Results of homogeneity testing for oregano 41

Table 11: Results of homogeneity testing for parsley 41

Table 12: Information from the laboratories on LOQ in the standard solution [ng/ml] or in the test materials [µg/kg] 43