SAFIA

Validation report of the mycoProfile 6 Kit

Your contact person

Dr. rer. nat. Peter Carl
Research & Development

T: +49 (0) 172 4260 922

M: carl@safia.tech

www.SAFIA.tech

Table of Contents

Executive summary: overview of main results 4
Introduction 6
- Mycotoxins 6
- Principle of the SAFIA Assay 6

Scope of this validation 7
Materials and Methods 7
- Materials and devices for Analysis 7
- Reference Materials and Blank Materials 7
- SAFIA Kits and Assay Procedure 9
  - Sample Preparation 9
  - Replicates and Plate Layouts 10
  - Assay Execution 10
- CyFlow Cube 6 V2m Measurement and Data Evaluation 10
  - Correction Factors Used in SAFIA 11

Results and Discussion 12
- Calibration 12
  - Linearity 13
- Selectivity (Cross reactivity) 14
- Limit of Detection 15
  - Estimation of LODs via repetitive measurement of sample blanks 15
  - Estimation of LODs via precision profile of the calibration curves 16
  - A consensus value for LODs from both approaches 19
- Limits of Quantification and working range 20
- Measurement of QC (Reference) Material – Estimation of Accuracy 25
  - Sum of Aflatoxins B1, B2 and G1, G2 (AFL) 26
  - Ochratoxin A (OTA) 27
  - Fumonisins (FUM) 27
  - Deoxynivalenol (DON) 28
  - Zearalenone (ZEN) 28
  - Sum of T-2/HT-2 Toxins (T-2/HT-2) 28
- Trueness 29
- Measurement of QC (Reference) Material – Estimation of Precision 29
  - Intra-assay precision (repeatability) 29
  - Intermediate precision (between-day) 31
  - Within-laboratory reproducibility (inter-analysis precision) 32
  - Precision limits (repeatability limit and within-laboratory limit) 33
- Measurement uncertainty 33

Trueness and method-wide bias contribution 33
Contributing of method-wide precision 34
Ruggedness 36
Observed effects and critical control parameters (CCP) 37
Annex 39
Plate Layouts 39
Results of individual samples and One-Way ANOVA 44
Results for Sum of Aflatoxins (AFL) 44
AFL Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95) 45
Results for Deoxynivalenol (DON) 51
DON Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95) 52
Results for Sum of Fumonisins (FUM) 57
FUM Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95) 58
Results for Ochratoxin A (OTA) 63
OTA Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95) 64
Results for sum of T-2 and HT-2 toxins (T-2/HT-2) 69
T-2/HT-2 Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95) 70
Results for sum of zearalenone (ZEN) 75
ZEN Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95) 76
Experimental Design plan for Ruggedness testing 81
ANOVA of DoE: Parameter AFL 82
ANOVA of DoE: Parameter Control 84
ANOVA for Doe: Parameter DON 86
ANOVA of DoE: Parameter FUM 88
ANOVA of DoE: Parameter OTA 90
ANOVA of DoE: Parameter T-2/HT-2 92
References 96
Contact 97

Executive summary: overview of main results

Table 1. Overview of main method performance data.

Validation Parameter	Subparameter	OTA	DON	FUM	ZEN	AFL	T-2/HT-2
Sensitivity	Limit of Detection [µg/kg]	1.16	47.0	39.4	3.49	0.883	11.3
Working range	Lower Limit of Quantification LLOQ [µg/kg]	1.20	50.0	50.0	3.50	0.90	11.5
	Upper Limit of Quantification ULOQ [µg/kg]	50	1500	700	300	15	300
Trueness	Trueness (Recovery/Bias) Mean (Min-Max) Target for Mean: 70 to 130 %	98% (61% - 158%)	90% (74% - 119%)	94% (77% - 113%)	100% (82% - 125%)	95% (63% - 124%)	97% (65% - 154%)
Precision	Intra-Assay Precision Mean (Min-Max) [%] Target for Mean: < 10%	2.6 (0.0 - 8.1)	5.3 (1.8 - 13.3)	6.1 (1.6 - 15.4)	3.8 (0.5 - 9.7)	6.2 (1.4 - 18.6)	7.5 (1.5 - 18.7)
	Inter-Assay Precision Mean (Min-Max) [%] Target for Mean: < 20 %	4.3 (1.7-10.3)	8.4 (3.3-13.4)	10.6 (6.9-18.8)	7.4 (2.8-16.4)	8.4 (3.1-19.3)	10.4 (2.8-19.5)
	Repeatability (Inter-Analysis Precision) Mean (Min-Max) [%] Target for Mean: < 20 %	8.0 (3.8 - 11.0)	10.6 (7.0 - 13.9)	13.8 (12.6 - 15.0)	12.0 (5.7 - 17.3)	9.1 (4.8 - 13.7)	13.1 (6.8 - 17.1)
Measurement Uncertainty MU95	MU95 (typical, RMS based)	49%	46%	39%	35%	48%	51%
	MU95 (conservative, 90 % percentile)	76%	61%	60%	54%	73%	77%
Ruggedness	Critical Control Parameters	Ethanol concentration in extraction, incubation time relative to calibration					

Table 2. Study design and acceptance criteria. CV denotes coefficient of variation; MU95 denotes expanded uncertainty at approximately 95% coverage; LLOQ and ULOQ denote the confirmed quantitative range.

Item	Value to report
Assay name and version	SAFIA mycoProfile 6
Analytes	Sum of Aflatoxins (AFB1, B1; G1; G2; AFL) Ochratoxin A (OTA) Deoxynivalenol (DON) Sum of Fumonisins (FB1, FB2, FB3; FUM) Zearalenone (ZEN) Sum of T-2/HT-2 Toxins (T-2/HT-2)
Measurement principle and instrument	Indirect competitive suspension array fluorescence immunoassay (SAFIA) with flow cytometric readout. Sysmex CyFlow Cube 6 V2m flow cytometer with CyFlow Robby 6 autoloader.
Matrices included in validation	High protein content (Grain and baking goods) High water content (Wine) High Fat content (Nuts) High sugar content (Dried fruits) Processed Food
Calibration model and number of standards	Sigmoidal 4 Parameters 8 standards per curve
Plate layout strategy	Block randomized, plate maps in Annex
Blanks included per plate	Reagent blank (70 % Ethanol, diluted 1:4 in sample buffer)
QC reference materials used	FAPAS QC Materials for covering all types of matrices
LOQ confirmation approach	Spiking at LLOQ and ULOQ levels under intermediate precision conditions.
Acceptance criteria, recovery	70 to 130% (mean)
Acceptance criteria, precision	Repeatability CV 10 %, intermediate precision RSD 20%
Measurement uncertainty reporting	MU95 typical and MU95 conservative, k = 2
Ruggedness design	Full factorial 2^4 DoE, factors and levels in Annex

Introduction

Mycotoxins

Mycotoxins are secondary metabolic products of molds or ergot fungi, primarily from species such as *Aspergillus, Alternaria, Fusarium, Penicillium,* and *Claviceps*. If consumed, they can cause acute poisoning, chronic illnesses, and even cancer. Mycotoxins represent one of the greatest contamination risks for the food industry and are heavily regulated by the EU (see Regulation (EC) No. 1126/2007 and (EC) No. 915/2023). The regulated mycotoxins that can be measured with the SAFIA kits include ochratoxin A (OTA), fumonisins (FUM, sum of isomers FB1, FB2, FB3), deoxynivalenol (vomitoxin, DON), zearalenone (ZEN), aflatoxins (AFL, sum of isomers AFB1, AFB2, AFG1, AFG2), and the sum of T-2 and HT-2 toxins (T-2/HT-2). The maximum permitted levels depend on the type of food.

Principle of the SAFIA Assay

The suspension array fluorescence immunoassay (SAFIA) is a particle-based multiplexing rapid test. Coded microparticles are used for multiplexing, with coding based on different amounts of a red fluorescent dye incorporated into the microparticles. Each code, represented by a specific dye concentration, corresponds to a measured analyte. The measurement principle for detecting mycotoxins is based on indirect competitive immunoassays. The mycotoxins are chemically immobilized on the surface of the particles. A sample or standard, along with a mixture of mycotoxin-specific antibodies and fluorophore-labeled antibodies, is added to the particles. The specific antibodies competitively bind either the immobilized mycotoxin or the mycotoxin present in the sample. Bound antibodies are stained with dye (green fluorescent) labeled antibodies to generate a measurable signal. Due to the competitive reaction, the concentration of the mycotoxin is inversely proportional to the signal and can be determined using a calibration curve.

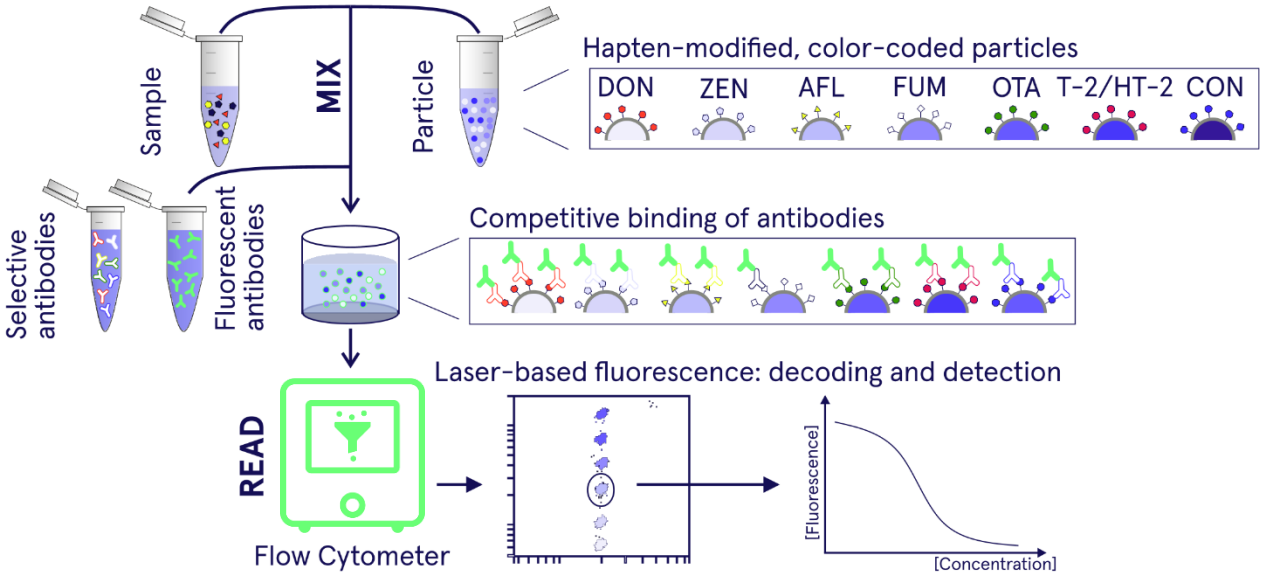


Figure 1. Schematic representation of the SAFIA working principle.

A flow cytometer reads the red fluorescence used for coding and the green fluorescence used for quantification. Within the flow cytometer, the SAFIA microparticles are hydrodynamically separated, and the fluorescence is measured independently for each particle using a blue laser as the excitation source, combined with two fluorescence detectors (green and red) and two scatter detectors (forward and sideward scatter).

Compared to classical immunoassays such as ELISA, SAFIA is a mix-and-read immunoassay. Washing steps, which are used to avoid high signal backgrounds, matrix interferences, or to stop signal increase, are not necessary.

In addition to mycotoxins, an internal control measurement ("Control") is carried out in the SAFIA. This indicates whether matrix effects interfere with the test during the measurement or if it was carried out correctly. Interpretation is done automatically in SAFIA-Score.

Scope of this validation

Building on previous validation work, this study evaluates the performance of the mycoProfile 6 Kit. The enhancements to the analytical procedure primarily comprise (1) the inclusion of the new sum parameter T-2/HT-2 and (2) an adjusted extraction procedure designed to reduce the sensitivity across all parameters.

The matrix scope is intended to cover a broad range of foodstuffs, excluding special matrices such as spices, cocoa, and coffee, for which a dedicated protocol is required. The validation includes both raw materials (for example, flours) and processed foods (for example, baked goods).

This validation assesses the full set of analytical performance characteristics, including calibration and working range, linearity, detection capability (Limit of Blank, Limit of Detection, and Limit of Quantification), accuracy (precision and trueness), measurement uncertainty, and the robustness and ruggedness of the method. This validation follows Eurachem Guide “The Fitness for Purpose of Analytical Methods” (3rd edition, 2025)”, when applicable to the method.

Materials and Methods

Materials and devices for Analysis

For the extraction of mycotoxins, denatured 99% ethanol (with MEK and Bitrex IPA, CHEMSOLUTE; Th.Geyer) was used. It was diluted to 70% (vol/vol) for the extraction of samples. Samples were extracted in 50 mL polypropylene centrifuge tubes and diluted in 2 mL polypropylene centrifuge tubes. Reagents were dispensed using reservoirs for multi-channel pipettes.

An analytical balance (Pioneer PX225D, Ohaus), a microtiter plate shaker (TITRAMAX 101, Heidolph Instruments), a rotator (LLG-UNILOOPMIX 2), and centrifuges (Megafuge 8R, Thermo Fisher, and MiniSpin, Eppendorf) were used for analysis. Sample Blank materials were ground and homogenized using a mill (A 11 basic Analysenmühle, IKA), dried fruits were frozen with liquid nitrogen prior milling. Single-channel pipettes (Research plus, Eppendorf) with matching polypropylene non-sterile tips from Brand LABSOLUTE and Eppendorf were used for dilution. For execution of the assay, electronic 8-channel pipettes (Xplorer, Eppendorf) were used in dispensing mode.

Reference Materials and Blank Materials

The Following reference/QC materials were used in the validation. The materials RM-15 to RM48 were obtained from the manufacturer as “QC materials” (materials from former interlaboratory comparison tests). RM-48 was solely used for the estimation of ruggedness.

Table 3. Overview of reference materials used in this study.

Sample Nr.	Name	Batch / Sample	Manufacturer	Matrix
RM-15	FCMO1-DRA13QC	T17215QC	FAPAS	Wine
RM-18	FCMM3-CCP32QC	T04482QC	FAPAS	Maize Flour
RM-22	FCMO1-DRA13QC	T04508QC	FAPAS	Oat Flour
RM-23	FCMF2-PRO17QC	T22206QC	FAPAS	Dried Pasta
RM-24	FCMM4-FRU38QC	T04484QC	FAPAS	Dried Figs (Slurry)
RM-32	FCMM4-NUT14QC	T04502QC	FAPAS	Pistachio (Slurry)
RM-33	FCMM10-CCP49QC	T22231QC	FAPAS	Wheat Flour
RM-34	FCMM3-CCP68QC	T04500QC	FAPAS	Durum Wheat
RM-36	FCMM4-NUT17QC	T04485QC	FAPAS	Sesame Paste (Tahini)
RM-46	FCMM4-CCP4QC	T04487QC	FAPAS	Barley Flour
RM-48	FCMM3-CCP68QC	T04535QC	FAPAS	Durum Wheat

Table 4. Mass fractions (Mean values) and target range values for the mycotoxins in the Reference materials. The target range of reference materials from Round-robin exercises are consistent within the range of -2 ≤ Z-Score ≤ +2. (n.d.: not defined)

	Mean Value of Mycotoxins and range of acceptable value (z = 2) [µg/kg]					
Sample Nr.	**Aflatoxins (AFL)**	**Deoxynivalenol (DON)**	**Zearalenone (ZEN)**	**Ochratoxin A (OTA)**	**Fumonisins (FUM)**	**T-2/HT-2 Toxin (T-2/HT-2)**
RM-15	n. d.	n. d.	n. d.	1.53 (0.86-2.20)	n. d.	n. d.
RM-18	4.72* (2.64 - 6.80)	1385 (963 - 1806)	186 (110 - 263)	2.52 (1.41 - 3.63)	640 (421 – 859)	104 (58 - 149)
RM-22	8.47 (4.74 - 12.19)	n. d.	n. d.	5.21 (2.92 - 7.50)	n. d.	190 (112 - 268)
RM-23	n. d.	769 (513 - 1024)	56 (31.4 - 80.6)	n. d.	n. d.	36.3 (20.3 - 52.2)
RM-24	7.72 (4.32 - 11.12)	n. d.	n. d.	9.72 (5.44 - 14.00)	n. d.	n. d.
RM-32	9.36 (5.24-13.48)	n. d.	n. d.	7.64 (4.28-11.0)	n.d.	n. d.
RM-33	8.63 (4.83-12.43)	796 (532-1059)	81.8 (45.8-117.8)	n. d.	n. d.	92.8 (52.0-133.6)
RM-34	2.31 * (1.29-3.33)	1240 (856-1624)	131 (74-188)	2.72 (1.52-3.92)	467 (299-635)	80.9 (45.3-116.5)
RM-36	8.56 (4.79-12.33)	n. d.	n. d.	8.25 (4.62-11.88)	n. d.	n. d.
RM-46	7.2 (4.03 - 10.37)	n. d.	n. d.	2.51 (1.40 - 3.61)	n. d.	n. d.
RM-48	1.88 * (1.05-2.71)	874 (589-1159)	122 (68-1159)	3.45 (1.93-4.96)	617 (405 – 830)	117 (66.0-168)

**Results given for aflatoxin B1. No indications for total aflatoxin content*

The following sample blank materials were used in the validation. All Materials were tested prior analysis and were found to be negative for the mycotoxins, except for LAB-13, which contained DON in quantities of approx. 50 µg/kg.

Table 5. Overviews of the blank materials.

Sample Nr.	Name	Brand	Manufacturer	Description/Ingredients
M-43	Zarte Haferflocken	K-Bio	Ceralia Getreideprodukte GmbH	whole grain oatmeal
M-53	Das Landknäcke nit Dinkel	Burger	Burger Knäcke GmbH + Co. KG	Whole grain rye flour, spelt flakes, whole grain spelt flour, table salt
M-101	Spätlese Prädikatswein	Rebenblick	Zimmermann-Graeff & Müller GmbH	White wine
M-104	Polenta Maisgrieß	K-Classic	Scheller Mühle GmbH	cornmeal
M-105	Maccaroni	K-Classic	Bon Pasta GmbH	durum wheat semolina
M-137	Butter Spritzgebäck	K Classic	Continental Bakeries Deutschland	Wheat flour, clarified butter (29%), milk, sugar, egg, table salt
LAB-1	-	-	-	Puff pastry pocket filled with Myzithra and feta cheese
LAB-2	-	-	-	Cheese rolls filled with Myzithra and feta cheese
LAB-13	-	-	-	spelt flour
LAB-14	-	-	-	wheat flour
LAB-17	-	-	-	Multigrain toast rolls
LAB-18	-	-	-	Frozen cake, plum

SAFIA Kits and Assay Procedure

The SAFIA kits with batch-number 45971-SCR/HT2-1L were used for analysis. Samples were extracted according to the kit protocol and measured at a fixed overall dilution factor of 8; all reported concentrations refer to the original sample. The procedure is described briefly.

Sample Preparation

1. 5 g of sample material was weighed in.
2. 10 mL of 70% (vol/vol) ethanol was added, and the samples were shaken for 15 min in an overhead shaker.
3. The mixture was centrifuged for 5 min at 1000 g, and 1 mL of the supernatant was collected.
4. 250 µL of sample was diluted in 750 µL of sample buffer. (The extract was diluted 1:4 (one part sample to 3 parts sample buffer). The dilution was briefly shaken and then centrifuged again at 12,000 g for 10 min.

Replicates and Plate Layouts

The analyses were performed by two operators to assess method robustness: operator 1 was highly experienced (more than 200 assays performed), whereas operator 2 had received training and had performed approximately 20 assays. Both were experienced in the execution of immunoassays.
On each plate, a full calibration curve comprising eight standards was included and measured in duplicate. A reagent blank (70% ethanol, diluted 1:4 in sample buffer) was also included on each plate and measured at least in duplicate. QC materials and negative matrices (sample blanks) were analyzed in total four times (two runs per operator) and measured in triplicate on each plate. LLOQ confirmation by spiking was performed in two runs per sample by operator 2 and measured in triplicate on each plate. ULOQ verification was conducted at two concentration levels; therefore, each sample was analyzed once per level in triplicate on each plate. For ruggedness evaluation, the 2^4 DoE runs were executed in duplicate, and each individual sample was measured in duplicate on a single plate.
The plate layout was block-randomized between runs to minimize systematic position effects, including potential plate-edge effects. Detailed plate layouts, including the distribution of standards, blanks, QC materials and samples for each run, are provided in the Annex.

Assay Execution

1. Initially, 25 µL of the diluted sample or standard was added to each well of the microtiter plate, following the selected plate layout.
2. Subsequently, the following were added sequentially:
    a. 10 µL particle working solution, which was thoroughly shaken for at least 20 seconds before filling
    b. 25 µL primary antibody (AK 1)
    c. 50 µL secondary antibody (AK 2)
3. The microtiter plate was then incubated for 20 minutes and shaken using a microtiter plate shaker.
4. Next, 50 µL of fixing solution was added, and the plate was incubated for an additional 5 minutes.
5. Subsequently, 140 µL of DI water was added.
6. The measurement in the flow cytometer was started

CyFlow Cube 6 V2m Measurement and Data Evaluation

After fixation of the assay, the measurement took place on the 1-laser flow cytometer CyFlow Cube 6 V2m. CyView software was used to control the cytometer. Device functionality was verified using SAFIA Check before starting the analysis. The provided configuration files for readout were utilized with specific settings: a flow rate of 0.5 µL/s and at least 10 µL per well or sample was analyzed.

The quantification of mycotoxins (processing raw FCS data) was performed using the automated gating function within SAFIA Score. Raw particle events were classified and analyzed using a density-based clustering algorithm. This method dynamically fits elliptical gates around population centroids within the dot-plots, ensuring that particle populations are correctly identified independent of potential signal drifts or operator bias.

Subsequently, the Median Fluorescence Intensities (MFI) were extracted from the corresponding histograms for each analyte. In SAFIA Score, these quantitative values were used to generate a 4-parametric calibration curve for the calculation of final sample concentrations.

$$MFI = \frac{A_1 - A_2}{1+\left(\frac{c(analyte)}{IC_{50}}\right)^p} + A_2 \qquad \textit{Equation 1}$$

, with MFI the measured signal in arb. u., A_1 the upper and A_2 the lower asymptote of the curve in arb. u., p the hillslope, IC_{50} the test-midpoint and c(analyte) the concentration in µg/L of the analyte.

The concentration of the analyte was calculated via equation:

$$c(analyte) = IC_{50}\left\{\left(\frac{A_1 - A_2}{MFI - A_2}\right) - 1\right\}^{\frac{1}{p}}$$ *Equation 2*

A total dilution factor of 8 was applied to all reported concentrations. A Grubbs Outlier test at significance level α = 0.05 was used to detect outlier measurements.

Analysis of Variance (ANOVA) was performed in jamovi (Version 2.6; The jamovi project, 2024), an open-source statistical spreadsheet built on the R language, using one way ANOVA for single factor comparisons and factorial ANOVA for the ruggedness DoE.

Correction Factors Used in SAFIA

For calibration, assay results for AFL and ZEN were adjusted using fixed recovery correction factors of 1.5 and 1.25, respectively.

The AFL parameter is designed to report the sum of aflatoxins B1, B2, G1 and G2. Because the antibody cross reactivity differs among the individual aflatoxins, the combined response is slightly biased towards the more strongly reacting congeners. In our system, cross reactivity ranges from approximately 100% for AFB1 down to about 62% for AFG2. Factor 1.5 is therefore applied to compensate for the lower response of the less reactive congeners and to improve agreement when mixed aflatoxin patterns are present.

The ZEN correction factor (1.25) was determined experimentally based on routine recovery observations, where recoveries around 75% were consistently obtained. A plausible explanation is the lower extraction efficiency of ZEN in the ethanol-based extraction procedure due to its comparatively lower polarity.

When interpreting results, these correction factors must be considered in spike experiments and QC materials. If samples are spiked exclusively with AFB1, applying the AFL correction factor would lead to an overestimation; therefore, results were back converted by dividing by 1.5 in this study for AFB1 only spike experiments and for QC materials indicating an AFB1 value only. For ZEN, the correction was applied to all samples, as all samples were affected by the lower extraction efficiency.

Results and Discussion

Calibration

SAFIA uses a sigmoidal calibration curve for quantification of the analytes. To assess precision and stability of the calibration curves, results from nine calibration curves were evaluated, four from the high skill operator and five from the medium skill operator. Each calibration point was measured in duplicate on each plate. Also, all calibration curves from both operators were averaged (Figure 2).

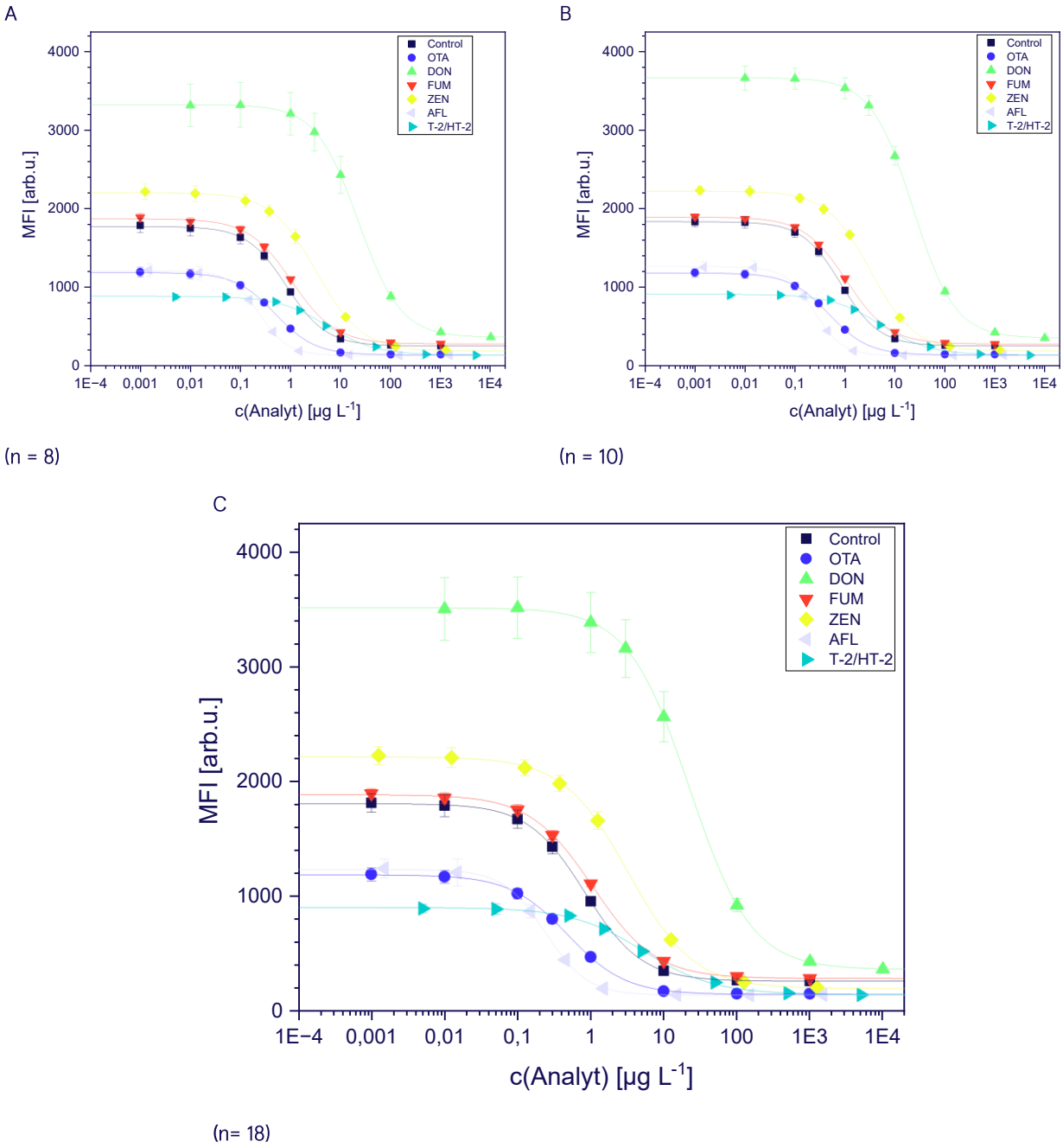


Figure 2. Average calibration curve of all SAFIA Assays. A) High Skill operator B) Medium Skill operator C) Both operators.

The calibration for both experiments was highly reproducible. Relative of standard deviation (RSD) values were only slightly lower for the highly skilled operator and still in the range below 8 % for both operators. This shows

that calibration has a high robustness. For a high-precision analysis this means that a calibration curve should be generated when switching between operators. The difference might be explainable e.g. in the Speed of the reagents added to the samples in the microtiter plate, since the incubation time is a critical factor for SAFIA, see section Ruggedness. If one operator performs SAFIA, it should be possible to run assays on different days, without the need for calibration on every single day. The repeatability of the calibration curves generated under the same conditions is thus very high.

Table 6. Minimal, Maximal and Median RSD values of median fluorescence intensity of each calibrant for all toxins.

Operator	RSD	Control	OTA	DON	FUM	ZEN	AFL	T-2/HT-2
High Skill	MIN	1.2%	3.0%	5.6%	2.0%	1.9%	2.8%	2.2%
	MAX	5.6%	5.3%	9.8%	2.8%	5.6%	7.3%	5.4%
	Mean	3.7%	4.1%	7.6%	2.6%	3.9%	4.8%	3.9%
Medium Skill	MIN	1.5%	2.5%	3.7%	1.8%	2.5%	3.1%	1.6%
	MAX	3.9%	7.3%	5.3%	3.4%	5.5%	9.1%	7.4%
	Mean	3.1%	4.8%	4.3%	2.4%	3.6%	5.9%	4.0%
Both	MIN	1.5%	2.9%	5.2%	2.4%	3.3%	3.5%	2.6%
	MAX	5.3%	6.1%	8.5%	3.2%	4.8%	9.7%	6.5%
	Mean	3.7%	4.6%	7.1%	2.6%	3.8%	6.1%	4.3%

Linearity

The linear region of the calibration curve corresponds approximately to the IC20 - IC80 values (20% and 80% inhibition compared to the upper asymptote). This region can be used as an estimate of the quantification limits, especially for untested matrices. As expected, no significant differences were found between operators. Hence, all values were averaged.

Table 7. IC20 and IC80 values for all analytes, generated by the average calibration curves (n = 18). The range from all individual curves is given in brackets.

Parameter	OTA (Min-Max)	DON (Min-Max)	FUM (Min-Max)	ZEN (Min-Max)	AFL (Min-Max)	T-2/HT-2 (Min-Max)
IC20 [µg/L]	0.136 (0.126 - 0.146)	5.85 (5.18 - 6.46)	2.69 (2.46 - 2.88)	0.786 (0.685 - 0.894)	0.0946 (0.0810 - 0.100)	1.10 (1.02 - 1.25)
IC80 [µg/L]	1.72 (1.59 - 1.84)	87.3 (81.7 - 92.1)	42.2 (39.9 - 45.1)	13.6 (12.0 - 14.7)	0.602 (0.548 - 0.635)	25.4 (22.8 - 26.4)

Selectivity (Cross reactivity)

Cross reactivity in an immunoassay is the extent to which the assay's antibody binds and responds to compounds other than the intended target analyte that share structural or epitope similarity. It is used as a practical measure of method selectivity because any binding to non-target compounds can generate a signal that is interpreted as target analyte and therefore causes positive bias.

Cross reactivity is typically expressed as a percentage relative to the target analyte, derived by comparing the concentrations that produce the same response, for example at 50% inhibition in a competitive assay:

$$CR = \frac{IC50_{Test}}{IC50_{Ref}} 100\ \% \qquad \textit{Equation 3,}$$

With $IC50_{Test}$ the IC50-value of the respective substance to be tested and $IC50_{Ref}$ the IC50-value of the corresponding reference substance for each analyte (i. e. AFB1 for AFL, Ochratoxin A for OTA, Deoxynivalenol for DON Fumonisin B1 for FUM, Zearelenon for ZEN and T-2 toxin for T-2/HT-2)

A higher cross reactivity indicates lower selectivity toward the target analyte, while low or non-detectable cross reactivity indicates high selectivity.

Cross reactivity was assessed using a two-step approach. First, a broad screening panel of potentially interfering compounds was tested at 10,000 µg/L and 100 µg/L. If no measurable inhibition was observed at 10,000 µg/L, cross reactivity was considered to be below 0.01%, as the IC50 Values would be higher than 10.000 µg/L. Compounds showing relevant inhibition in the screening step, defined as an apparent target concentration above the LOD, were evaluated in a second step by recording a full calibration curve using the respective compound. Cross reactivity was then quantified according to Equation 2 by comparing the compound's IC50 to the IC50 of the target analyte.

For chaetoglobosin A, citrinin, cyclopiazonic acid, citreoviridin, diacetoxyscirpenol, fumitremorgin B, fumitremorgin C, fumagillin, fusarenon X, gliotoxin, mycophenolic acid, penitrem A, roridin E, roridin L2, roquefortin C, rugulosin, satratoxin G, satratoxin H, stachybotrylactam, verruculogen and verrucarin A, no cross reactivity was observed for any analyte under the screening conditions. For AFL, DON, ZEN, FUM and OTA, cross reactivity towards the compounds listed in Table 6 had been evaluated previously. As the antibodies used in the kit are unchanged and cross reactivity is primarily an antibody specific parameter, no relevant differences are expected under the current assay conditions (70% ethanol extract diluted 1:4 in sample buffer). For DON, an updated cross reactivity of 0.05% was determined for nivalenol. For the T-2/HT-2 parameter, cross reactivity was observed for T-2 triol (4%) and HT-2 (95%); the high cross reactivity towards HT-2 is intended, as the assay is designed to detect the sum of T-2 and HT-2.

Important specific cross-reactivities (CR in %) are shown in Table 8.

Table 8. Cross-reactivity of selected compounds. The cross-reactivity for the unspecified parameters was less than 0.01 %.

Substance	CR (%)	Reference Parameters
Ochratoxin-B	49	OTA
Aflatoxin-B1	100	AFL
Aflatoxin-B2	64	AFL
Aflatoxin-G1	120	AFL
Aflatoxin-G2	62	AFL
Sterigmatocystin	0,02	AFL
Zearalanon	180	ZEN
α-Zearalenol	69	ZEN
α-Zearalanon	340	ZEN
β-Zearalenol	8	ZEN
β-Zearalanon	51	ZEN

Substance	CR (%)	Reference Parameters
15-Acetyl-DON	16	DON
3-Acetyl-DON	270	DON
Nivalenol	0.05	DON
Fumonisin B1	100	FUM
Fumonisin B2	140	FUM
Fumonisin B3	75	FUM
HT-2 toxin	95	T-2/HT-2
T-2 toxin	100	T-2/HT-2
T-2 triol	4	T-2/HT-2

Limit of Detection

Estimation of LODs via repetitive measurement of sample blanks

The Limit of Blank (LOB), Limit of Detection (LOD) were calculated according to DIN:ISO 32645 via

$$LOB = c_{blank} + 1{,}645\ \sigma_{blank}$$ *Equation 4,*

$$LOD = c_{blank} + 3\ \sigma_{blank}$$ *Equation 5,*

with c_{blank} the found mean concentration of the analyte in the blank material and σ_{blank} the standard deviation of the blank measurements. Note, that this approach was chosen instead of e. g. LOD = 3 σ_{Blank} because the mean values of blanks are usually different from zero. Values above the upper asymptote A_1 were set to "0.00 µg/kg". Values of all experimental runs were included in the calculation. The following values were generated from all measurements.

Table 9. Results of the negative samples. Given are the means of all measurements (4 x 3 = 12) and the standard deviation

Sample	OTA [µg/kg]	DON [µg/kg]	FUM [µg/kg]	ZEN [µg/kg]	AFL [µg/kg]	T-2/HT-2 [µg/kg]
Blank	0.111 ± 0.136	4.59 ± 4.21	0.243 ± 0.442	0.596 ± 0.622	0.148 ± 0.151	0.500 ± 0.695
M-43	0.0465 ± 0.0569	15.0 ± 5.11	0.932 ± 1.39	0.0987 ± 0.297	0.216 ± 0.181	7.20 ± 2.32
M-53	0.283 ± 0.182	25.0 ± 7.48	2.01 ± 2.56	0.554 ± 0.541	0.541 ± 0.180	7.37 ± 1.72
M-101	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
M-104	0.0345 ± 0.0559	14.6 ± 6.61	25.0 ± 12.9	0.184 ± 0.266	0.107 ± 0.138	2.14 ± 1.57
M-105	0.115 ± 0.149	18.0 ± 6.25	1.18 ± 1.53	0.597 ± 0.640	0.126 ± 0.141	2.18 ± 1.42
M-137	0.199 ± 0.173	27.6 ± 4.83	2.68 ± 1.81	0.911 ± 0.868	0.318 ± 0.202	2.51 ± 1.30
LAB-1	0.685 ± 0.191	9.11 ± 6.62	16.9 ± 3.27	1.16 ± 0.947	0.225 ± 0.222	0.893 ± 0.853
LAB-2	0.904 ± 0.158	20.4 ± 5.04	28.7 ± 3.05	2.22 ± 0.914	0.205 ± 0.185	1.08 ± 0.785
LAB-13	0.0885 ± 0.0966	> LOD	1.64 ± 1.97	1.20 ± 0.857	0.352 ± 0.218	6.03 ± 2.30
LAB-14	0.385 ± 0.143	24.5 ± 5.27	4.00 ± 2.23	1.52 ± 0.777	0.148 ± 0.154	0.935 ± 0.800
LAB-17	0.207 ± 0.200	8.17 ± 7.85	6.83 ± 3.19	0.00 ± 0.00	0.317 ± 0.200	0.941 ± 0.913
LAB-18	0.000510 ± 0.00169	0.00 ± 0.00	1.86 ± 2.34	0.00 ± 0.00	0.0332 ± 0.0752	0.544 ± 0.565

From these values, the following LODs were derived, see Table 10.

Table 10. Limit of blank and limit of detection according to DIN:ISO 32645, estimated via repetitive measurements of the negative samples and detection limits estimated via precision profile.

Parameter	OTA [µg/kg]	DON [µg/kg]	FUM [µg/kg]	ZEN [µg/kg]	AFL [µg/kg]	T-2/HT-2 [µg/kg]
$c_{blank} \pm \sigma_{blank}$	0.247 ± 0.306	14.8 ± 10.7	7.70 ± 10.6	0.704 ± 0.927	0.216 ± 0.223	2.65 ± 2.90
LOB	0.750	32.4	25.1	2.23	0.582	7.42
LOD	1.16	47.0	39.4	3.49	0.883	11.3

Estimation of LODs via precision profile of the calibration curves

For the estimation of the limit of detection, a precision profile approach can also be applied rather than the DIN:ISO 32645 blank value method. DIN:ISO 32645 is primarily formulated for the linear calibration case under repeatability conditions and relies on assumptions that are not fulfilled for indirect competitive immunoassays, namely a sigmoidal dose response with inverse proportionality between signal and concentration and a concentration dependent, heteroskedastic error structure across the calibration range. Under these conditions, transferring the DIN 32645 concept to a logistic calibration typically leads to overly pessimistic

detection limits because the error variance increases toward the low and high ends of the sigmoidal curve and is not constant as assumed in simple regression-based approaches. Therefore, the LOD was derived from the calibration-based precision profile obtained from all replicate measurements along the fitted logistic curve, expressing precision as a function of concentration. The Precision profile was calculated according to

$$x = -100\frac{\sigma(MFI)}{p(A_2-A_1)}\left[\left(\frac{IC_{50}}{c}\right)^p + 2 + \left(\frac{c}{IC_{50}}\right)^p\right]$$ *Equation 6,*

where x is the relative error of the concentration, and σ(MFI) is the standard deviation of the signal intensity of each calibration point. The precision profile was calculated for both operators separately and from the averaged calibration curves, see Figure 2 . To obtain a U-shaped precision profile, the following equation was used and fitted to the data. To avoid overfitting, extreme values at the ends of the calibration ranges were truncated.

$$PP = Ax^B + C + \frac{D}{x^E}$$ *Equation 7,*

where PP is the fitted precision profile, and A, B, C, D, E are the obtained parameters of the curve. The upper and lower detection ranges can be estimated at intersection at 30 % relative error of calibration. The LOD of the entire method can then be derived by multiplying the obtained value with the dilution factor of extraction and further dilution steps, which is for this method a factor of eight. As the precision profile shows a typical U-shape, there are two Limits of detection: A lower LOD (LLOD) and an upper LOD (ULOD). The range between LLOD and ULOD can be considered the working range of the assay.

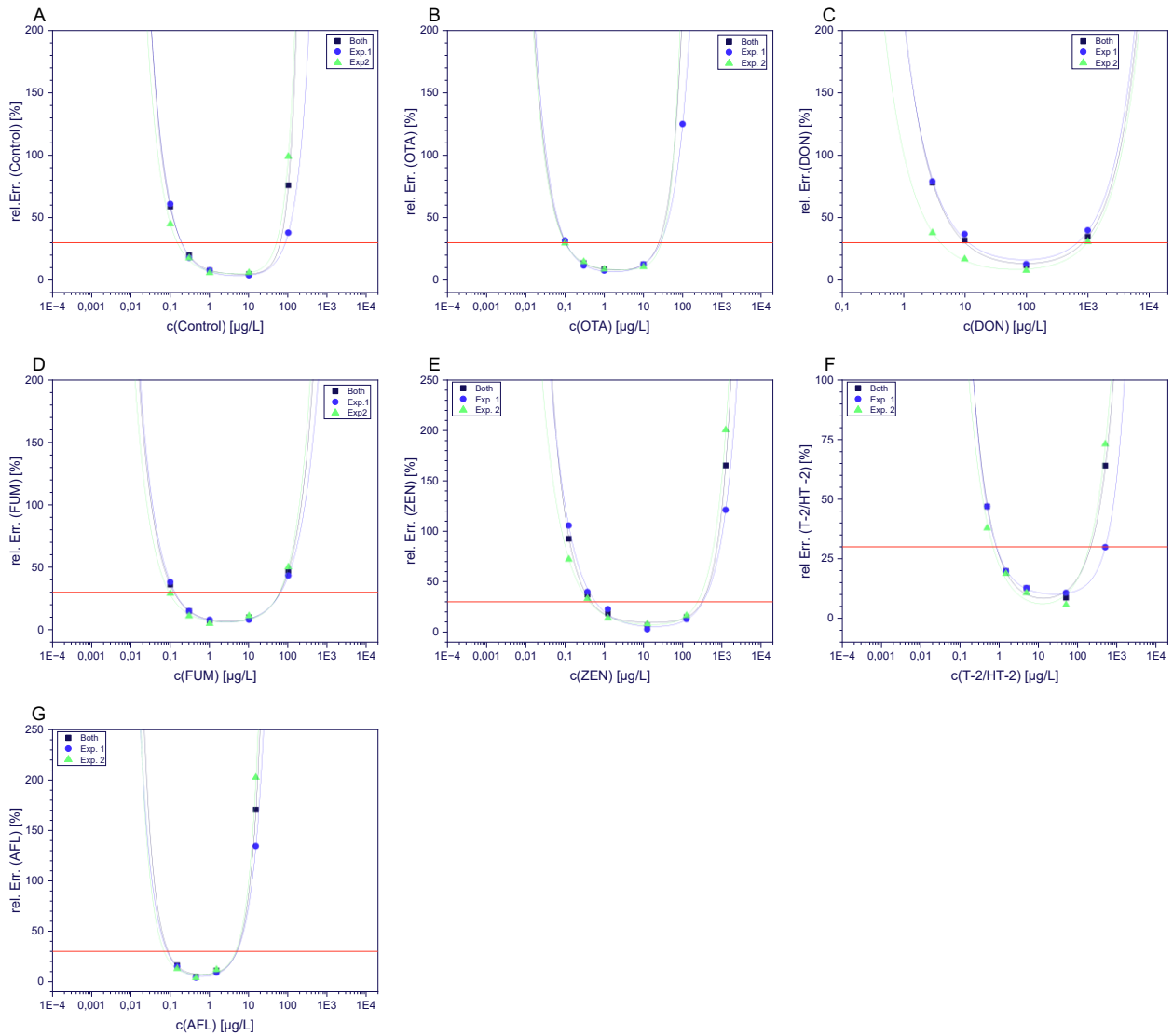


Figure 3. Precision profiles of the SAFIA calibration curves.

The precision profiles obtained by the two operators show very good agreement for the parameter Control, OTA, FUM, ZEN and AFL. Consequently, only minor differences in the estimated LOQs are expected for these analytes. For T-2/HT-2 and DON, small but noticeable deviations between the two precision profiles were observed. For DON, the precision profile of operator 2 suggests a lower LLOD, whereas for T-2/HT-2 the precision profile of operator 1 indicates a higher ULOD. These differences likely reflect normal operator-to-operator variability, for example slight differences in pipetting and handling, rather than systematic methodological effects. In line with this, operator 2 showed slightly lower standard deviations across the DON calibration curve, while operator 1 exhibited lower standard deviations in the upper concentration range of the T-2/HT-2 curve. Using the averaged precision profile across both operators is therefore considered the most representative approach for estimating LLOD and ULOD, as it reflects routine performance and yields conservative, realistic limits.

Table 11. Lower Limits of Detection derived from the precision profiles of the calibration curves. For comparison, the IC20 multiplied by the dilution factor (beginning of quasi-linear range) and values of the LOD from repetitive matrix measurements are given.

Parameter	OTA [µg/kg]	DON [µg/kg]	FUM [µg/kg]	ZEN [µg/kg]	AFL [µg/kg]	T-2/HT-2 [µg/kg]
Lower LOD Precision Profile	0.106	10.1	1.24	0.480	0.0891	0.861
8 x LLOD Precision Profile	0.846	81.0	9.91	3.84	0.713	6.89
For comparison						
8 x IC20	1.09	46.8	21.5	6.29	0.757	8.82
LOD (Blank Matrix)	1.16	47.0	39.4	3.49	0.883	11.3

Only for AFL and ZEN comparable values were achieved via both methods. The LLOD of the precision profile received for OTA, FUM and T-2/HT-2 are smaller, compared to the LODs of the DIN:ISO 32645 method. The reason for this is the higher background, and thus higher standard deviation received from repetitive blank measurements, leading to a higher LOD. For DON, a lower LOD is achieved via the precision profile. The reason for this lies within the higher standard deviations of the averaged calibration curves of both operators. Hence, the influence of the operator on the DON-Curve is higher than for all other curves, which results in a higher LOD.

The Upper LODs of the averaged curve are far below the IC80, even far below the IC90 (90 % inhibition against the upper asymptote of the curve) points of the calibration curves. For Measurement in real samples, it must be considered that also matrix-effects might influence the measurement, leading to measurements with high uncertainty in this "shallow" region of the calibration curves. To give a realistic estimate of the ULOD we conducted spiking experiments, also to include effects of the matrix, see section Limits of Quantification and working range, and include a much smaller range. The values obtained here should be therefore considered as absolute maximum values only obtainable under ideal conditions, i.e. absence of a matrix.

Table 12. Upper Limits of Detection derived from the precision profiles of the calibration curves. For comparison, the IC80 values (end of quasi-linear range) and IC90 multiplied by the dilution factor are given.

Parameter	OTA [µg/kg]	DON [µg/kg]	FUM [µg/kg]	ZEN [µg/kg]	AFL [µg/kg]	T-2/HT-2 [µg/kg]
Upper LOD Precision Profile	23.8	814	607	297	4.77	210
8 x ULOD Precision Profile	190	6510	4860	2380	38	1680
For Comparison						
8 x IC80	13.7	698	338	109	4.81	200
8 x IC90	28.8	1540	755	201	8.26	500.0

A consensus value for LODs from both approaches

According to Eurachem, the limit of detection (LOD) is the lowest analyte concentration that can be detected with a specified level of confidence. In statistical terms, the LOD is commonly treated as a decision threshold

chosen to control the probability α of falsely classifying a blank or negative sample as positive, with α often set to 5% (corresponding to 95% confidence). To select an LOD definition suitable for routine testing, several candidate thresholds were compared against the empirical blank distribution (n = 148 to 160 per analyte). The criterion LOD = $c_{blank} + 3\sigma_{blank}$ resulted in at most one exceedance per analyte (≤ 0.6%), consistent with a low false positive risk. In contrast, IC20 based and precision profile-based thresholds were exceeded more frequently for some analytes, indicating an unacceptably permissive decision threshold when interpreted as LOD. We therefore report the blank based LOD as the detection decision threshold and use spiking confirmed LLOQ and ULOQ to define the quantitative working range.

Table 13. Comparison of LOD estimates (blank-based, IC20, and precision-profile LLOD) and the fraction of measurements below each threshold for OTA, DON, FUM, ZEN, AFL, and T-2/HT-2.

Parameters	OTA	DON	FUM	ZEN	AFL	T-2/HT-2
Total measurements	160	148	160	160	160	160
$LOD = c_{blank} + 3\,\sigma_{blank}$						
LOD	**1.16**	**47.0**	**39.4**	**3.49**	**0.883**	**11.3**
Number of measurements > LOD	0	0	0	1	0	1
Fraction < LOD	100%	100%	100%	99.4%	100%	99.4%
LOD = IC20						
8 x IC20	1.09	46.8	21.5	6.29	0.757	8.82
	1	0	18	0	2	7
Fraction < IC20	100%	99.4%	88.8	100%	100%	98.8%
LLOD of precision profile						
8 x LLOD Precision Profile	0.846	81.0	9.91	3.84	0.713	6.89
Number of measurements > LLOD Precision Profile	11	0	36	0	3	19
Fraction < LLOD Precision Profile	93.1%	100%	77.5%	100%	98.1%	88.1%

Limits of Quantification and working range

The limit of quantification (LOQ) is the lowest measurand level at which the analyte can be quantified with acceptable performance, where acceptability is defined by the fitness for purpose requirements, typically expressed in terms of precision, trueness and or measurement uncertainty. For the SAFIA assay, LOQ was assessed following the Eurachem method validation guide using experimental confirmation as described in Quick Reference 4(c). Candidate quantification limits derived from the calibration function and the precision profile were verified by spiking an appropriate blank matrix at the proposed lower and upper LOQ levels and analyzing replicate samples under intermediate precision conditions. Trueness (recovery), precision (RSD) and the expanded measurement uncertainty were evaluated against predefined performance criteria. The lower

and upper LOQ were defined as the lowest and highest concentrations at which these criteria were fulfilled, thereby establishing the working range of the assay.

In competitive immunoassays, the error variance is typically heteroskedastic and strongly concentration dependent due to the sigmoidal calibration function and the indirect response format. Therefore, a simple blank based rule such as $LOQ=c_{blank} + 10 \cdot \sigma_{blank}$ can be overly conservative, as it is dominated by the comparatively high variability observed at or near the blank and may underestimate the practically achievable quantification performance at low but nonzero concentrations. For example, the relative standard deviation of OTA derived from blank measurements was 124%, whereas the typical precision observed for nonzero materials in this study, such as the QC materials, averaged 8%.

For LLOQ estimation, the following concentration levels were selected, see Table 14. First, they are above the IC20 region and therefore fall within the quasi-linear part of the calibration curve. Second, except for DON, they also lie within the LOD ranges derived from the precision profiles (the higher LLOD observed for DON is discussed above). Third, the selected levels are at or above the estimated LOD, ensuring that quantification is assessed at concentrations where reliable detection has already been demonstrated.

For ULOQ estimation, upper LOQ values were determined in a two-step approach. First, provisional ULOQ levels were selected based on the IC90 values, as the precision profiles indicated that quantification should be feasible in this concentration range. During experimental confirmation, however, it became evident that these initial ULOQ estimates were too high for AFL, DON, ZEN, FUM and T-2/HT-2, as performance criteria were no longer met. The proposed ULOQ levels were therefore reduced and reassessed until acceptable recovery, intermediate precision and measurement uncertainty were achieved (Table 14).

Table 14. Estimated LLOQS and ULOQS of the SAFIA Assay

Parameter	Spiked LLOQ [µg/kg]	Spiked ULOQ 1 {µg/kg]	Spiked ULOQ 2 [µg/kg]
AFL	0.900	20.0	15.0
DON	50.0	2000	1500
FUM	50.0	1000	700
OTA	1.20	50.0	50
T-2/HT-2	11.0	500	300
ZEN	3.00	400	300

Consequently, blank matrices were spiked at the target analyte concentrations and analyzed in triplicate. In addition, a reagent blank (70% ethanol) was spiked in parallel to assess potential bias introduced by the spiking procedure. To evaluate intermediate precision, the complete experiment was repeated on a second measurement day, for the LLOQ Levels.

For the LLOQ experiments, results were blank-corrected using matrix-matched blanks. Specifically, for each matrix the mean result of the corresponding unspiked blank material was subtracted from the mean result of the spiked material (e.g., ethanol blank from ethanol spike, M-43 blank from M-43 spike and so on). Blank and spiked samples were not measured within the same analytical run; therefore, blank correction was applied on the level of averaged results across runs to mitigate systematic background contributions observed for some matrices at very low concentrations. For the ULOQ experiments, no blank correction was applied, since the

spiking levels were at least one order of magnitude higher than the responses observed in negative samples and the influence of the blank was therefore negligible. Results for the individual samples are provided in the Annex.

During these measurements, two samples showed clear, analyte specific and level dependent matrix interferences. Sample LAB-18 (plum cake) resulted in a strong underestimation of DON and, in particular, ZEN, while sample LAB-17 (multigrain toast rolls) showed an underestimation of ZEN. The effect occurred predominantly at the lower spiking levels, where ZEN even yielded an apparent zero signal, which is unusual for competitive immunoassays where positive bias is more commonly observed. The level dependence suggests a concentration dependent interference, for example reduced extraction efficiency close to the LOQ or specific binding or chemical transformation that becomes less relevant at higher analyte levels. A general fluorescence related effect, such as increased background or quenching, is unlikely because none of the other analytes were affected. The underlying cause remains unclear and may reflect a matrix component acting selectively on DON and or ZEN. For data evaluation, these two samples were excluded for DON and ZEN.

As a benchmark for LOQ confirmation, the accuracy study results were used, because quantification of all analytes in those samples was satisfactory (all z-scores were within −2 to +2). The LOQ was therefore considered confirmed when measurements at the proposed LOQ levels achieved medium recovery rate between 70 and 130 %, the target range for SAFIA. The trueness (recovery), (intermediate) precision, and expanded measurement uncertainty should also be comparable to the one obtained from measurement of QC materials. One-way ANOVA was used to identify and document potential systematic performance differences between the concentration levels investigated. However, the selection of LLOQ and ULOQ was primarily based on compliance with the predefined fitness-for-purpose criteria (recovery, intermediate precision, and MU). Statistical significance between levels was reported as supporting information and does not, by itself, imply that a level is unsuitable if the acceptance criteria are met. The results are summarized in Table 15.

The expanded measurement uncertainty (MU95, k = 2) was estimated using a top-down approach based on the observed performance in the spiking experiments. For low-level experiments where blank correction was applied, the corrected concentration was calculated as the difference between the mean result of the spiked sample and the mean result of the matrix-matched blank. The associated standard uncertainty was then estimated from the combined contributions of the bias and the variability of both terms, i.e.

$$u \approx \sqrt{b^2 + s_{sample}^2 + s_{blank}^2}$$ *Equation 8,*

where b is the deviation from the nominal spike level and s_{sample} and s_{blank} are the standard deviations of the measured spiked and blank samples (as intermediate precision of all runs), respectively. For higher concentration levels where no blank correction was applied, the uncertainty was estimated analogously but without the blank term, i.e.

$$u \approx \sqrt{b^2 + s_{sample}^2}.$$ *Equation 9.*

The expanded uncertainty was obtained as

$$MU_{95} = k * u, k = 2$$ *Equation 10*

The uncertainty of the spiking standards (purity and preparation of spiking solutions) was not included. This is justified because the standards were supplied with high stated purity and the spiking solutions were prepared using the same balance, pipettes and procedure throughout the study. Consequently, any residual systematic contribution from standard preparation would be largely common mode across all spike levels and matrices

and would not affect the relative comparison of performance between levels; the reported MU therefore reflects the analytical method performance (extraction and measurement) rather than the uncertainty of the reference spike value.

Table 15. Comparison of Recovery, Precision and measurement uncertainty of the different spiking levels (LLOQ, ULOQ 1 and ULOQ 2). Given are the mean values and the range of the minimum to maximum. For ULOQS, instead of intermediate precision, the intra-assay precision is reported.

Accuracy Criterion	Parameter	AFL	DON	FUM	OTA	T-2/HT-2	ZEN
Recovery [%]	LLOQ Level	98% (75% - 124%)	88% (74% - 119%)	96% (79% - 112%)	94% (78% - 113%)	102% (85% - 154%)	102% (82% - 125%)
	ULOQ Level 1	93 (74 - 116)	85 (80 - 90)	86 (77 - 98)	**123 (96 - 158)****	**80 (65 - 99)***	101 (86 - 119)
	ULOQ Level 2	**72 (65 - 81)****	**63 (54 - 69)****	**34 (26 - 40)****	93 (61 - 158)	**39 (34 - 50)****	**87 (82 - 98)***
	QC-material	92 (63 - 119)	102 (86 - 112)	109 (105 - 113)	93 (69 - 112)	111 (94 - 126)	95 (83 - 104)
Precision (RSD, [%])	LLOQ Level	8 (4 - 13)	14 (10 - 26)	11 (4 - 22)	9 (5 - 19)	15 (10 - 21)	11 (5 - 17)
	ULOQ Level 1	10 (2 - 16)	**5 (1 - 7)****	6 (4 - 11)	6 (2 - 15)	7 **(3 - 13)****	**4 (2 - 7)***
	ULOQ Level 2	13 (5 - 22)	**7 (3 - 10)****	**5 (2 - 7)****	8 (3 - 15)	**11 (7 - 13)****	7 (3 - 10)
	QC-material	9 (5 - 14)	11 (7 - 14)	14 (13 - 15)	8 (4 - 11)	13 (7 - 17)	10 (3 - 17)
MU_{95} (k =2) [%]	LLOQ Level	49 (30 - 82)	**63 (35 - 92)****	40 (23 - 92)	40 (20 - 59)	51 (31 - 84)	**60 (41 - 89)****
	ULOQ Level 1	40 (9 - 74)	**37 (22 - 51)****	38 (11 - 62)	38 (11 - 80)	55 (14 - 109)	**20 (12 - 34)****
	ULOQ Level 2	85 (54 - 114)	**122 (90 - 171)****	**403 (297 - 560)****	55 (17 - 129)	**315 (200 - 382)****	38 (16 - 68)
	QC-material	60 (17 - 118)	**29 (14 - 38)****	32 (26 - 37)	43 (12 - 92)	36 (23 - 48)	38 (16 - 68)

*** Indicate statistically difference ($p < 0.05$) to one other Level or the QC Material form the One-Way ANOVA.*

All proposed LOQS and ULOQs were experimentally confirmed, as recovery, intermediate or intra-assay precision, and MU95 at the selected upper levels were comparable to the QC material and met the predefined fitness-for-purpose criteria. For AFL, DON and FUM, the initially proposed highest levels (20 µg/kg, 2000 µg/kg and 1000 µg/kg, respectively) showed significantly reduced recoveries and or increased uncertainty; therefore, the refined ULOQs were set to 15 µg/kg (AFL) and 1500 µg/kg (DON). (For AFL, performance at 20 µg/kg only marginally met the recovery criterion; therefore, a more conservative ULOQ of 15 µg/kg was selected.) For FUM, performance at the highest level was not acceptable and the ULOQ was limited to the lower confirmed

level accordingly. OTA showed different recoveries between the two 50 µg/kg spike sets, which is attributed to spiking variability; however, results remained within the performance range observed for QC materials and the ULOQ of 50 µg/kg was confirmed. For T-2/HT-2, recoveries were significantly lower at both investigated upper levels, but the refined level (300 µg/kg) still met the target recovery range (70 to 130%) and was therefore selected as ULOQ. For ZEN, MU95 at the LLOQ level was higher, likely due to the applied blank correction; although recovery at the higher ZEN level was statistically different, it remained within the acceptance limits, and the refined ULOQ was set to 300 µg/kg (conservative) based on the overall performance across levels. For ZEN at the LLOQ level, the higher MU95 is plausibly explained by the applied blank correction. The corrected result is calculated as a difference (spiked sample minus matrix-matched blank), and the variability of both terms contributes to the variance of the corrected value. Consequently, blank correction can increase the relative uncertainty at very low concentration levels even when trueness is not degraded; this behavior is therefore expected and is reflected in the MU results.

At the higher spike levels used for ULOQ assessment, precision was generally very good, which is typical for competitive immunoassays at the upper end of the calibration range. However, despite low variability, the bias can increase substantially in this region due to reduced sensitivity and saturation effects of the sigmoidal calibration function (signal compression). Consequently, a sample with a true concentration above the ULOQ may still yield an apparent result within the working range, leading to underestimation, and the expanded measurement uncertainty may become large because it is dominated by the bias contribution rather than by imprecision. (See to spiking Levels in Table 15, that exceed the ULOQ in Table 16.) This behavior is consistent with the U-shaped precision profile commonly observed for competitive immunoassays, where uncertainty is lowest in the quasi-linear mid-range and increases towards both ends of the working range. As a practical recommendation for routine testing, results close to the ULOQ or flagged as '> ULOQ' should be confirmed by re-measurement at a higher dilution and checked for dilution conformity (back calculated concentrations should agree within the expected tolerance). Quantitative reporting should be based on the dilution level that places the analyte well within the confirmed working range.

The validated working ranges (LLOQ - ULOQ) for the assay is summarized in Table 16. Because the LOQ is conceptually expected to be equal to or higher than the LOD, and given that the determined values for ZEN (LLOQ 3.0 µg/kg vs LOD 3.49 µg/kg) and T-2/HT-2 (LLOQ 11.0 µg/kg vs LOD 11.3 µg/kg) are within a very small margin that cannot be meaningfully resolved under routine conditions, the LOQs were aligned to the corresponding LOD estimates. Therefore, the reported LLOQ values were conservatively set to 3.5 µg/kg for ZEN and 11.5 µg/kg for T-2/HT-2 to ensure internal consistency of performance characteristics and to avoid implying quantification capability below the demonstrated detection capability.

Table 16. Working ranges of the SAFIA (LLOQ - ULOQ).

Analyte	Working range (LLOQ – ULOQ) [µg/kg]
AFL	0.900 – 15
DON	50.0 – 1500
FUM	50.0 - 700
OTA	1.20 – 50
T-2/HT-2	11.5 - 300
ZEN	3.50 - 300

Measurement of QC (Reference) Material – Estimation of Accuracy

The following results summarize all measurements of the QC Materials, excluding RM-48, which were solely used for ruggedness tests. The QC Materials were analyzed 4 times (two by each operator) in triplicates in each run. Mean values and standard deviations are calculated from all measurements executed. The individual values were derived from the calibration curve executed on the day of measurement. The results are presented for each mycotoxin. Recovery was calculated according to

$$Recovery = \frac{c(Found)}{c(Reference)} \cdot 100\%$$ *Equation 11*

with c(found) the mass fraction measured in SAFIA (n =15) and c(Reference) the reference mass fraction of the materials. **The target range for recovery in SAFIA is 70 - 130 %.** Z- Scores were calculated according to

$$Z = \frac{c(Found) - c(Reference)}{\sigma(Reference)}$$ *Equation 12*

with σ(Reference) the provided standard deviation of the reference concentration.
The combined standard measurement uncertainty u(c) was estimated using a top-down approach based on within laboratory performance data, considering both intermediate precision and measurement bias. The combined uncertainty was calculated as

$$u(c) = \sqrt{{u_{prec}}^2 + {u_{bias}}^2}$$ *Equation 13,*

where u_{prec} and u_{bias} are the standard uncertainty contributions from precision and bias, respectively. The precision contribution u_{prec} was taken as the intermediate precision standard deviation of the set of independent results obtained across all measurement days and both operators (inter run variability under routine conditions). The uncertainty contribution associated with bias was estimated as

$$u_{bias} = \sqrt{b^2 + u(X_a)^2 + \frac{\sigma_{IA}^2}{n}}$$ *Equation 14,*

with the observed bias b

$$b = \bar{X} - X_A$$ *Equation 15,*

where X is the mean result obtained by SAFIA for the QC material and X_A is the assigned value of the QC material. The term $\frac{\sigma_{IA}^2}{n}$ accounts for the uncertainty of the estimated mean X, with n being the number of independent results contributing to X.

Because the materials used were QC items originating from proficiency testing rounds rather than certified reference materials, $u(X_a)$ was approximated from the proficiency test consensus statistics. Following ISO 13528, the standard uncertainty of an assigned value derived from a robust consensus can be estimated as

$$u(X_a) = \frac{1{,}25 * \sigma_{PT}}{\sqrt{p}}$$ *Equation 16,*

where p is the number of participant results used to produce the assigned value and σ_{PT} is the standard deviation for proficiency assessment. In the absence of a directly reported σ_{PT}, it was approximated from the satisfactory range given for $|z| \leq 2$. Since the z score is defined as

$$z = \frac{x - X_A}{\sigma_{PT}}$$ *Equation 17,*

the reported interval corresponds approximately to $X_A \pm 2\,\sigma_{PT}$. Therefore,

$$\sigma_{PT} = \frac{Range\ for\ |z|<2}{4} \qquad \textit{Equation 18}$$

Finally, the expanded measurement uncertainty at approximately 95% coverage was calculated as

$$U(c, 95\%) = k * u(c),\ \ k = 2, \qquad \textit{Equation 19.}$$

Sum of Aflatoxins B1, B2 and G1, G2 (AFL)

Table 17. Results of the AFL determination in the reference materials (n =12).

Matrix	Sample	Result [µg/kg]	RSD	MU (k= 2) [µg/kg]	MU (k= 2) [%]	Reference [µg/kg]	Recovery	Z-Score
Maize Flour	RM-18	5.28 ± 0.455*	8.6%	1.32	25%	4.72 (2.64 - 6.80)	112%	0.54
Oat Flour	RM-22	8.74 ± 0.877	10.0%	1.47	17%	8.47 (4.74 - 12.19)	103%	0.14
Dried Figs	RM-24	4.87 ± 0.340	7.0%	5.74	118%	7.72 (4.32 - 11.12)	63%	-1.67
Pistachio	RM-32	6.98 ± 0.953	13.7%	5.10	73%	9.36 (5.24 - 13.48)	75%	-1.15
Wheat Flour	RM-33	10.2 ± 0.850	8.4%	6.19	56%	8.63 (4.83 - 12.43)	118%	0.81
Durum Wheat	RM-34*	2.75 ± 0.352	12.8%	0.93	35%	2.31 (1.29 - 3.33)	119%	0.86
Sesame Paste (Tahini)	RM-36	6.16 ± 0.483	7.9%	5.12	83%	8.56 (4.79 - 12.33)	72%	-1.28
Barley Flour	RM-46	5.44 ± 0.263	4.8%	3.84	71%	7.2 (4.03 - 10.37)	76%	-1.11

** Values corrected for AF-B1 results.*

For RM-18 and RM-34, the reference material provider reported assigned values for aflatoxin B1 only. Results are reported as total aflatoxins. For comparison to assigned values spiked with AFB1 only, results were converted to AFB1 equivalents by dividing by 1.5. This conversion is only justified in this specific case because the materials were spiked exclusively with aflatoxin B1 and the manufacturer indicated that no other aflatoxins were present in these samples. In total, the mean recovery rate for AFL was 92%, spanning from 63 – 119%. All Z-Score were in the satisfactory range from -2 to 2.

Ochratoxin A (OTA)

Table 18. Results of the OTA determination in the reference materials (n =12).

Matrix	Sample	Result [µg/kg]	RSD	MU (k= 2) [µg/kg]	MU (k= 2) [%]	Reference [µg/kg]	Recovery	Z-Score
Wine	RM-15	1.06 ± 0.114	10.7%	0.971	92%	1.53 (0.86 - 2.20)	69%	-1.40
Maize Flour	RM-18	2.76 ± 0.281	10.2%	0.78	28%	2.52 (1.41 - 3.63)	110%	0.44
Oat Flour	RM-22	4.65 ± 0.182	3.9%	1.34	29%	5.21 (2.92 - 7.50)	89%	-0.49
Dried Figs	RM-24	9.93 ± 0.379	3.8%	1.17	12%	9.72 (5.44 - 14.00)	102%	0.10
Pistachio	RM-32	6.80 ± 0.746	11.0%	2.42	36%	7.64 (4.28 - 11.0)	89%	-0.50
Durum Wheat	RM-34	3.04 ± 0.275	9.1%	0.903	30%	2.72 (1.52 - 3.92)	112%	0.53
Sesame Paste (Tahini)	RM-36	5.75 ± 0.301	5.2%	5.32	92%	8.25 (4.62 - 11.88)	70%	-1.38
Barley Flour	RM-46	2.60 ± 0.260	10.0%	0.73	28%	2.51 (1.40 - 3.61)	103%	0.16

The average recovery for OTA was very good in the ranges between 69 - 112 %, with an average of 93 %. The found Z-Scores are also on the ranges between -2 and 2. Therefore, it can be expected that the determination of OTA in real samples will proceed with high accuracy and precision.

Fumonisins (FUM)

Table 19. Results of the FUM determination in the reference materials (n =12).

Matrix	Sample	Result [µg/kg]	RSD	MU (k= 2) [µg/kg]	MU (k= 2) [%]	Reference [µg/kg]	Recovery	Z-Score
Maize Flour	RM-18	674 ± 84.5	12.6%	196	37%	640 (421 - 859)	105%	0.31
Durum Wheat	RM-34	527 ± 79.1	15.0%	177	26%	467 (299 - 635)	114%	0.71

The average recovery for FUM was 110%, with both Z-Scores also within the acceptable range.

Deoxynivalenol (DON)

Table 20. Results of the DON determination in the reference materials (n =12).

Matrix	Sample	Result [µg/kg]	RSD	MU (k= 2) [µg/kg]	MU (k= 2) [%]	Reference [µg/kg]	Recovery	Z-Score
Maize Flour	RM-18	1190 ± 131	11.0%	449	38%	1385 (963 - 1806)	86%	-0.93
Dried Pasta	RM-23	771 ± 53.7	7.0%	106	14%	769 (513 - 1024)	100%	0.02
Wheat Flour	RM-33	878 ± 122	13.9%	300	34%	796 (532 - 1059)	110%	0.62
Durum Wheat	RM-34	1390 ± 149	10.7%	420	30%	1240 (856 - 1624)	112%	0.76

Also, DON could be measured in the QC materials with high accuracy, with a mean recovery of 102 % and all Z-Score between -1 and +1.

Zearalenone (ZEN)

Table 21. Results of the ZEN determination in the reference materials (n =12).

Matrix	Sample	Result [µg/kg]	RSD	MU (k= 2) [µg/kg]	MU (k= 2) [%]	Reference [µg/kg]	Recovery	Z-Score
Maize Flour	RM-18	194 ± 32.9	17.0%	71.4	37%	186 (110 - 263)	104%	0.21
Dried Pasta	RM-23	46.3 ± 2.52	5.4%	21.6	47%	56 (31.4 - 80.6)	83%	-0.79
Wheat Flour	RM-33	77.0 ± 12.6	16.3%	30.2	39%	81.8 (45.8 - 117.8)	94%	-0.27
Durum Wheat	RM-34	123 ± 10.4	8.5%	27.9	23%	131 (74 - 188)	94%	-0.28

ZEN was quantified with high accuracy and precision in the QC-samples, with an average recovery rate of 94% and all Z-Score lying within -1 and +1.

Sum of T-2/HT-2 Toxins (T-2/HT-2)

Table 22. Results of the T-2/HT-2 determination in the reference materials (n =12).

Matrix	Sample	Result [µg/kg]	RSD	MU (k= 2) [µg/kg]	MU (k= 2) [%]	Reference [µg/kg]	Recovery	Z-Score
Maize Flour	RM-18	112 ± 19.1	17.1%	36.7	33%	104 (58 - 149)	107%	0.34
Oat Flour	RM-22	178 ± 12.1	6.8%	41.8	23%	190 (112 - 268)	94%	-0.31
Dried Pasta	RM-23	45.7 ± 5.88	12.9%	22.2	48%	36.3 (20.3 - 52.2)	126%	1.18
Wheat Flour	RM-33	99.2 ± 15.8	16.0%	35.7	36%	92.8 (52.0 - 133.6)	107%	0.31
Durum Wheat	RM-34	97.5 ± 12.5	12.8%	38.4	39%	80.9 (45.3 - 116.5)	121%	-0.43

The sum of T-2 and HT-2 toxins was quantified with a mean recovery of 111 % and all Z-Sore being in the range of -1 to +2.

Trueness

The combined trueness, expressed as Recovery form all samples -spiked and QC Materials is summarized in Table 23.

Table 23. Trueness, expressed as recovery from all samples 8QC and spiked) within this study.

Parameter	N (Samples)	Mean Recovery [%]	Median Recovery [%]	Min Recovery [%]	Max Recovery [%]	N out of 70-130 % range
AFL	28	95%	96%	63%	124%	1
DON	20	90%	85%	74%	119%	0
FUM	22	94%	91%	77%	113%	0
OTA	36	98%	95%	61%	158%	5
T-2/HT-2	25	97%	94%	65%	154%	2
ZEN	21	100%	97%	82%	125%	0

Trueness of the SAFIA assay was summarized across all recovery datasets within the confirmed working range, including QC materials and LLOQ and ULOQ spiking experiments. Mean recoveries ranged from 90% (DON) to 100% (ZEN), and medians were close to the means, indicating no pronounced systematic skew. For most analytes, all or nearly all results fell within the target range of 70 to 130%. A small number of out-of-range recoveries was observed, mainly for OTA (5 of 36), and to a lesser extent for T-2/HT-2 (2 of 25) and AFL (1 of 28). Most OTA deviations occurred at the higher spiking level (50 µg/kg), suggesting that part of the spread may originate from spiking or sample preparation variability rather than a systematic method bias, as small absolute pipetting or mixing errors can translate into noticeable recovery shifts at elevated levels. These results were retained for the overall trueness and uncertainty evaluation to reflect routine conditions and potential matrix or level dependent effects, particularly towards the edges of the working range.

Measurement of QC (Reference) Material – Estimation of Precision

Intra-assay precision (repeatability)

Repeatability was evaluated as intra-assay precision on each microtiter plate using all positive reference materials, meaning reference concentrations above the LOQ for the respective analyte. For each plate, the coefficient of variation (CV) was calculated from replicate determinations and summarized by analyte and operator group. The acceptance criterion for repeatability was a mean intra-assay CV below 10%. As shown in Figure 4 and Table 24, this criterion was met for all analytes.

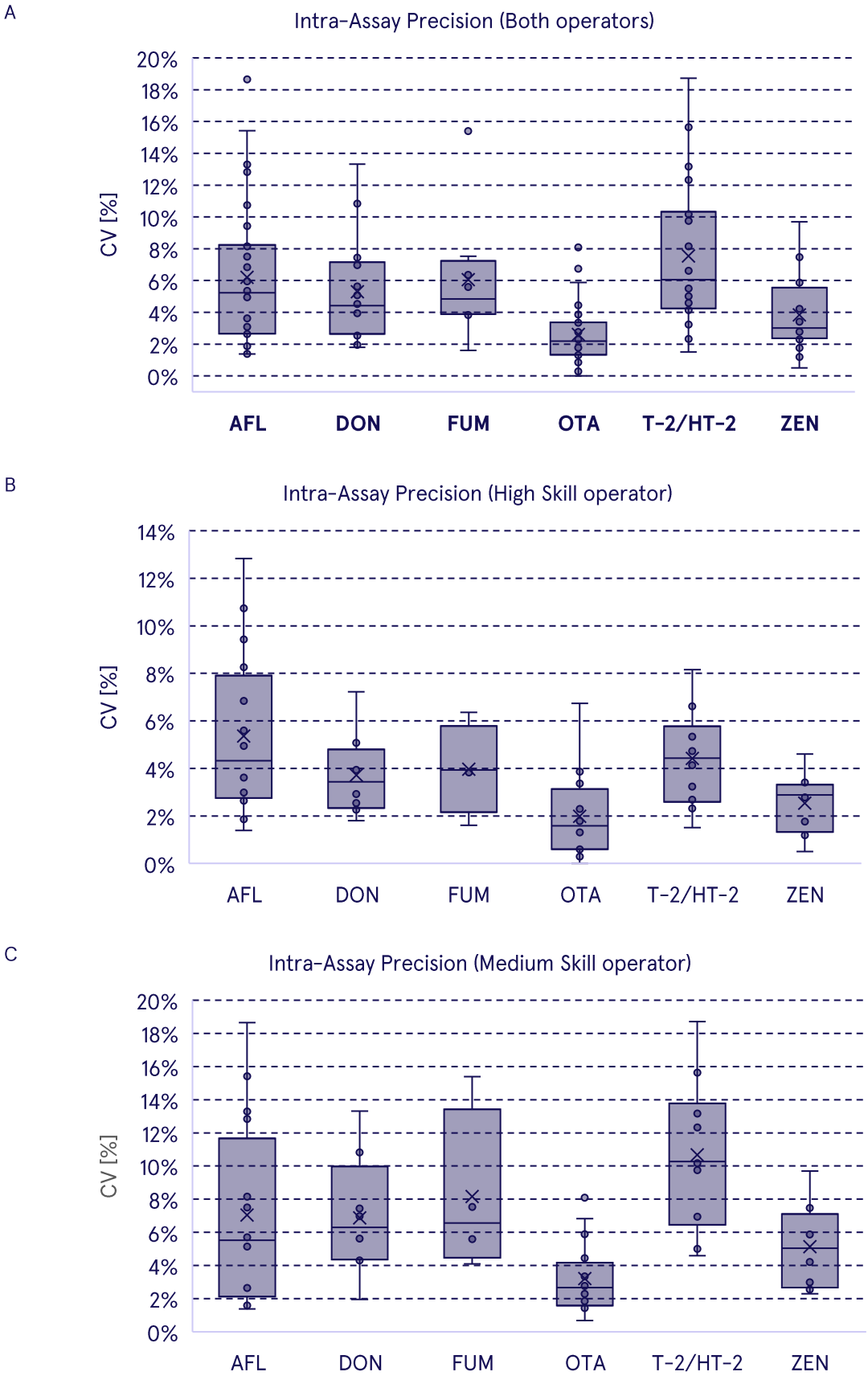


Figure 4. Intra assay coefficient of variation (CV). A) All operators. B) High skill operator. C) Medium skill operator.

Table 24. Statistical analysis of the intra assay coefficient of variation (CV).

	CV	AFL	DON	FUM	OTA	T-2/HT-2	ZEN
all operators	Mean [%]	6,2	5,3	6,1	2,6	7,5	3,8
	SD [%]	4,4	3,1	3,9	2,0	4,6	2,3
	n	32	16	8	32	20	16
High skilled operator	Mean [%]	5,4	3,7	4,0	2,0	4,4	2,5
	SD [%]	3,3	1,7	1,7	1,7	2,0	1,2
	Min [%]	1,4	1,8	1,6	0,0	1,5	0,5
	MAX [%]	12,8	7,2	6,4	6,7	8,2	4,6
	n	16	8	4	16	10	8
Medium skilled operator	Mean [%]	7,0	6,9	8,2	3,2	10,7	5,1
	SD [%]	5,2	3,5	4,4	2,0	4,3	2,4
	Min [%]	1,4	1,9	4,1	0,7	4,6	2,3
	MAX [%]	18,6	13,3	15,4	8,1	18,7	9,7
	n	16	8	4	16	10	8

Across all operators, mean intra-assay CVs ranged from 2.6% (OTA) to 7.5% (T-2/HT-2). The medium-skill operator group showed slightly higher mean CVs than the high-skill group, but the differences were small, indicating robust repeatability and limited dependence on operator skill under routine conditions.

Intermediate precision (between-day)

Intermediate precision was assessed by repeating the procedure on different measurement days. The comparison of between-day variability for the operator groups is shown in Figure 5. Overall, the observed between-day variation remained within the fitness-for-purpose criterion for intermediate precision, defined as RSD at or below 20%. For most analytes, the medium-skill operator group showed slightly higher variability than the high-skill group, which is consistent with normal operator-to-operator variation and does not indicate a systematic limitation of the method.

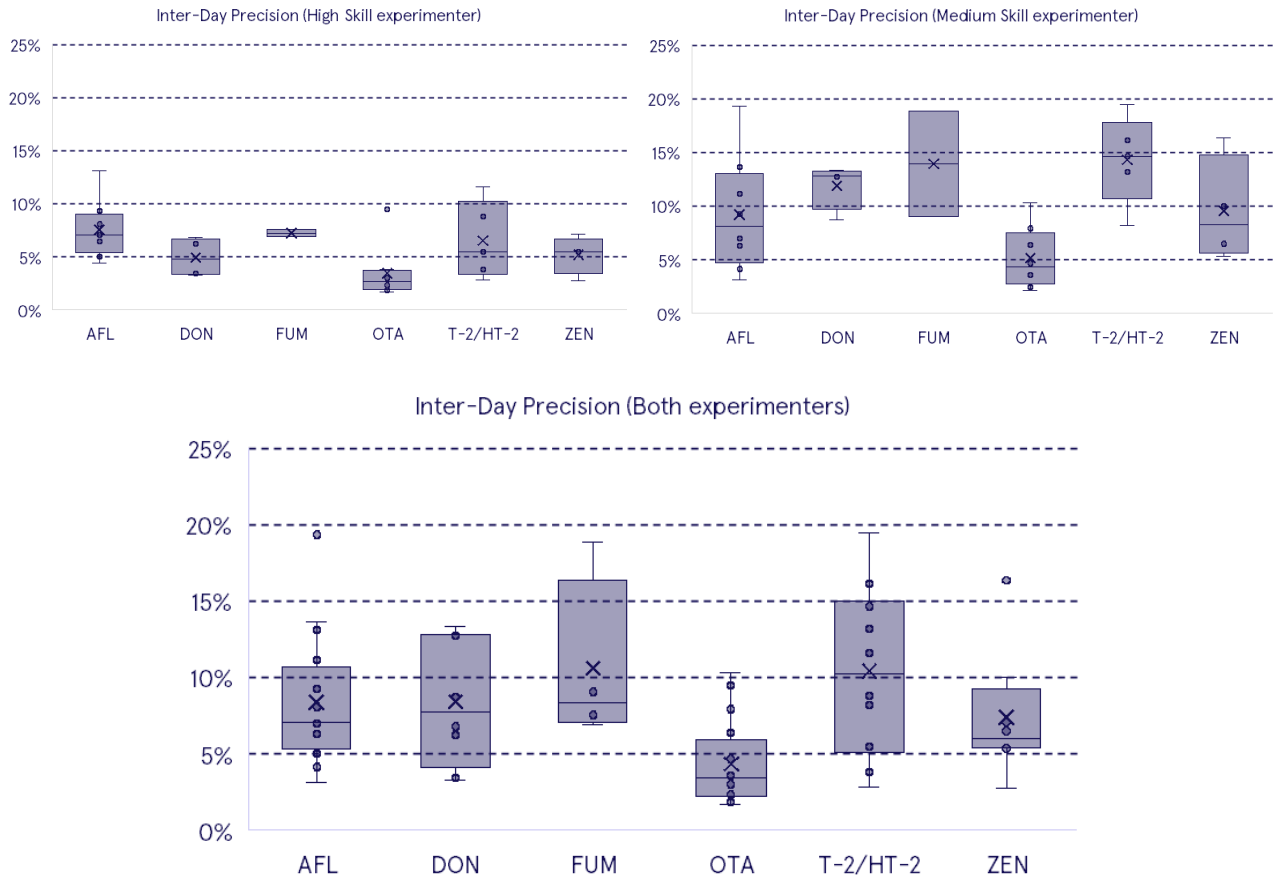


Figure 5. Comparison of inter-day precision for different operators.

Within-laboratory reproducibility (inter-analysis precision)

Within-laboratory reproducibility was evaluated across all analyses performed during the study by calculating inter-analysis CVs from all positive results, meaning results above LOQ. The mean CV of the complete analytical procedure ranged from 8.0% (OTA) to 13.8% (FUM) and remained below 20% for all analytes, meeting the predefined acceptance criterion. Figure 6 and Table 25 summarize the distribution of inter-analysis CVs.

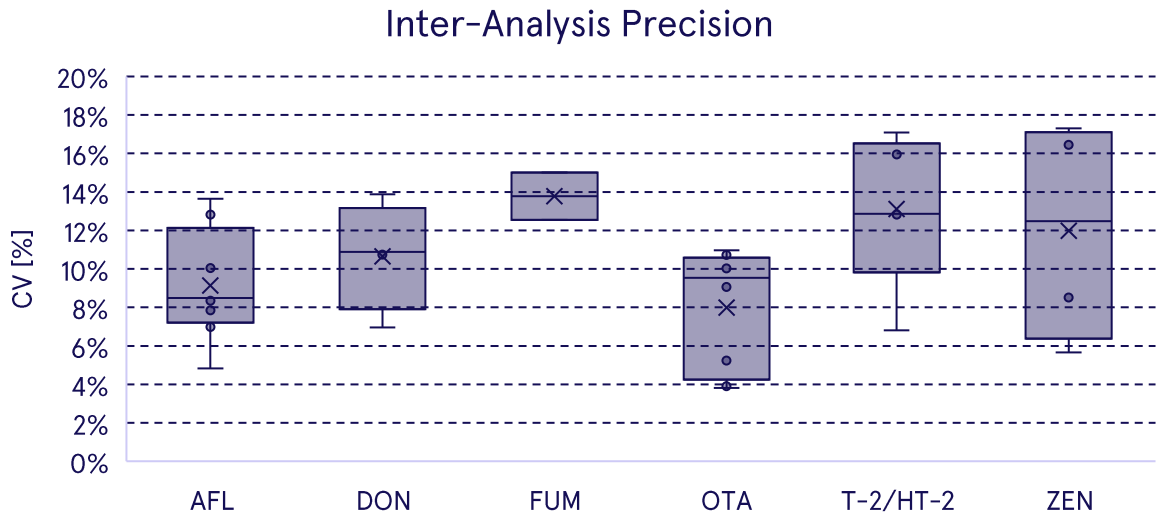


Figure 6. Inter analysis coefficient of variation (CV).

Table 25. Statistical analysis of the inter analysis coefficient of variation (CV).

CV	AFL	DON	FUM	OTA	T-2/HT-2	ZEN
Mean [%]	9.1	10.6	13.8	8.0	13.1	12.0
SD [%]	2.8	2.5	1.2	2.9	3.6	5.0
Min [%]	4.8	7.0	12.6	3.8	6.8	5.7
MAX [%]	13.7	13.9	15.0	11.0	17.1	17.3
n	8	4	2	8	5	4

Precision limits (repeatability limit and within-laboratory limit)

To provide operational precision limits, the relative repeatability limit and the relative within-laboratory reproducibility limit were calculated using the ISO 5725 approximation. The relative repeatability limit is defined as

$$r_{rel}\,[\%] = 2.8\,RSD_r[\%]$$ *Equation 20,*

where RSD_r is the intra-assay CV. The relative within-laboratory limit is defined as

$$R_{rel}[\%] = 2.8\,RSD_{IP}[\%]$$ *Equation 21,*

where RSD_{IP} is the inter-analysis CV representing within-laboratory reproducibility.

These limits are provided as indicative, practical guidance on the expected maximum difference between two results under repeatability and within-laboratory conditions; final acceptance criteria and decision limits should be defined by the implementing laboratory according to its scope and ISO/IEC 17025 requirements.

Table 26 repeatability limit (r_{rel}) and within-laboratory limit (R_{rel}) for each parameter based on mean CVs.

Parameter	r_{rel} from intra-assay CV [%]	R_{rel} from inter analysis CV [%]
AFL	17.4	25.5
DON	14.8	29.7
FUM	17.1	38.6
OTA	7.3	22.4
T-2/HT-2	21.0	36.7
ZEN	10.6	33.6

Measurement uncertainty

Trueness and method-wide bias contribution

Trueness of the SAFIA assay was assessed using both QC-material measurements with assigned values and spiking experiments across the working range. Detailed recoveries and their variability are reported in the dedicated sections and in Table 15 and Table 23 are therefore not repeated here. To provide a general uncertainty estimate applicable to unknown routine samples, the sample-specific bias term was replaced by a

method-wide bias component derived from the trueness dataset. For each dataset, the relative bias was calculated as

$$b_{rel} = \frac{\bar{X} - X_{ref}}{X_{ref}}$$ *Equation 22,*

using the assigned value for QC materials and the nominal spike level for spiking experiments. A "typical" bias component was defined as the root-mean-square value

$$b_{RMS} = \sqrt{\frac{1}{n}\sum_{n}^{i=1} b_{rel}^2}$$ *Equation 23,*

which quantifies the typical magnitude of bias without cancellation of positive and negative deviations. In addition, a "conservative" bias component was defined as an upper quantile of the $|b_{\mathrm{rel}}|$ distribution (e.g., 90th percentile) to reflect less favourable matrices.

Contributing of method-wide precision

Intermediate precision was estimated from run-to-run variability under within-laboratory conditions. For each eligible dataset (LLOQ spiking experiments and QC materials), triplicate determinations were first averaged per run to obtain run means. Intermediate precision for each dataset was then calculated as the relative standard deviation of these run means. LLOQ datasets comprised two independent runs (each analyzed in triplicate), whereas QC materials comprised four independent runs (each analyzed in triplicate). ULOQ datasets were not included in the intermediate precision estimate because they were measured in a single run only and therefore reflect repeatability rather than intermediate precision. A method-wide precision component was derived by aggregating the individual relative standard deviation values across all included datasets within the confirmed working range, using

$$RSD_{RMS} = \sqrt{\frac{1}{n}\sum_{n}^{i=1} RSD_{rel}^2}$$ *Equation 24,*

as the RSD of a "typical" estimate and an upper quantile (90th percentile) as a conservative estimate. The following values were obtained.

Table 27. Method-wide relative bias and intermediate precision components used for typical (RMS) and conservative (90th percentile) uncertainty estimates across all analytes.

Uncertainty Parameter	Typical/ conservative	AFL	DON	FUM	OTA	T-2/HT-2	ZEN
U_{bias}	b_{RMS}	22%	18%	15%	23%	20%	13%
	$\lvert b_{rel} \rvert$ (90 %)	34%	25%	25%	36%	34%	21%
U_{prec}	RSD_{RMS}	9%	14%	12%	9%	15%	12%
	RSD (90 %)	13%	17%	16%	12%	19%	17%

In this dataset, the uncertainty budget is dominated by the bias component.
To provide a practical, method-wide estimate of measurement uncertainty for routine samples, a "typical" and a "conservative" relative uncertainty were derived for each analyte by combining a method-wide bias

component and a method-wide intermediate precision component. The typical relative standard uncertainty was defined as

$$U_{typ} = \sqrt{u_{bias,RMS}^2 + u_{prec,RMS}^2}, \qquad \textit{Equation 25,}$$

where $u_{bias,RMS}$ is the RMS of the relative bias distribution and $u_{prec,RMS}$ is the RMS of the relative intermediate precision (RSD_{RMS}) across all eligible datasets within the confirmed working range. The corresponding typical expanded relative uncertainty at approximately 95% coverage was calculated as

$$MU95_{typ} = k \cdot MU_{typ},\ with\ k = 2 \qquad \textit{Equation 26.}$$

In addition, a conservative relative standard uncertainty was calculated as

$$U_{cons} = \sqrt{u_{bias,90}^2 + u_{prec,90}^2} \qquad \textit{Equation 27,}$$

using the 90th percentile of $|\ b_{rel}\ |$ and the 90th percentile of RSD to reflect less favourable matrices. The corresponding conservative expanded relative **uncertainty** was then calculated accordingly

$$MU95_{cons} = k \cdot MU_{cons}, with\ k = 2 \qquad \textit{Equation 28.}$$

The resulting analyte-specific values are summarized in Table 28.

Table 28. Method-wide relative measurement uncertainty for each analyte, reported as typical (RMS-based) and conservative (90th percentile-based) standard uncertainty (U_{typ}, U_{cons}) and expanded uncertainty at 95% coverage ($MU95_{typ}$, $MU95_{cons}$)

Uncertainty Parameter	AFL	DON	FUM	OTA	T-2/HT-2	ZEN
U_{typ}	24%	23%	20%	24%	25%	18%
$MU95_{typ}$	**48%**	**46%**	**39%**	**49%**	**51%**	**35%**
U_{cons}	37%	30%	30%	38%	39%	27%
$MU95_{cons}$	**73%**	**61%**	**60%**	**76%**	**77%**	**54%**

The typical expanded uncertainties $MU95_{typ}$ ranged from 35% (ZEN) to 51% (T-2/HT-2), while the conservative expanded uncertainties $MU95_{cons}$ ranged from 54% (ZEN) to 77% (T-2/HT-2). These estimates are consistent with the distribution of individually calculated uncertainties reported in Table 15 and visualized as boxplot in Figure 7. In particular, the box plot of individual MU95 values shows that $MU95_{typ}$ lies close to the central tendency of the dataset, whereas $MU95_{cons}$ approximates the upper part of the distribution and therefore provides a robust, fit-for-purpose uncertainty estimate for unknown routine samples across diverse matrices.

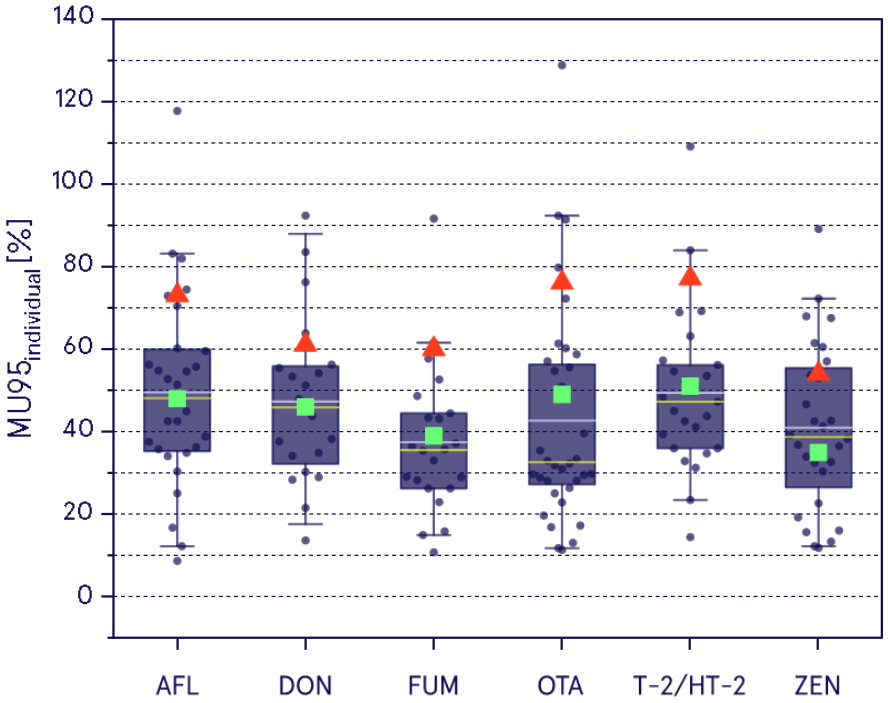


Figure 7. Distribution of analyte-specific expanded relative measurement uncertainties (MU95, k = 2) across all individual datasets within the confirmed working range; boxes show the median and interquartile range, whiskers indicate the spread 5-95%, dots indicate the individual data points and outlier, the yellow line indicates the mean and the pale blue line the median values. The typical (green filled squares) and conservative (red filled triangle) method-wide MU95 estimates are shown for comparison.

Overall, the agreement between the method-wide estimates and the empirically observed spread of individual MU95values supports that the proposed approach yields plausible and generally applicable uncertainty figures for the SAFIA assay within the confirmed working range.

Ruggedness

Ruggedness of the SAFIA assay was evaluated using a full factorial 2^4 design of experiments with four method parameters: ethanol concentration in the extraction solvent (60% v/v and 80% v/v), extraction time (10 min and 20 min), sample buffer dilution (1:9 vs 1:11), and assay incubation time (−5 min and +5 min relative to the 20 min incubation used for calibration). The QC material RM-48 was used for this study and showed normal quantification within the reported accuracy under non-varied conditions. Sixteen factor combinations were tested and each combination was repeated twice; the complete design plan is provided in the Annex. Duplicate measurements per experimental run were averaged prior to statistics, resulting in n = 32 observations per analyte. Data were evaluated in jamovi using a fixed-effects ANOVA including all main effects and all two-way interactions, with recovery (%) as the primary response. The results are summarized in Table 29.

ΔRecovery is defined as the difference between the mean recoveries at the high and low setting of the respective factors, averaged over all remaining factors. For Control, the absolute concentration was used instead of Recovery, as for Control no recovery is defined. Main effects are listed in Table 29; statistically significant effects are indicated by $p < 0.05$. All other main effects and two-way interactions were not significant.

Table 29. Ruggedness DoE (24) summary: statistically significant effects on recovery (%) and estimated effect sizes by analyte (fixed-effects ANOVA; RM-48).

Analyte	Significant factor(s)	ΔRecovery (High−Low) [%]	p-value
AFL	Incubation time	-50.1	<0.001
	c(EtOH)	-5.1	0.573
DON	Incubation time	-41.1	<0.001
	c(EtOH)	-24	0.002
FUM	Incubation time	-48.5	<0.001
	c(EtOH)	-57	<0.001
OTA	Incubation time	-100.7	<0.001
	c(EtOH)	-42	0.004
T-2/HT-2	Incubation time	-89.9	<0.001
	c(EtOH)	-11	0.225
ZEN	Incubation time	-65.2	<0.001
	c(EtOH)	-23	<0.001
Control	Incubation time	-2.9	<0.001
	EtOH concentr.	-0.296	0.001
	c(EtOH) x Incubation time	0.56	0.002

Observed effects and critical control parameters (CCP)

The estimated effect sizes (ΔRecovery, high minus low) confirm that assay incubation time was the dominant ruggedness driver for all analytes, thus can be considered a critical control parameter (CCP). A shift from 15 min to 25 min incubation relative to the 20 min calibration caused pronounced decreases in recovery, ranging from −41% (DON) to −101% (OTA), with all effects being highly significant ($p < 0.001$). This demonstrates that even small systematic deviations in incubation timing can lead to substantial bias in quantitative results. This is readily explained by the assay kinetics: if the incubation time is shorter than in the calibration, signal intensities are reduced, which tends to shift results towards higher calculated concentrations (overestimation). Conversely, longer incubation increases signal intensities due to continued antibody binding to the beads, which tends to shift results towards lower calculated concentrations (underestimation). For practical SAFIA use, this implies that incubation times for calibrators and samples should be kept as consistent as possible, and reagent addition should be performed without systematic delays between standards and

samples. Time control should be managed as a documented process parameter. Incubation start and stop times for calibrators and samples should be recorded (e.g., via a simple time log or a worksheet template) to ensure traceability and to demonstrate compliance with the defined procedure. In practice, this can be achieved by using dedicated timers, applying a fixed pipetting sequence, and limiting the maximum time offset between the first and last well of standards and samples. If a relevant deviation in incubation timing is observed or cannot be excluded (e.g., interruptions during pipetting, operator change, or measurements performed on different days), a new calibration curve should be generated under the prevailing timing conditions, and the deviation should be documented as a potential contributor to uncertainty.

Ethanol concentration showed analyte specific effects and is therefore a second CCP. Significant negative ethanol effects were observed for DON (−24%, $p = 0.002$), FUM (−57%, $p < 0.001$), OTA (−42%, $p = 0.004$) and ZEN (−23%, $p < 0.001$), whereas no significant ethanol dependence was found for AFL (−5.1%, $p = 0.573$) or T-2/HT-2 (−11%, $p = 0.225$). For the control parameter, both incubation time (−2.9, $p < 0.001$) and ethanol concentration (−0.296, $p = 0.001$) were significant, and a significant EtOH × incubation interaction ($\Delta = +0.56$, $p = 0.002$) indicates that the impact of incubation time on the control response depends on the ethanol level. This is plausible because antibodies differ in their tolerance to organic solvents and may partially lose binding performance at elevated ethanol concentrations. Accordingly, careful preparation of the extraction solvent is recommended. Analytical grade absolute ethanol should be used and diluted with deionized water to 70% (v/v) using calibrated volumetric glassware (e.g., class A) or equivalently validated volumetric equipment.

Extraction time and sample buffer dilution did not show significant main effects, and two-way interactions were generally not relevant; only the control parameter showed a significant interaction between ethanol concentration and incubation time. This indicates that extraction is rapid for the tested grain-type matrices and that small deviations in the preparation of the sample buffer dilution do not materially affect method performance within the investigated range.

Assumption checks indicated that residuals were consistent with normality for aflatoxins, DON, and T-2/HT-2, whereas control, FUM, OTA, and ZEN showed deviations and most datasets indicated variance heterogeneity. Overall, the method can be considered rugged with respect to extraction time and sample buffer dilution within the tested ranges. In contrast, assay incubation time and, for several analytes, ethanol concentration should be treated as critical process parameters requiring tighter procedural control.

Annex

Plate Layouts

The Samples DLA-A-2025 and DLA-B are sample from a round robin exercise and were not part of this study, but analyzed within the experiments.

Plate Name/Analysis ID: MKS215

#1	1	2	3	4	5	6	7
A	Kal-1	Kal-1	RM-15	RM-15	RM-15	RM-36	RM-36
B	Kal-2	Kal-2	RM-18	RM-18	RM-18	RM-46	RM-46
C	Kal-3	Kal-3	RM-22	RM-22	RM-22	DLA-A 2025	DLA-A 2025
D	Kal-4	Kal-4	RM-23	RM-23	RM-23	DLA-B 2025	DLA-B 2025
E	Kal-5	Kal-5	RM-24	RM-24	RM-24	RM-36	blank
F	Kal-6	Kal-6	RM-32	RM-32	RM-32	RM-46	blank
G	Kal-7	Kal-7	RM-33	RM-33	RM-33	DLA-A 2025	blank
H	Kal-8	Kal-8	RM-34	RM-34	RM-34	DLA-B 2025	blank

Plate Name/Analysis ID: MKS216

#2	1	2	3	4	5	6	7
A	Kal-1	Kal-1	RM-15	RM-15	RM-15	blank	RM-36
B	Kal-2	Kal-2	RM-18	RM-18	RM-18	blank	RM-46
C	Kal-3	Kal-3	RM-22	RM-22	RM-22	blank	DLA-A 2025
D	Kal-4	Kal-4	RM-23	RM-23	RM-23	blank	DLA-B 2025
E	Kal-5	Kal-5	RM-24	RM-24	RM-24	RM-36	RM-36
F	Kal-6	Kal-6	RM-32	RM-32	RM-32	RM-46	RM-46
G	Kal-7	Kal-7	RM-33	RM-33	RM-33	DLA-A 2025	DLA-A 2025
H	Kal-8	Kal-8	RM-34	RM-34	RM-34	DLA-B 2025	DLA-B 2025

.

Plate Name/Analysis ID: MKS217

#3	6	7	8	9	10	11	12
A	Kal-1	Kal-1	RM-36	RM-36	RM-15	RM-15	RM-15
B	Kal-2	Kal-2	RM-46	RM-46	RM-18	RM-18	RM-18
C	Kal-3	Kal-3	DLA-A 2025	DLA-A 2025	RM-22	RM-22	RM-22
D	Kal-4	Kal-4	DLA-B 2025	DLA-B 2025	RM-23	RM-23	RM-23
E	Kal-5	Kal-5	RM-36	blank	RM-24	RM-24	RM-24
F	Kal-6	Kal-6	RM-46	blank	RM-32	RM-32	RM-32
G	Kal-7	Kal-7	DLA-A 2025	blank	RM-33	RM-33	RM-33
H	Kal-8	Kal-8	DLA-B 2025	blank	RM-34	RM-34	RM-34

Plate Name/Analysis ID: MKS219

#4	1	2	3	4	5	6
A	RM-46	RM-34	RM-34	RM-34	Kal-1	Kal-1
B	RM-46	RM-33	RM-33	RM-33	Kal-2	Kal-2
C	RM-46	RM-32	RM-32	RM-32	Kal-3	Kal-3
D	RM-36	RM-24	RM-24	RM-24	Kal-4	Kal-4
E	RM-36	RM-23	RM-23	RM-23	Kal-5	Kal-5
F	RM-36	RM-22	RM-22	RM-22	Kal-6	Kal-6
G	blank	RM-18	RM-18	RM-18	Kal-7	Kal-7
H	blank	RM-15	RM-15	RM-15	Kal-8	Kal-8

Plate Name/Analysis ID: MKS220

#5	1	2	3	4	5	6	7
A	Kal-1	Kal-1	LAB-1	LAB-1	LAB-1	M-101	M-101
B	Kal-2	Kal-2	LAB-2	LAB-2	LAB-2	M-104	M-104
C	Kal-3	Kal-3	LAB-13	LAB-13	LAB-13	M-105	M-105
D	Kal-4	Kal-4	LAB-14	LAB-14	LAB-14	M-137	M-137
E	Kal-5	Kal-5	LAB-17	LAB-17	LAB-17	M-101	blank
F	Kal-6	Kal-6	LAB-18	LAB-18	LAB-18	M-104	blank
G	Kal-7	Kal-7	M-43	M-43	M-43	M-105	blank
H	Kal-8	Kal-8	M-53	M-53	M-53	M-137	blank

Plate Name/Analysis ID: MKS221

#6	1	2	3	4	5	6	7
A	LAB-1	LAB-1	LAB-1	M-101	M-101	Kal-1	Kal-1
B	LAB-2	LAB-2	LAB-2	M-104	M-104	Kal-2	Kal-2
C	LAB-13	LAB-13	LAB-13	M-105	M-105	Kal-3	Kal-3
D	LAB-14	LAB-14	LAB-14	M-137	M-137	Kal-4	Kal-4
E	LAB-17	LAB-17	LAB-17	M-101	blank	Kal-5	Kal-5
F	LAB-18	LAB-18	LAB-18	M-104	blank	Kal-6	Kal-6
G	M-43	M-43	M-43	M-105	blank	Kal-7	Kal-7
H	M-53	M-53	M-53	M-137	blank	Kal-8	Kal-8

Plate Name/Analysis ID: MKS222

#3	1	2	3	4	5	6	7
A	Kal-1	Kal-1	M-101	M-101	LAB-1	LAB-1	LAB-1
B	Kal-2	Kal-2	M-104	M-104	LAB-2	LAB-2	LAB-2
C	Kal-3	Kal-3	M-105	M-105	LAB-13	LAB-13	LAB-13
D	Kal-4	Kal-4	M-137	M-137	LAB-14	LAB-14	LAB-14
E	Kal-5	Kal-5	M-101	blank	LAB-17	LAB-17	LAB-17
F	Kal-6	Kal-6	M-104	blank	LAB-18	LAB-18	LAB-18
G	Kal-7	Kal-7	M-105	blank	M-43	M-43	M-43
H	Kal-8	Kal-8	M-137	blank	M-53	M-53	M-53

Plate Name/Analysis ID: MKS223

#4	1	2	3	4	5	6	7
A	M-101	M-101	LAB-1	LAB-1	LAB-1	Kal-1	Kal-1
B	M-104	M-104	LAB-2	LAB-2	LAB-2	Kal-2	Kal-2
C	M-105	M-105	LAB-13	LAB-13	LAB-13	Kal-3	Kal-3
D	M-137	M-137	LAB-14	LAB-14	LAB-14	Kal-4	Kal-4
E	M-101	blank	LAB-17	LAB-17	LAB-17	Kal-5	Kal-5
F	M-104	blank	LAB-18	LAB-18	LAB-18	Kal-6	Kal-6
G	M-105	blank	M-43	M-43	M-43	Kal-7	Kal-7
H	M-137	blank	M-53	M-53	M-53	Kal-8	Kal-8

Plate Name/Analysis ID: MKS224

Note: The samples reported here are spiked at LLOQ-Level

	1	2	3	4	5	6	7
A	Kal-1	Kal-1	M-104	LAB-1	LAB-1	LAB-1	M-104
B	Kal-2	Kal-2	M-105	LAB-2	LAB-2	LAB-2	M-105
C	Kal-3	Kal-3	M-137	LAB-13	LAB-13	LAB-13	M-137
D	Kal-4	Kal-4	Spiked Blank	LAB-14	LAB-14	LAB-14	M-104
E	Kal-5	Kal-5	Spiked Blank	LAB-17	LAB-17	LAB-17	M-105
F	Kal-6	Kal-6	Spiked Blank	LAB-18	LAB-18	LAB-18	M-137
G	Kal-7	Kal-7	Spiked Blank	M-43	M-43	M-43	blank
H	Kal-8	Kal-8	Spiked Blank	M-53	M-53	M-53	blank

Plate Name/Analysis ID: MKS225

Note: The samples reported here are spiked at LLOQ-Level

	1	2	3	4	5	6	7
A	Kal-1	Kal-1	M-104	LAB-1	LAB-1	LAB-1	M-104
B	Kal-2	Kal-2	M-105	LAB-2	LAB-2	LAB-2	M-105
C	Kal-3	Kal-3	M-137	LAB-13	LAB-13	LAB-13	M-137
D	Kal-4	Kal-4	Spiked Blank	LAB-14	LAB-14	LAB-14	M-104
E	Kal-5	Kal-5	Spiked Blank	LAB-17	LAB-17	LAB-17	M-105
F	Kal-6	Kal-6	Spiked Blank	LAB-18	LAB-18	LAB-18	M-137
G	Kal-7	Kal-7	Spiked Blank	M-43	M-43	M-43	blank
H	Kal-8	Kal-8	Spiked Blank	M-53	M-53	M-53	blank

Plate Name/Analysis ID: MKS229

Note: The samples reported here are spiked at ULOQ-Level 1

	1	2	3	4	5
A	Kal-1	Kal-1	LAB-1	LAB-1	LAB-1
B	Kal-2	Kal-2	LAB-14	LAB-14	LAB-14
C	Kal-3	Kal-3	LAB-17	LAB-17	LAB-17
D	Kal-4	Kal-4	LAB-18	LAB-18	LAB-18
E	Kal-5	Kal-5	M-43	M-43	M-43
F	Kal-6	Kal-6	M-104	M-104	M-104
G	Kal-7	Kal-7	M-105	M-105	M-105
H	Kal-8	Kal-8	Spiked Blank	Spiked Blank	Spiked Blank

Plate Name/Analysis ID: MKS231

Note: The samples reported here are spiked at ULOQ-Level 2

	1	2	3	4	5
A	Kal-1	Kal-1	LAB-1	LAB-1	LAB-1
B	Kal-2	Kal-2	LAB-14	LAB-14	LAB-14
C	Kal-3	Kal-3	LAB-17	LAB-17	LAB-17
D	Kal-4	Kal-4	LAB-18	LAB-18	LAB-18
E	Kal-5	Kal-5	M-43	M-43	M-43
F	Kal-6	Kal-6	M-104	M-104	M-104
G	Kal-7	Kal-7	M-105	M-105	M-105
H	Kal-8	Kal-8	Spiked Blank	Spiked Blank	Spiked Blank

Plate Name/Analysis ID: MKS228

Note: These are the experimental runs from the 2^4 DoE. Sample RM-48

	Incubation time 20 min		Incubation time 15 min		Incubation time 25 min		Incubation time 15 min		Incubation time 25 min	
Row	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
A	KAL-1	KAL-1	1	2	5	6	1	2	5	6
B	KAL-2	KAL-2	3	4	7	8	3	4	7	8
C	KAL-3	KAL-3	9	10	13	14	9	10	13	14
D	KAL-4	KAL-4	11	12	15	16	11	12	15	16
E	KAL-5	KAL-5	17	18	21	22	17	18	21	22
F	KAL-6	KAL-6	19	20	23	24	19	20	23	24
G	KAL-7	KAL-7	25	26	29	30	25	26	29	30
H	KAL-8	KAL-8	27	28	31	32	27	28	31	32

Results of individual samples and One-Way ANOVA

Results for Sum of Aflatoxins (AFL)

Table 30. Summarized results of all spiked samples for AFL.

Sample	Result ± SD AFL [µg/kg]	RSD [%]	Blank corrected result [µg/kg]	Ref. conc. [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	1.22 ± 0.133	11	1.11	0.900	124	0.21	19	43
M-105	1.09 ± 0.118	11	0.97	0.900	107	0.07	7	39
M-137	1.35 ± 0.173	13	1.03	0.900	114	0.13	13	55
M-43	1.11 ± 0.0900	8	0.90	0.900	100	0.00	0	45
M-53	1.40 ± 0.125	9	0.86	0.900	95	-0.04	-5	51
LAB-1	1.02 ± 0.0582	6	0.80	0.900	89	-0.10	-13	60
LAB-13	1.02 ± 0.0501	5	0.67	0.900	75	-0.23	-34	82
LAB-14	1.07 ± 0.0657	6	0.93	0.900	103	0.03	3	36
LAB-17	1.10 ± 0.0470	4	0.78	0.900	87	-0.12	-15	56
LAB-18	0.794 ± 0.0607	8	0.76	0.900	85	-0.14	-18	30
LAB-2	0.972 ± 0.0525	5	0.77	0.900	85	-0.13	-17	55
Spiked Blank	1.19 ± 0.0897	8	1.04	0.900	116	0.14	14	38
M-43	15.6 ± 0.735	5	-	15.0	104	0.6	4	12
M-104	11.6 ± 0.532	5	-	15.0	77	-3.4	-29	60
M-105	17.4 ± 2.81	16	-	15.0	116	2.4	14	43
LAB-1	11.1 ± 1.51	14	-	15.0	74	-3.9	-35	74
LAB-14	12.1 ± 1.30	11	-	15.0	81	-2.9	-24	53
LAB-17	14.4 ± 0.257	2	-	15.0	96	-0.6	-4	9
LAB-18	13.0 ± 1.23	9	-	15.0	87	-2.0	-15	36
Spiked Blank	15.8 ± 2.57	16	-	15.0	105	0.8	5	34
M-43	13.0 ± 2.32	18	-	20.0	65	-7.0	-54	114

Sample	Result ± SD AFL [µg/kg]	RSD [%]	Blank corrected result [µg/kg]	Ref. conc. [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	16.3 ± 2.41	15	-	20.0	81	-3.7	-23	54
M-105	13.8 ± 1.09	8	-	20.0	69	-6.2	-45	91
LAB-1	14.6 ± 0.853	6	-	20.0	73	-5.4	-37	75
LAB-14	13.1 ± 2.82	22	-	20.0	65	-6.9	-53	114
LAB-17	14.6 ± 2.18	15	-	20.0	73	-5.4	-37	79
LAB-18	13.8 ± 2.11	15	-	20.0	69	-6.2	-45	96
Spiked Blank	15.6 ± 0.802	5	-	20.0	78	-4.4	-29	58
RM-18	5.28 ± 0.455	8.6	-	4.72	112	0.56	11	25
RM-22	8.74 ± 0.877	10.0	-	8.47	103	0.27	3	17
RM-24	4.87 ± 0.340	7.0	-	7.72	63	-2.85	-58	118
RM-32	6.98 ± 0.953	13.7	-	9.36	75	-2.38	-34	73
RM-33	10.2 ± 0.850	8.4	-	8.63	118	1.54	15	56
RM-34	2.75 ± 0.352	12.8	-	2.31	119	0.44	16	35
RM-36	6.16 ± 0.483	7.9	-	8.56	72	-2.40	-39	83
RM-46	5.44 ± 0.263	4.8	-	7.20	76	-1.76	-32	71

AFL Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95)

One-Way ANOVA (Welch's)

	F	df1	df2	p
Measurement Uncertainty	5.97	3	14.7	0.007

Group Descriptives

	Spike Level	N	Mean	SD	SE
Measurement Uncertainty	QC-Mat	8	59.8	33.6	11.88
	0.9	12	49.2	13.9	4.02
	15	8	40.1	22.5	7.94

Group Descriptives

	Spike Level	N	Mean	SD	SE
	20	8	85.1	22.9	8.09

Assumption Checks

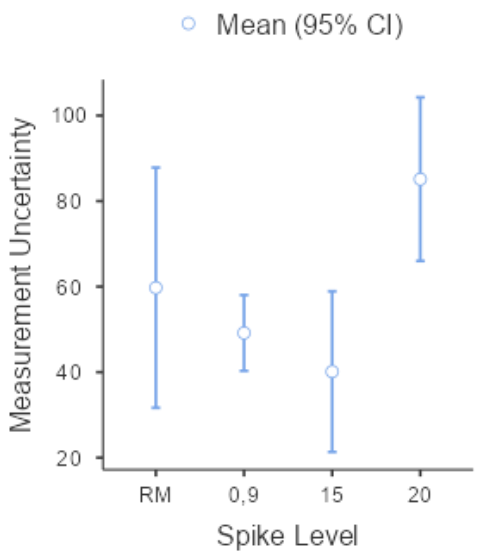
Normality Test (Shapiro-Wilk)

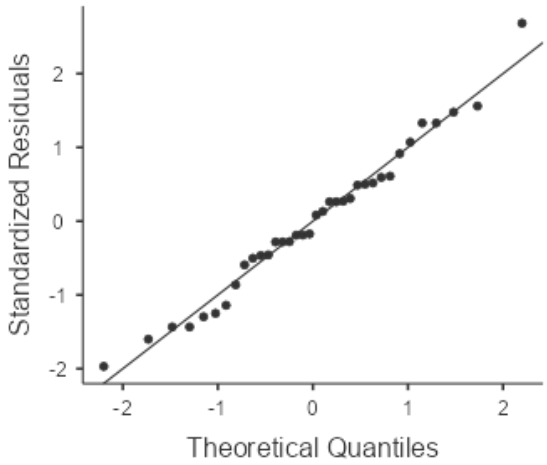
	W	p
Measurement Uncertainty	0.982	0.803

Note. A low p-value suggests a violation of the assumption of normality

Plots

Measurement Uncertainty





Post Hoc Tests

Games-Howell Post-Hoc Test – Measurement Uncertainty

		RM	0,9	15	20
RM	Mean difference	—	10.6	19.63	-25.4
	p-value	—	0.832	0.537	0.334
0,9	Mean difference		—	9.04	-36.0*
	p-value		—	0.744	0.011
15	Mean difference			—	-45.0**
	p-value			—	0.007
20	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA

One-Way ANOVA (Welch's)

	F	df1	df2	p
Recovery	12.8	3	15.5	<.001

Assumption Checks

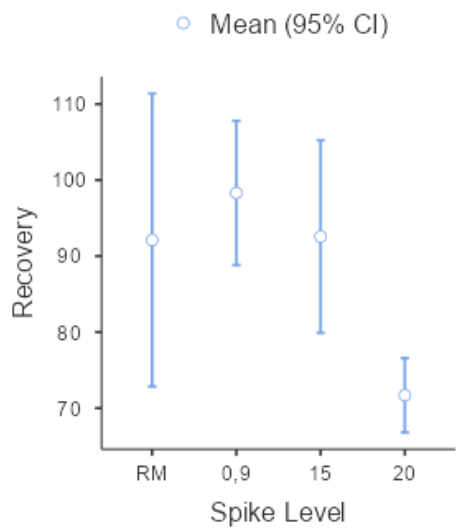
Normality Test (Shapiro-Wilk)

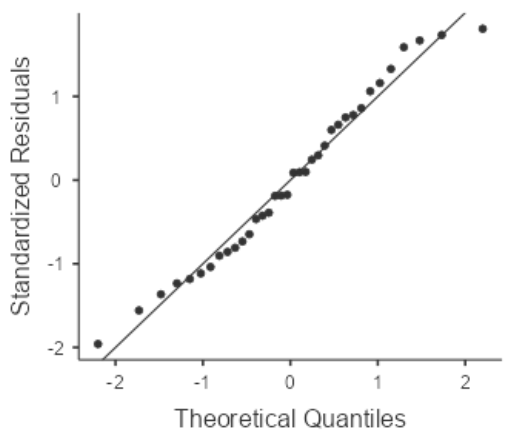
	W	p
Recovery	0.973	0.515

Note. A low p-value suggests a violation of the assumption of normality

Plots

Recovery





Post Hoc Tests

Games-Howell Post-Hoc Test – Recovery

		RM	0,9	15	20
RM	Mean difference	—	-6.20	-0.463	20.4
	p-value	—	0.905	1.000	0.149
0,9	Mean difference		—	5.733	26.6***
	p-value		—	0.838	<.001
15	Mean difference			—	20.9*
	p-value			—	0.023
20	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA

One-Way ANOVA (Welch's)

	F	df1	df2	p
Precision	1.83	3	15.1	0.185

Assumption Checks

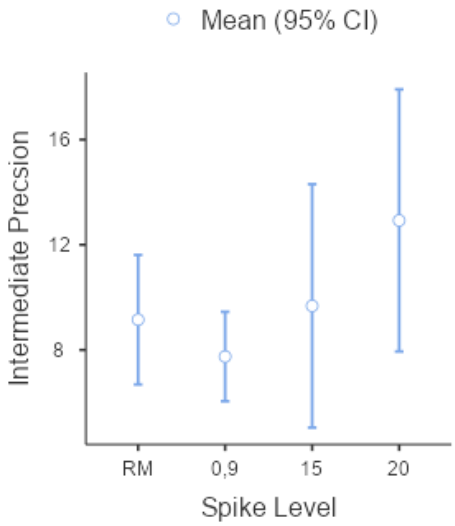
Normality Test (Shapiro-Wilk)

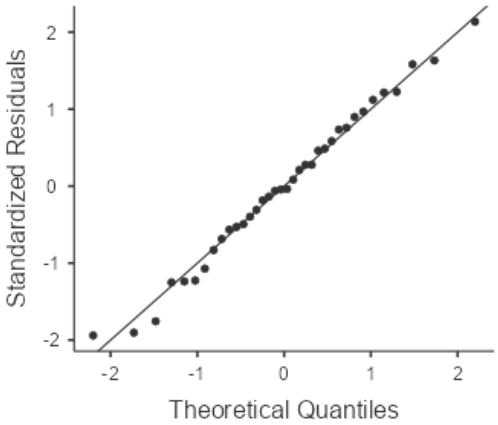
	W	p
Precision	0.986	0.924

Note. A low p-value suggests a violation of the assumption of normality

Plots

Precision





Post Hoc Tests

Games-Howell Post-Hoc Test – Intermediate Precision

		RM	0,9	15	**20**
RM	Mean difference	—	1.40	-0.525	-3.78
	p-value	—	0.707	0.995	0.417
0,9	Mean difference		—	-1.925	-5.18
	p-value		—	0.798	0.168
15	Mean difference			—	-3.25
	p-value			—	0.678
20	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

Results for Deoxynivalenol (DON)

Table 31. Summarized results of all spiked samples for DON.

Sample	Result ± SD DON [µg/kg]	RSD [%]	Blank corrected result [µg/kg]	Ref. conc. [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	74.0 ± 19.5	26	59.44	50	119	9.44	16	76
M-105	55.1 ± 9.47	17	37.13	50	74	-12.9	-35	92
M-137	74.1 ± 11.4	15	46.51	50	93	-3.49	-8	55
M-43	57.3 ± 6.45	11	42.27	50	85	-7.73	-18	53
M-53	63.4 ± 8.18	13	38.40	50	77	-11.6	-30	84
LAB-1	51.3 ± 6.08	12	42.22	50	84	-7.78	-18	56
LAB-14	66.6 ± 9.48	14	42.07	50	84	-7.93	-19	64
LAB-17*	41.6 ± 7.76	19	33.40	50	67	-16.6	-50	119
LAB-18*	10.6 ± 5.93	56	10.60	50	21	-39.4	-372	752
LAB-2	62.5 ± 6.49	10	42.08	50	84	-7.92	-19	54
Spiked Blank	50.3 ± 5.22	10	45.69	50	91	-4.31	-9	35
M-43	1.355 ± 17.7	1.3	-	1500	90	-145	-11	22
M-104	1.236 ± 59.2	4.8	-	1500	82	-264	-21	44
M-105	1.269 ± 75.6	6.0	-	1500	85	-231	-18	38
LAB-1	1.222 ± 91.3	7.5	-	1500	81	-278	-23	48
LAB-14	1.318 ± 57.8	4.4	-	1500	88	-182	-14	29
LAB-17	1.202 ± 78.1	6.5	-	1500	80	-298	-25	51
LAB-18*	849 ± 62.4	7.4	-	1500	57	-651	-77	154
Spiked Blank	1.315 ± 21.5	1.6	-	1500	88	-185	-14	28
M-43	1.387 ± 119	8.6	-	2000	69	-613	-44	90
M-104	1.310 ± 75.8	5.8	-	2000	66	-690	-53	106
M-105	1.213 ± 96.2	7.9	-	2000	61	-787	-65	131

Sample	Result ± SD DON [µg/kg]	RSD [%]	Blank corrected result [µg/kg]	Ref. conc. [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
LAB-1	1.292 ± 135	10.5	-	2000	65	-708	-55	112
LAB-14	1.079 ± 74.0	6.9	-	2000	54	-921	-85	171
LAB-17	1.238 ± 73.4	5.9	-	2000	62	-762	-62	124
LAB-18*	870 ± 59.3	6.8	-	2000	44	-1130	-130	260
Spiked Blank	1.277 ± 44.4	3.5	-	2000	64	-723	-57	114
RM-18	1.190 ± 131	11.0	-	1385	86	-195	-16	38
RM-23	771 ± 53.7	7.0	-	769	100	2	0	14
RM-33	878 ± 122	13.9	-	796	110	82	9	34
RM-34	1.386 ± 149	10.7	-	1240	112	146	11	30

Results marked with * were excluded due to possible matrix interference, see chapter Limits of Quantification and working range

DON Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95)

One-Way ANOVA (Welch's)

	F	df1	df2	p
Recovery DON	33.5	3	9.52	<.001

Assumption Checks

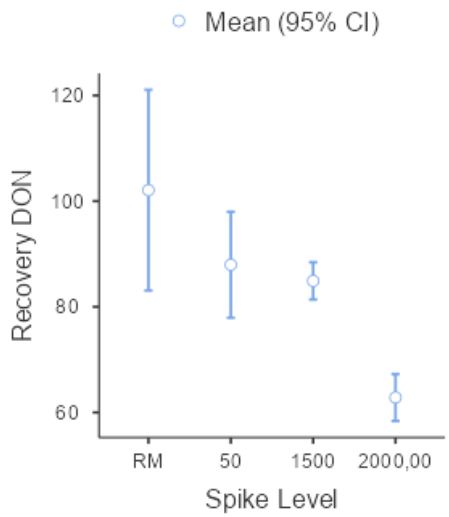
Normality Test (Shapiro-Wilk)

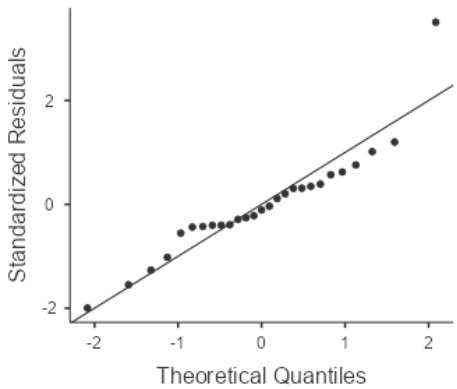
	W	p
Recovery DON	0.894	0.010

Note. A low p-value suggests a violation of the assumption of normality

Plots

Recovery DON





Post Hoc Tests

Games-Howell Post-Hoc Test – Recovery DON

		QC-Mat	50	1500	2000,00
QC-Mat	Mean difference	—	14.1	17.17	39.2*
	p-value	—	0.311	0.172	0.015
50	Mean difference		—	3.06	25.1**
	p-value		—	0.907	0.001

Games-Howell Post-Hoc Test – Recovery DON

		QC-Mat	50	1500	2000,00
1500	Mean difference			—	22.1***
	p-value			—	<.001
2000	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
Precision DON	9.72	3	10.5	0.002

Assumption Checks

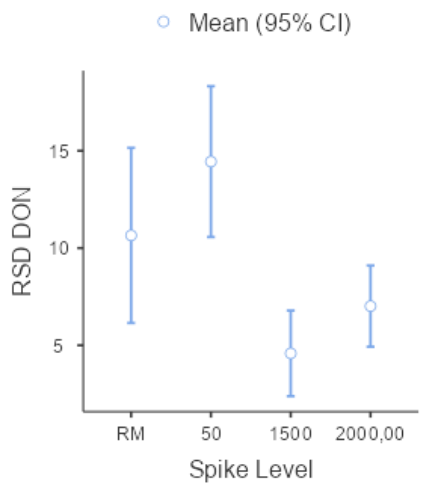
Normality Test (Shapiro-Wilk)

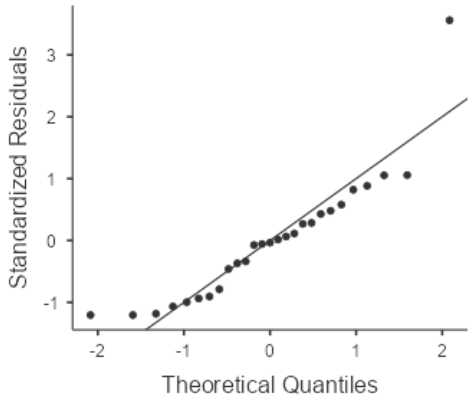
	W	**p**
Precision DON	0.864	0.002

Note. A low p-value suggests a violation of the assumption of normality

Plots

Precision DON





Post Hoc Tests

Games-Howell Post-Hoc Test – RSD DON

		QC-Mat	50	1500	2000,00
QC-Mat	Mean difference	—	-3.79	6.06*	3.64
	p-value	—	0.360	0.048	0.237
50	Mean difference		—	9.86**	7.43**
	p-value		—	0.001	0.010
1500	Mean difference			—	-2.43
	p-value			—	0.256
2000,00	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
MU95 DON	24.6	3	11.2	<.001

Assumption Checks

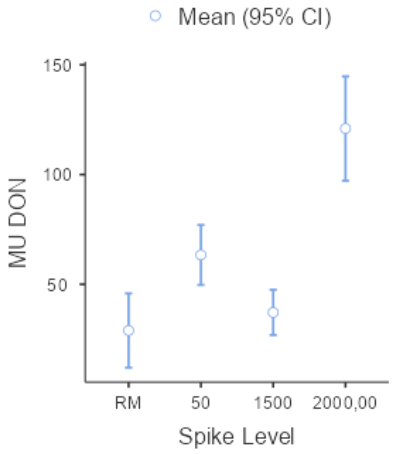
Normality Test (Shapiro-Wilk)

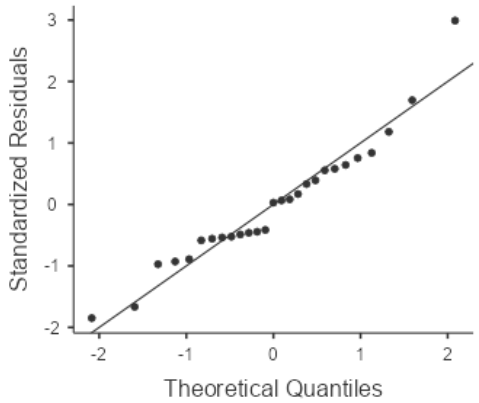
	W	p
MU DON	0.948	0.190

Note. A low p-value suggests a violation of the assumption of normality

Plots

MU95 DON





Post Hoc Tests

Games-Howell Post-Hoc Test – MU95 DON

		QC-Mat	50	1500	2000,00
QC-Mat	Mean difference	—	-34.4**	-8.22	-92.0***
	p-value	—	0.008	0.640	<.001
50	Mean difference		—	26.19*	-57.6**
	p-value		—	0.014	0.002
1500	Mean difference			—	-83.8***
	p-value			—	<.001
2000,00	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

Results for Sum of Fumonisins (FUM)

Table 32. Summarized results of all spiked samples for FUM.

Sample	Result ± SD FUM [µg/kg]	RSD [%]	Blank corrected result [µg/kg]	Ref. conc. [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	70.3 ± 15.6	22	45.3	50	91	-4,70	-10	92
M-105	44.1 ± 6.28	14	43.0	50	86	-7,05	-16	45
M-137	54.5 ± 9.09	17	51.8	50	104	1,83	4	36
M-43	40.4 ± 4.10	10	39.4	50	79	-10,6	-27	58
M-53	56.1 ± 6.22	11	54.1	50	108	4,13	8	29
LAB-1	71.6 ± 2.69	4	54.7	50	109	4,66	9	23
LAB-13	46.0 ± 5.12	11	44.4	50	89	-5,62	-13	35
LAB-14	48.3 ± 4.07	8	44.3	50	89	-5,68	-13	33
LAB-17	52.7 ± 4.13	8	45.9	50	92	-4,13	-9	29
LAB-18	43.1 ± 4.24	10	41.2	50	82	-8,79	-21	49
LAB-2	83.6 ± 4.34	5	54.9	50	110	4,90	9	26
Spiked Blank	56.4 ± 5.06	9	56.1	50	112	6,12	11	28
M-43	536 ± 22.8	4.2	-	700	77	-164	-31	62
M-104	584 ± 50.5	8.7	-	700	83	-116	-20	43
M-105	591 ± 67.2	11.4	-	700	84	-109	-18	43
LAB-1	688 ± 35.2	5.1	-	700	98	-12,0	-2	11
LAB-14	557 ± 34.5	6.2	-	700	80	-143	-26	53
LAB-17	657 ± 23.9	3.6	-	700	94	-43,3	-7	15
LAB-18	597 ± 30.5	5.1	-	700	85	-103	-17	36
Spiked Blank	744 ± 39.6	5.3	-	700	106	43,9	6	16
M-43	319 ± 15.7	4.9	-	1000	32	-681	-214	428
M-104	329 ± 17.6	5.4	-	1000	33	-671	-204	408
M-105	287 ± 16.0	5.5	-	1000	29	-713	-248	496

Sample	Result ± SD FUM [µg/kg]	RSD [%]	Blank corrected result [µg/kg]	Ref. conc. [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
LAB-1	402 ± 26.8	6.7	-	1000	40	-598	-149	297
LAB-14	263 ± 18.5	7.0	-	1000	26	-737	-280	560
LAB-17	357 ± 8.25	2.3	-	1000	36	-643	-180	360
LAB-18	356 ± 20.8	5.8	-	1000	36	-644	-181	362
Spiked Blank	391 ± 11.9	3.1	-	1000	39	-609	-156	311
RM-18	674 ± 84.5	12.6	-	640	105	33,5	5	37
RM-34	527 ± 79.1	15.0	-	467	113	60,0	11	26

FUM Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95)

One-Way ANOVA (Welch's)

	F	df1	df2	p
Recovery FUM	152	3	5.15	<.001

Assumption Checks

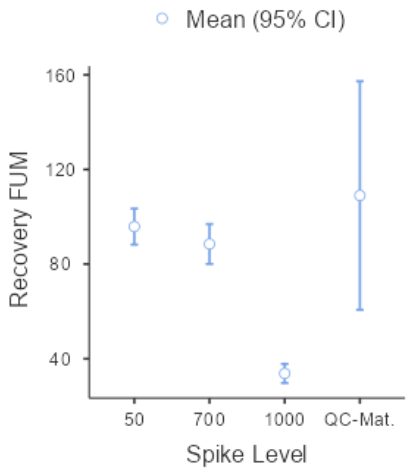
Normality Test (Shapiro-Wilk)

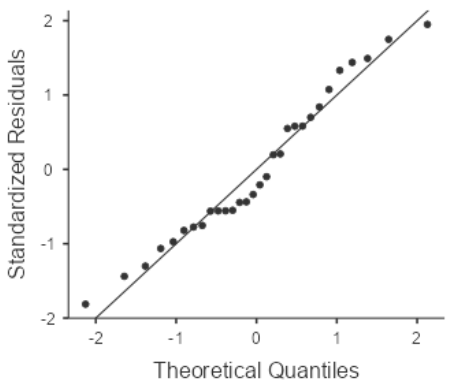
	W	p
Recovery FUM	0.963	0.362

Note. A low p-value suggests a violation of the assumption of normality

Plots

Recovery FUM





Post Hoc Tests

Games-Howell Post-Hoc Test – Recovery FUM

		50	700	1000	QC-Mat.
50	Mean difference	—	7.36	62.0***	-13.2
	p-value	—	0.469	<.001	0.218
700	Mean difference		—	54.7***	-20.5
	p-value		—	<.001	0.078
1000	Mean difference			—	-75.2
	p-value			—	n. a.†
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

† p values not available (n.a.) because at least one group contains insufficient variance or too few observations for the test statistics.

One-Way ANOVA (Welch's)

	F	df1	df2	p
Precision FUM	13.5	3	5.28	0.007

Assumption Checks

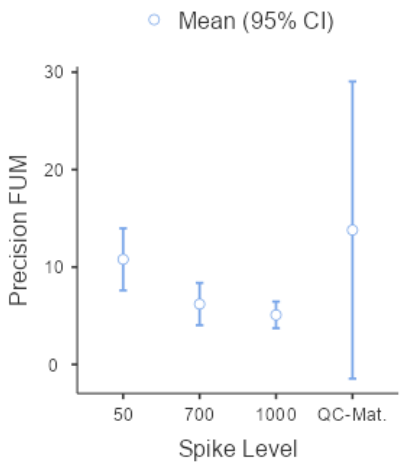
Normality Test (Shapiro-Wilk)

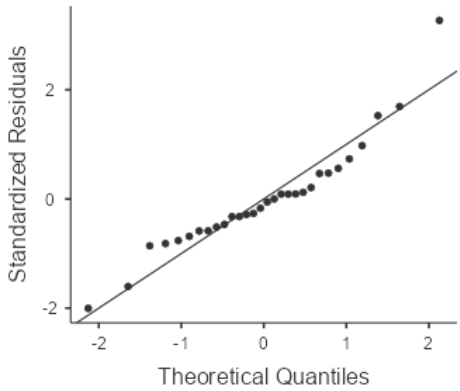
	W	p
Precision FUM	0.917	0.022

Note. A low p-value suggests a violation of the assumption of normality

Plots

Precision FUM





Post Hoc Tests

Games-Howell Post-Hoc Test – Precision FUM

		50	700	1000	QC-Mat.
50	Mean difference	—	4.59	5.70*	-3.01
	p-value	—	0.068	0.012	0.454
700	Mean difference		—	1.11	-7.60
	p-value		—	0.739	0.068
1000	Mean difference			—	-8.71
	p-value			—	n. a.
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

† p values not available (n.a.) because at least one group contains insufficient variance or too few observations for the test statistics.

One-Way ANOVA (Welch's)

	F	df1	df2	p
MU95 FUM	37.6	3	7.72	<.001

Assumption Checks

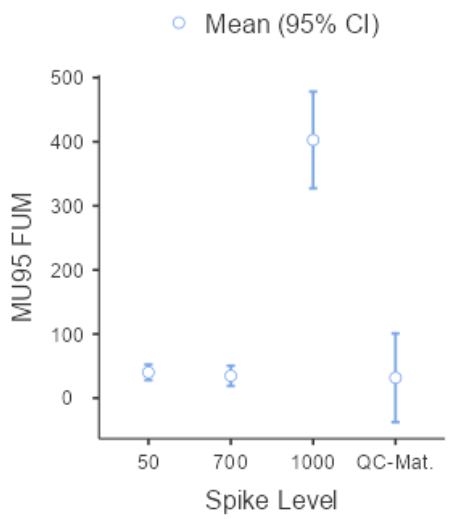
Normality Test (Shapiro-Wilk)

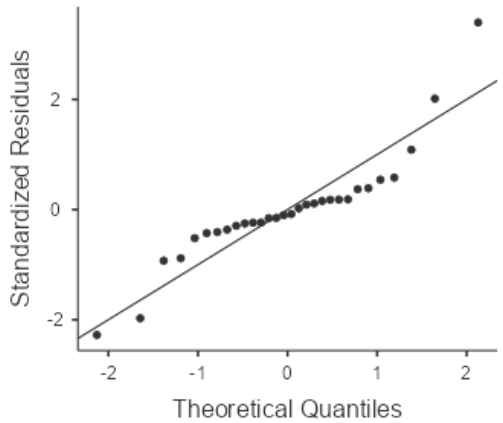
	W	**p**
MU95 FUM	0.858	<.001

Note. A low p-value suggests a violation of the assumption of normality

Plots

MU95 FUM





Post Hoc Tests

Games-Howell Post-Hoc Test – MU95 FUM

		50	700	1000	QC-Mat.
50	Mean difference	—	5.48	-362***	8.52
	p-value	—	0.920	<.001	0.710
700	Mean difference		—	-368***	3.05
	p-value		—	<.001	0.983
1000	Mean difference			—	370.98***
	p-value			—	<.001
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

Results for Ochratoxin A (OTA)

Table 33. Summarized results of all spiked samples for OTA.

Sample	Result ± SD OTA [µg/kg]	RSD [%]	Ref. conc. [µg/kg]	Blank corrected result [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	0.976 ± 0.112	11	1.2	0.94	78	-0,258	-26	59
M-105	1.09 ± 0.136	12	1.2	0.98	82	-0,221	-20	55
M-137	1.56 ± 0.293	19	1.2	1.36	113	0,156	10	48
M-43	1.07 ± 0.0925	9	1.2	1.02	85	-0,181	-17	40
M-53	1.60 ± 0.142	9	1.2	1.32	110	0,115	7	32
LAB-1	1.90 ± 0.103	5	1.2	1.21	101	0,012	1	23
LAB-13	1.05 ± 0.0634	6	1.2	0.96	80	-0,241	-23	51
LAB-14	1.45 ± 0.127	9	1.2	1.07	89	-0,131	-9	32
LAB-17	1.33 ± 0.0637	4.8	1.2	1.13	94	-0,074	-6	33
LAB-18	0.982 ± 0.109	11.1	1.2	0.98	82	-0,218	-22	50
LAB-2	2.17 ± 0.128	6	1.2	1.26	105	0,0648	3	20
Spiked Blank	1.46 ± 0.133	9	1.2	1.35	112	0,147	10	33
M-43	52.4 ± 3.74	7.1	50	-	105	2,36	5	17
M-104	47.9 ± 2.30	4.8	50	-	96	-2,14	-4	13
M-105	57.5 ± 1.30	2.3	50	-	115	7,47	13	26
LAB-1	78.1 ± 2.54	3.3	50	-	156	28,1	36	72
LAB-14	50.4 ± 2.85	5.7	50	-	101	0,38	1	11
LAB-17	69.1 ± 1.58	2.3	50	-	138	19,1	28	56
LAB-18	56.7 ± 5.00	8.8	50	-	113	6,7	12	29
Spiked Blank	79.1 ± 12.1	15.3	50	-	158	29,1	37	80
M-43	38.7 ± 2.92	7.6	50	-	77	-11,3	-29	60
M-104	39.0 ± 1.49	3.8	50	-	78	-11,0	-28	57
M-105	38.3 ± 1.45	3.8	50	-	77	-11,7	-30	61

Sample	Result ± SD OTA [µg/kg]	RSD [%]	Ref. conc. [µg/kg]	Blank corrected result [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
LAB-1	54.3 ± 1.89	3.5	50	-	109	4,29	8	17
LAB-14	30.5 ± 2.27	7.4	50	-	61	-19,5	-64	129
LAB-17	45.9 ± 5.47	11.9	50	-	92	-4,1	-9	30
LAB-18	45.0 ± 4.81	10.7	50	-	90	-5,04	-11	31
Spiked Blank	47.3 ± 5.29	11.2	50	-	95	-2,68	-6	25
RM-15	1.06 ± 0.114	10.7	1.53	-	69	-0,469	-44	92
RM-18	2.76 ± 0.281	10.2	2.52	-	110	0,245	9	28
RM-22	4.65 ± 0.182	3.9	5.21	-	89	-0,565	-12	29
RM-24	9.93 ± 0.379	3.8	9.72	-	102	0,206	2	12
RM-32	6.80 ± 0.746	11.0	7.64	-	89	-0,839	-12	36
RM-34	3.04 ± 0.275	9.1	2.72	-	112	0,319	11	30
RM-36	5.75 ± 0.301	5.2	8.25	-	70	-2,498	-43	92
RM-46	2.60 ± 0.260	10.0	2.51	-	103	0,0859	3	28

OTA Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95)

One-Way ANOVA (Welch's)

	F	df1	df2	p
Recovery OTA	4.32	3	15.8	0.021

Assumption Checks

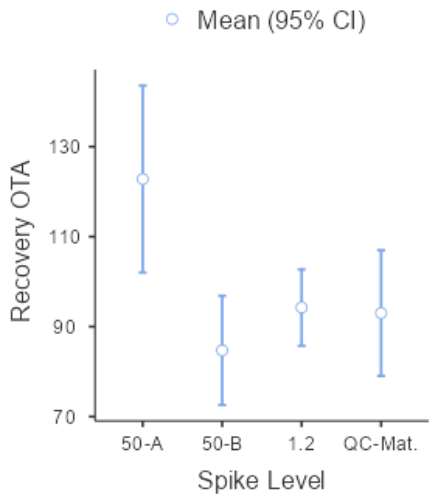
Normality Test (Shapiro-Wilk)

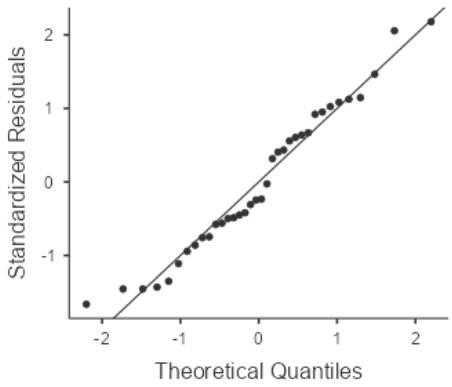
	W	p
Recovery OTA	0.964	0.284

Note. A low p-value suggests a violation of the assumption of normality

Plots

Recovery OTA





Post Hoc Tests

Games-Howell Post-Hoc Test – Recovery OTA

		50-A	50-B	1.2	QC-Mat.
50-A	Mean difference	—	38.1*	28.56	29.79
	p-value	—	0.014	0.058	0.064
50-B	Mean difference		—	-9.52	-8.29
	p-value		—	0.473	0.719
1.2	Mean difference			—	1.23
	p-value			—	0.998
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
Precision OTA	0.879	3	16.8	0.472

Assumption Checks

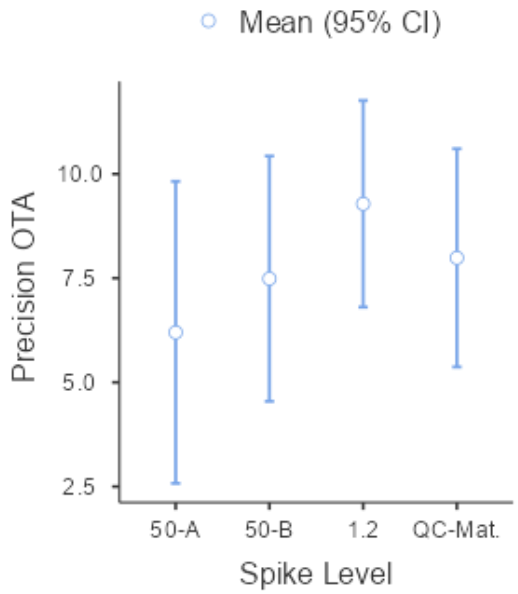
Normality Test (Shapiro-Wilk)

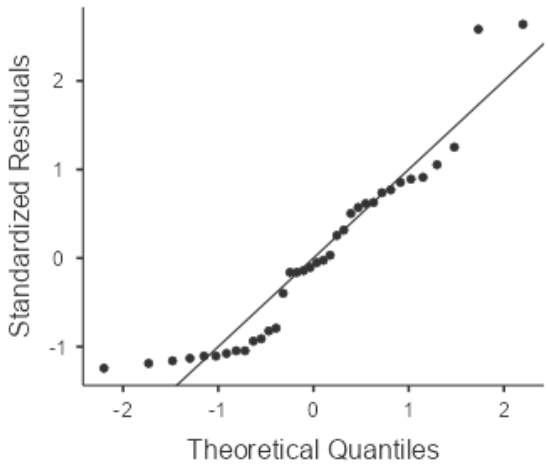
	W	p
Precision OTA	0.907	0.005

Note. A low p-value suggests a violation of the assumption of normality

Plots

Precision OTA





Post Hoc Tests

Games-Howell Post-Hoc Test – Precision OTA

		50-A	50-B	1.2	QC-Mat.
50-A	Mean difference	—	-1.29	-3.08	-1.787
	p-value	—	0.913	0.399	0.781
50-B	Mean difference		—	-1.80	-0.500
	p-value		—	0.712	0.990
1.2	Mean difference			—	1.296
	p-value			—	0.844
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
MU95 OTA	0.280	3	13.7	0.839

Assumption Checks

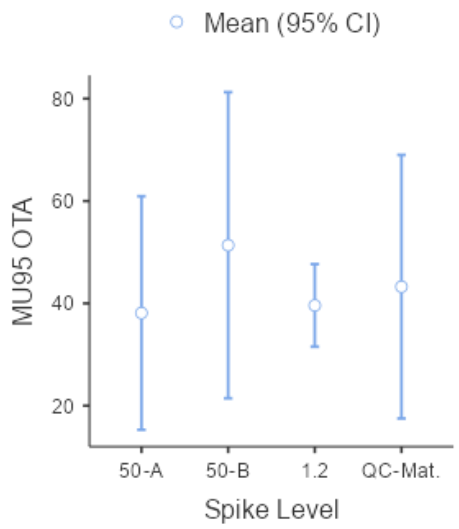
Normality Test (Shapiro-Wilk)

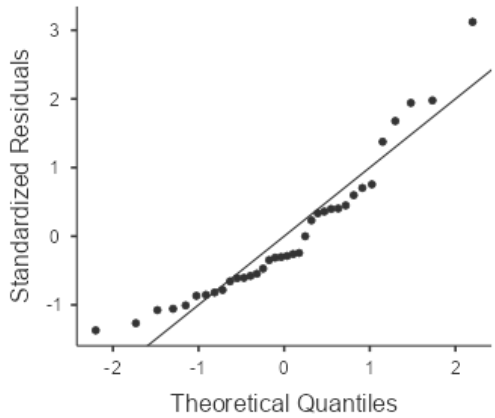
	W	p
MU95 OTA	0.902	0.004

Note. A low p-value suggests a violation of the assumption of normality

Plots

MU95 OTA





Post Hoc Tests

Games-Howell Post-Hoc Test – MU95 OTA

		50-A	50-B	1.2	QC-Mat.
50-A	Mean difference	—	-13.2	-1.48	-5.14
	p-value	—	0.838	0.999	0.984
50-B	Mean difference		—	11.76	8.10
	p-value		—	0.809	0.961
1.2	Mean difference			—	-3.66
	p-value			—	0.988
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

Results for sum of T-2 and HT-2 toxins (T-2/HT-2)

Table 34. Summarized results of all spiked samples for T-2/HT-2

Sample	Result ± SD T-2/HT-2 [µg/kg]	RSD [%]	Ref. conc. [µg/kg]	Blank corrected result [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	13.2 ± 2.52	19	11	11.1	101	0.0863	4	54
M-105	12.3 ± 2.23	18	11	10.1	92	-0.850	-39	55
M-137	14.0 ± 1.44	10	11	11.5	104	0.457	18	35
M-43	18.3 ± 2.18	12	11	11.1	101	0.130	2	57
M-53	24.3 ± 3.50	14	11	17.0	154	5.96	81	84
LAB-1	10.3 ± 1.28	12	11	9.40	85	-1.60	-180	47
LAB-13	19.4 ± 2.66	14	11	13.3	121	2.32	38	63
LAB-14	12.3 ± 1.55	13	11	11.4	104	0.398	43	31
LAB-17	11.6 ± 1.64	14	11	10.6	97	-0.362	-38	36
LAB-18	10.1 ± 2.16	21	11	9.52	87	-1.480	-272	56
LAB-2	11.3 ± 2.01	18	11	10.2	93	-0.799	-74	45
Spiked Blank	10.1 ± 1.91	19	11	9.55	87	-1.45	-289	52
M-43	296 ± 21.2	7.1	300	-	99	-3.50	-1	14
M-104	226 ± 25.0	11.1	300	-	75	-74.2	-33	69
M-105	253 ± 22.4	8.9	300	-	84	-47.0	-19	41
LAB-1	223 ± 7.23	3.2	300	-	74	-76.7	-34	69
LAB-14	240 ± 12.1	5.0	300	-	80	-59.7	-25	51
LAB-17	248 ± 13.9	5.6	300	-	83	-52.4	-21	44
LAB-18	196 ± 25.0	12.7	300	-	65	-104	-53	109
Spiked Blank	249 ± 13.9	5.6	300	-	83	-51.1	-21	43
M-43	251 ± 32.2	12.8	500	-	50	-249	-99	200
M-104	193 ± 15.0	7.8	500	-	39	-307	-159	319
M-105	190 ± 22.3	11.7	500	-	38	-310	-163	326

Sample	Result ± SD T-2/HT-2 [µg/kg]	RSD [%]	Ref. conc. [µg/kg]	Blank corrected result [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
LAB-1	177 ± 20.5	11.6	500	-	35	-323	-183	367
LAB-14	172 ± 21.2	12.3	500	-	34	-328	-191	382
LAB-17	218 ± 24.1	11.1	500	-	44	-282	-129	259
LAB-18	179 ± 12.9	7.2	500	-	36	-321	-180	359
Spiked Blank	198 ± 21.3	10.8	500	-	40	-302	-152	305
RM-18	112 ± 19.1	17.1	104	-	107	7.59	7	33
RM-22	178 ± 12.1	6.8	190	-	94	-12.0	-7	23
RM-23	45.7 ± 5.88	12.9	36.3	-	126	9.41	21	48
RM-33	99.2 ± 15.8	16.0	92.8	-	107	6.37	6	36
RM-34	97.5 ± 12.5	12.8	80.9	-	121	16.6	17	39

T-2/HT-2 Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95)

One-Way ANOVA (Welch's)

	F	df1	df2	p
Recovery T-2/HT-2	88.3	3	12.4	<.001

Assumption Checks

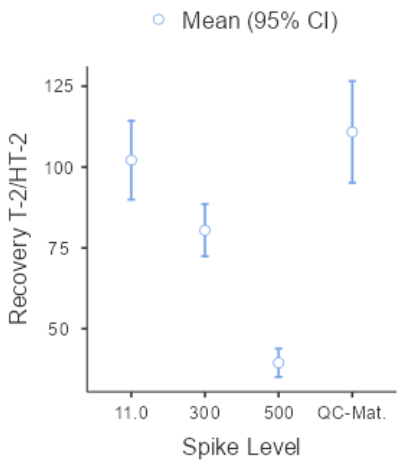
Normality Test (Shapiro-Wilk)

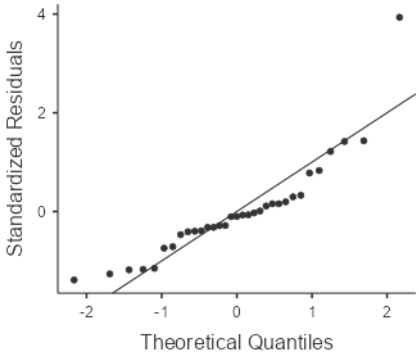
	W	p
Recovery T-2/HT-2	0.841	<.001

Note. A low p-value suggests a violation of the assumption of normality

Plots

Recovery T-2/HT-2





Post Hoc Tests

Games-Howell Post-Hoc Test – Recovery T-2/HT-2

		11.0	300	500	QC-Mat.
11.0	Mean difference	—	21.7*	62.7***	-8.73
	p-value	—	0.019	<.001	0.696
300	Mean difference		—	41.0***	-30.40*
	p-value		—	<.001	0.011
500	Mean difference			—	-71.40***
	p-value			—	<.001
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
Precision T-2/HT-2	8.90	3	12.7	0.002

Assumption Checks

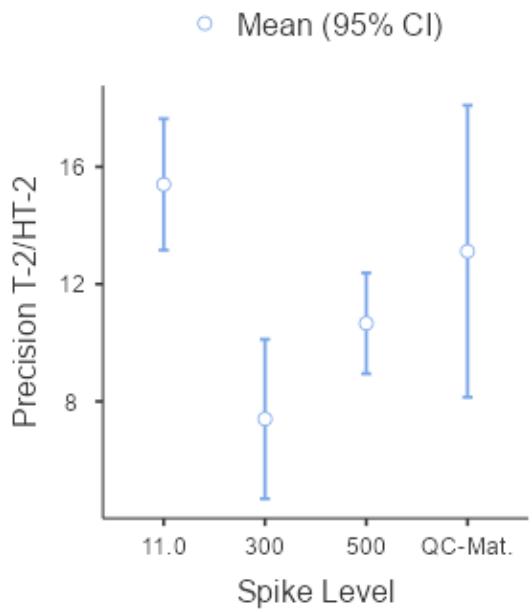
Normality Test (Shapiro-Wilk)

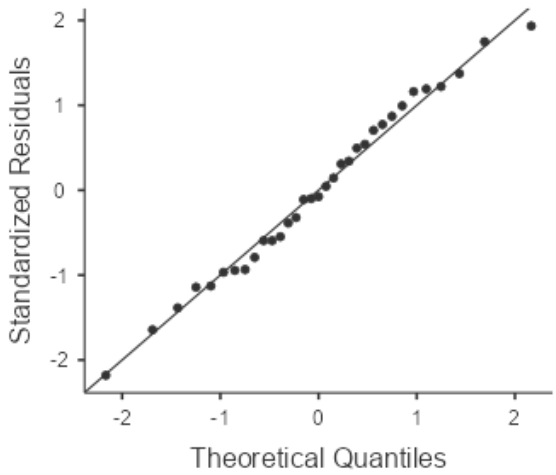
	W	p
Precision T-2/HT-2	0.986	0.939

Note. A low p-value suggests a violation of the assumption of normality

Plots

Precision T-2/HT-2





Post Hoc Tests

Games-Howell Post-Hoc Test – Precision T-2/HT-2

		11.0	300	500	QC-Mat.
11.0	Mean difference	—	8.00***	4.74**	2.28
	p-value	—	<.001	0.007	0.698
300	Mean difference		—	-3.26	-5.72
	p-value		—	0.131	0.110
500	Mean difference			—	-2.46
	p-value			—	0.614
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
MU95 T-2/HT-2	49.6	3	13.9	<.001

Assumption Checks

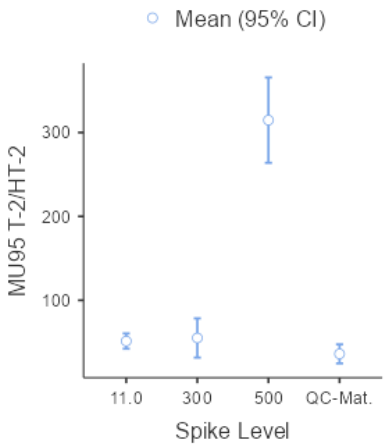
Normality Test (Shapiro-Wilk)

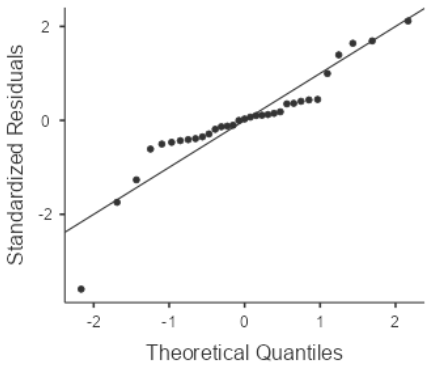
	W	p
MU95 T-2/HT-2	0.878	0.001

Note. A low p-value suggests a violation of the assumption of normality

Plots

MU95 T-2/HT-2





Post Hoc Tests

Games-Howell Post-Hoc Test – MU95 T-2/HT-2

		11.0	300	500	QC-Mat.
11.0	Mean difference	—	-3.72	-263***	15.3
	p-value	—	0.985	<.001	0.090
300	Mean difference		—	-260***	19.0
	p-value		—	<.001	0.343
500	Mean difference			—	278.7***
	p-value			—	<.001
QC-Mat.	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

Results for sum of zearalenone (ZEN)

Table 35. Summarized results of all spiked samples for ZEN

Sample	Result ± SD ZEN [µg/kg]	RSD [%]	Ref. conc. [µg/kg]	Blank corrected result [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
M-104	3.51 ± 0.571	16	3.00	3.33	111	0.327	10	43
M-105	3.33 ± 0.456	14	3.00	2.74	91	-0.264	-10	61
M-137	4.36 ± 0.635	15	3.00	3.45	115	0.445	13	68
M-43	2.99 ± 0.511	17	3.00	2.90	97	-0.103	-4	41
M-53	4.30 ± 0.409	9.5	3.00	3.74	125	0.745	20	54
LAB-1	4.43 ± 0.203	4.6	3.00	3.27	109	0.274	8	61
LAB-13	3.85 ± 0.258	7	3.00	2.65	88	-0.346	-13	72
LAB-14	4.72 ± 0.435	9	3.00	3.20	107	0.195	6	57
LAB-17*	1.07 ± 0.353	33	3.00	1.07	36	-1.93	-182	369
LAB-18*	0 ± 0	0	3.00	0.00	0	-3.00	n.d.	n.d.
LAB-2	4.69 ± 0.304	6	3.00	2.47	82	-0.533	-22	89
Spiked Blank	3.49 ± 0.434	12	3.00	2.89	96	-0.110	-4	53
M-104	297 ± 18.1	6.1	300	-	99	-2.56	-1	12
M-105	331 ± 9.31	2.8	300	-	110	30.6	9	19
M-43	312 ± 21.3	6.8	300	-	104	12.0	4	16
LAB-1	259 ± 6.56	2.5	300	-	86	-41.5	-16	32
LAB-14	286 ± 10.1	3.5	300	-	95	-13.7	-5	12
LAB-17	282 ± 4.27	1.5	300	-	94	-18.4	-7	13
LAB-18*	194 ± 3.28	1.7	300	-	0	-300	-155	109
Spiked Blank	358 ± 8.26	2.3	300	-	119	58.0	16	34
M-104	344 ± 35.9	10.4	400	-	86	-55.7	-16	38
M-105	347 ± 24.5	7.1	400	-	87	-53.4	-15	33
M-43	360 ± 29.1	8.1	400	-	90	-40.2	-11	30

Sample	Result ± SD ZEN [µg/kg]	RSD [%]	Ref. conc. [µg/kg]	Blank corrected result [µg/kg]	Recovery [%]	Bias [µg/kg]	B_{rel} [%]	$MU95_{rel}$ (k= 2) [%]
LAB-1	327 ± 12.4	3.8	400	-	82	-73.1	-22	43
LAB-14	348 ± 31.3	9.0	400	-	87	-51.7	-15	37
LAB-17	331 ± 23.7	7.2	400	-	83	-68.8	-21	68
LAB-18*	214 ± 3.16	1.5	400	-	0	-400	-187	95
Spiked Blank	391 ± 12.8	3.3	400	-	98	-8.91	-2	16
RM-18	194 ± 32.9	17.0	186	-	104	8.07	4	37
RM-23	46.3 ± 2.52	5.4	56	-	83	-9.74	-21	47
RM-33	77.0 ± 12.6	16.3	81.8	-	94	-4.82	-6	39
RM-34	123 ± 10.4	8.5	131	-	94	-8.04	-7	23

Results marked with * were excluded due to possible matrix interference, see chapter Limits of Quantification and working range.

ZEN Level Comparison by One-Way ANOVA (Recovery, Precision, and MU95)

One-Way ANOVA (Welch's)

	F	df1	df2	p
Recovery ZEN	4.67	3	10.2	0.027

Assumption Checks

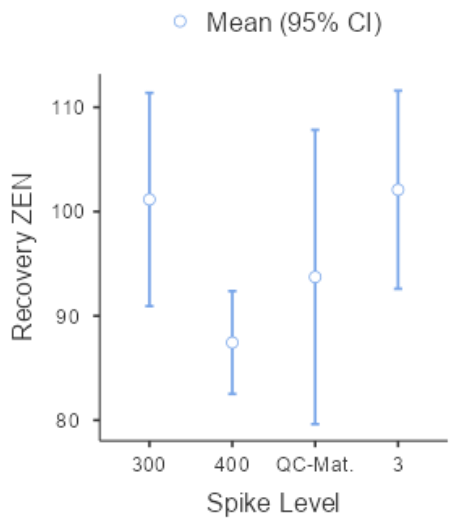
Normality Test (Shapiro-Wilk)

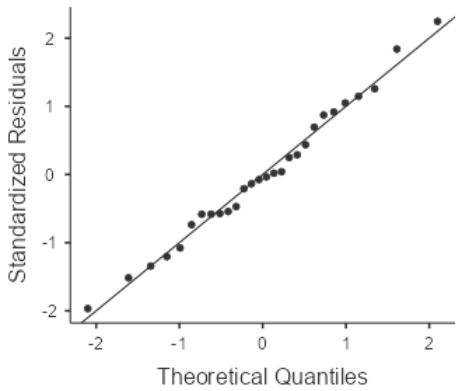
	W	p
Recovery ZEN	0.987	0.975

Note. A low p-value suggests a violation of the assumption of normality

Plots

Recovery ZEN





Post Hoc Tests

Games-Howell Post-Hoc Test – Recovery ZEN

		300	400	QC-Mat.	3
300	Mean difference	—	13.7	7.43	-0.933
	p-value	—	0.066	0.633	0.999
400	Mean difference		—	-6.28	-14.647*
	p-value		—	0.611	0.035
QC-Mat.	Mean difference			—	-8.365
	p-value			—	0.547
3	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
Precision ZEN	8.20	3	9.79	0.005

Assumption Checks

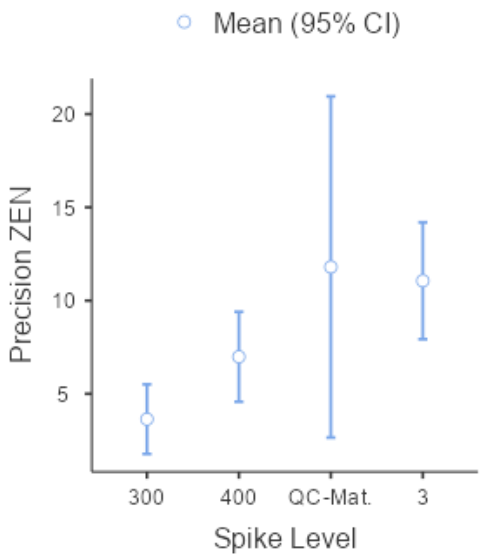
Normality Test (Shapiro-Wilk)

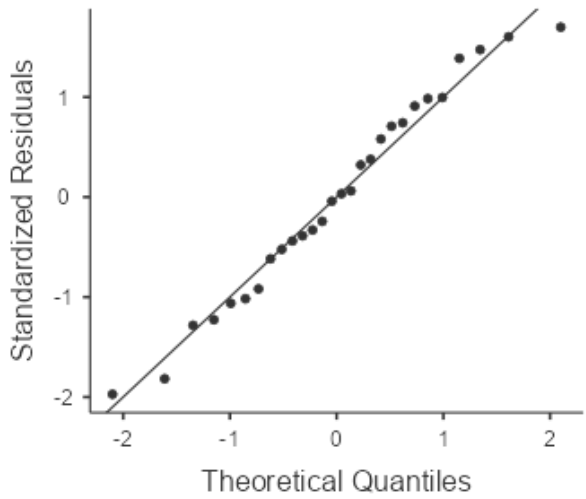
	W	p
Precision ZEN	0.973	0.670

Note. A low p-value suggests a violation of the assumption of normality

Plots

Precision ZEN





Post Hoc Tests

Games-Howell Post-Hoc Test – Precision ZEN

		300	400	QC-Mat.	3
300	Mean difference	—	-3.34	-8.16	-7.417**
	p-value	—	0.084	0.177	0.002
400	Mean difference		—	-4.81	-4.074
	p-value		—	0.481	0.121
QC-Mat.	Mean difference			—	0.740
	p-value			—	0.995
3	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

One-Way ANOVA (Welch's)

	F	df1	df2	p
MU95 ZEN	14.5	3	10.9	<.001

Assumption Checks

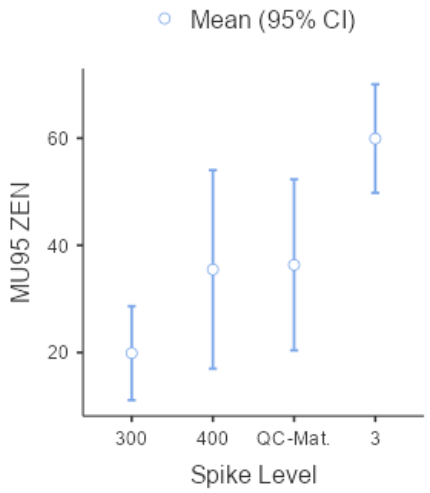
Normality Test (Shapiro-Wilk)

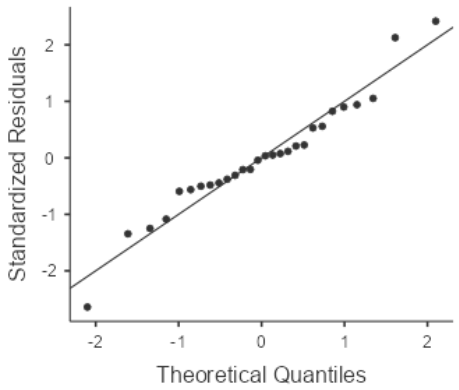
	W	p
MU95 ZEN	0.957	0.296

Note. A low p-value suggests a violation of the assumption of normality

Plots

MU95 ZEN





Post Hoc Tests

Games-Howell Post-Hoc Test – MU95 ZEN

		300	400	QC-Mat.	3
300	Mean difference	—	-15.6	-16.479	-40.0***
	p-value	—	0.309	0.127	<.001
400	Mean difference		—	-0.850	-24.4
	p-value		—	1.000	0.078
QC-Mat.	Mean difference			—	-23.6*
	p-value			—	0.033
3	Mean difference				—
	p-value				—

Note. * p < .05, ** p < .01, *** p < .001

Experimental Design plan for Ruggedness testing

Table 36. 2^4 DoE for Ruggedness Testing. Experimental plan with factors: Ethanol concentration c(EtOH), Extraction Time (Extr. T), Dilution of sample buffer (SB dil.) and Incubation time (Inc. T.). as well as achieved recovery rates for alle parameters.

	Factors				Recovery [%]						
Run	c(EtOH) (%vol)	Extr. T (min)	SB dil.	Inc. T. (min)	AFL	OTA	DON	FUM	ZEN	T-2/HT-2	Conc. Control
1	60	10	1:9	15	200	313	179	267	195	213	3.11
2	80	10	1:9	15	143	127	116	94	116	178	2.77
3	60	20	1:9	15	148	185	150	167	151	182	3.09
4	80	20	1:9	15	152	129	110	79	119	197	2.53
5	60	10	1:9	25	94	67	87	92	77	87	0.10
6	80	10	1:9	25	64	19	54	36	47	63	0.00
7	60	20	1:9	25	84	52	82	83	75	84	0.00
8	80	20	1:9	25	108	50	85	58	65	91	0.00
9	60	10	1:11	15	147	165	144	152	140	199	3.05
10	80	10	1:11	15	145	130	114	84	118	192	2.58
11	60	20	1:11	15	142	156	140	145	134	182	3.05
12	80	20	1:11	15	87	77	69	45	83	132	2.58
13	60	10	1:11	25	111	76	100	104	76	99	0.00
14	80	10	1:11	25	112	51	80	60	63	91	0.00
15	60	20	1:11	25	63	26	55	45	61	62	0.00
16	80	20	1:11	25	68	25	65	40	52	74	0.00
17	60	10	1:9	15	102	129	87	90	127	121	2.89
18	80	10	1:9	15	139	132	94	80	119	159	2.42
19	60	20	1:9	15	129	159	121	136	130	154	2.95
20	80	20	1:9	15	137	130	108	78	117	151	2.55
21	60	10	1:9	25	94	100	92	115	76	77	0.00
22	80	10	1:9	25	103	54	75	56	65	78	0.00
23	60	20	1:9	25	99	78	89	99	72	85	0.00
24	80	20	1:9	25	104	53	77	56	66	79	0.00
25	60	10	1:11	15	134	170	128	134	134	154	2.92
26	80	10	1:11	15	145	131	97	83	117	129	2.59
27	60	20	1:11	15	170	189	132	152	165	206	4.25
28	80	20	1:11	15	163	163	116	124	130	158	2.68
29	60	10	1:11	25	103	73	92	100	70	82	0.04
30	80	10	1:11	25	109	55	74	60	65	77	0.00
31	60	20	1:11	25	102	72	89	92	74	85	0.00

32	80	20	1:11	25	64	22	50	36	51	55	0.00

ANOVA of DoE: Parameter AFL

ANOVA - Recovery rate

	Sum of Squares	df	Mean Square	F	p
Overall model	22100.8125	10	2210.0812	3.7146	0.005
Ethanol concentration (%vol)	195.0312	1	195.0312	0.3278	0.573
Extraction time	488.2812	1	488.2812	0.8207	0.375
concentration sample buffer	38.2813	1	38.2813	0.0643	0.802
Incubation time assay	20050.0312	1	20050.0312	33.6991	<.001
Ethanol concentration (%vol) ✻ Extraction time	26.2813	1	26.2813	0.0442	0.836
Ethanol concentration (%vol) ✻ concentration sample buffer	195.0313	1	195.0313	0.3278	0.573
Extraction time ✻ concentration sample buffer	892.5313	1	892.5313	1.5001	0.234
Ethanol concentration (%vol) ✻ Incubation time assay	57.7813	1	57.7813	0.0971	0.758
Extraction time ✻ Incubation time assay	157.5313	1	157.5313	0.2648	0.612
concentration sample buffer ✻ Incubation time assay	0.0313	1	0.0313	5.25e-5	0.994
Residuals	12494.4062	21	594.9717		

Assumption Checks

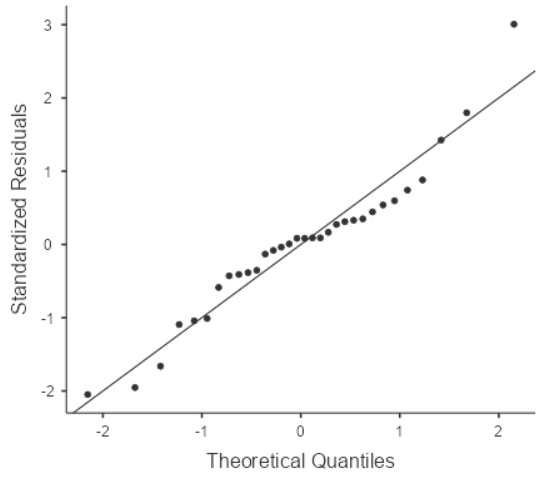
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
4.22	15	16	0.003

Normality Test (Shapiro-Wilk)

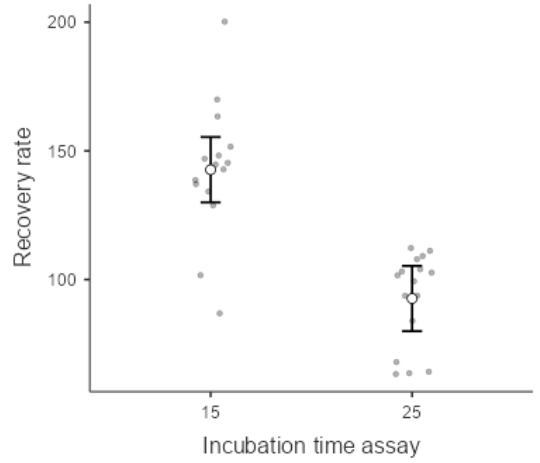
Statistic	p
0.945	0.104

Q-Q Plot

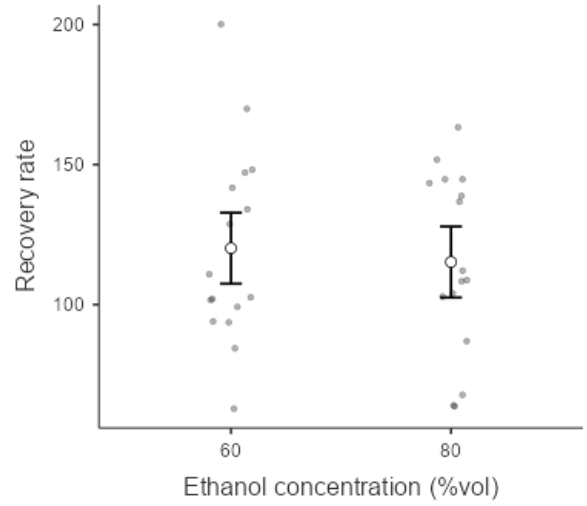


Estimated Marginal Means

Incubation time assay



Ethanol concentration (%vol)



ANOVA of DoE: Parameter Control

ANOVA – Concentration "Control"

	Sum of Squares	df	Mean Square	F	p
Overall model	67.4732	10	6.7473	130.450	<.001
Extraction time	0.0458	1	0.0458	0.885	0.358
concentration sample buffer	0.0553	1	0.0553	1.069	0.313
incubation time assay	65.7518	1	65.7518	1271.220	<.001
Ethanol concentration	0.7051	1	0.7051	13.632	0.001
Ethanol concentration ✻ Extraction time	0.0488	1	0.0488	0.944	0.342
Ethanol concentration ✻ concentration sample buffer	0.0319	1	0.0319	0.616	0.441
Extraction time ✻ concentration sample buffer	0.0751	1	0.0751	1.452	0.242
Ethanol concentration ✻ incubation time assay	0.6244	1	0.6244	12.072	0.002
Extraction time ✻ incubation time assay	0.0694	1	0.0694	1.341	0.260
concentration sample buffer ✻ incubation time assay	0.0657	1	0.0657	1.270	0.272
Residuals	1.0862	21	0.0517		

Assumption Checks

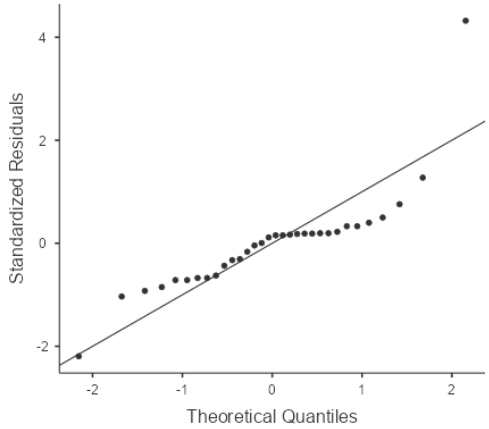
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
5.40	15	16	<.001

Normality Test (Shapiro-Wilk)

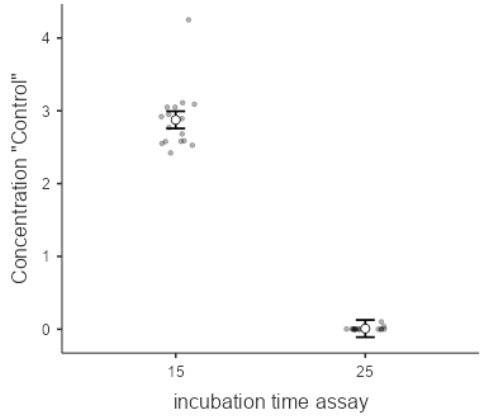
Statistic	p
0.765	<.001

Q-Q Plot

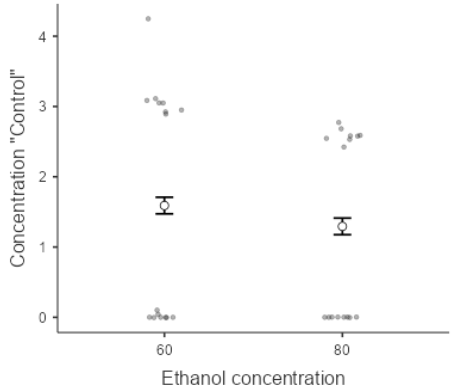


Estimated Marginal Means

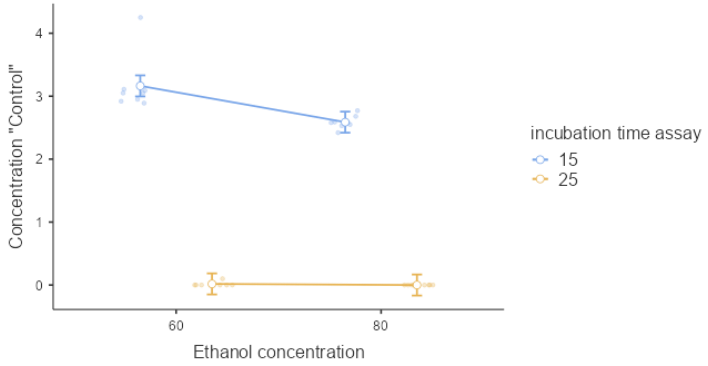
incubation time assay



Ethanol concentration



Ethanol concentration ✻ incubation time assay



ANOVA for Doe: Parameter DON

ANOVA – recovery rate

	Sum of Squares	df	Mean Square	F	p
Overall model	19866.81	10	1986.68	5.3384	<.001
Extraction time	175.78	1	175.78	0.4723	0.499
concentration sample buffer	116.28	1	116.28	0.3125	0.582
incubation time assay	13571.28	1	13571.28	36.4672	<.001
Ethanol concentration	4584.03	1	4584.03	12.3177	0.002
Ethanol concentration ✻ Extraction time	22.78	1	22.78	0.0612	0.807
Ethanol concentration ✻ concentration sample buffer	69.03	1	69.03	0.1855	0.671
Extraction time ✻ concentration sample buffer	712.53	1	712.53	1.9146	0.181
Ethanol concentration ✻ incubation time assay	536.28	1	536.28	1.4410	0.243
Extraction time ✻ incubation time assay	75.03	1	75.03	0.2016	0.658
concentration sample buffer ✻ incubation time assay	3.78	1	3.78	0.0102	0.921
Residuals	7815.16	21	372.15		

Assumption Checks

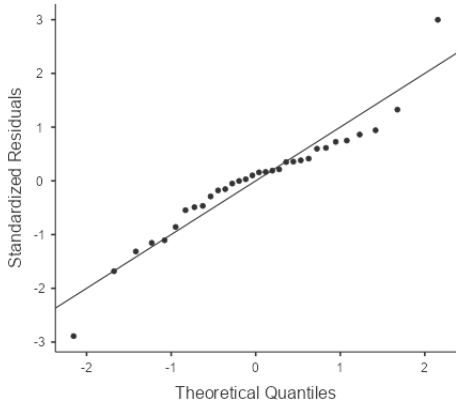
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
10.7	15	16	<.001

Normality Test (Shapiro-Wilk)

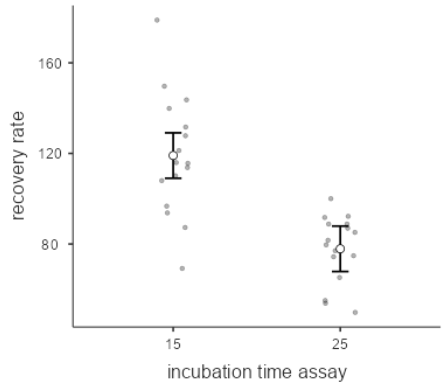
Statistic	p
0.936	0.058

Q-Q Plot

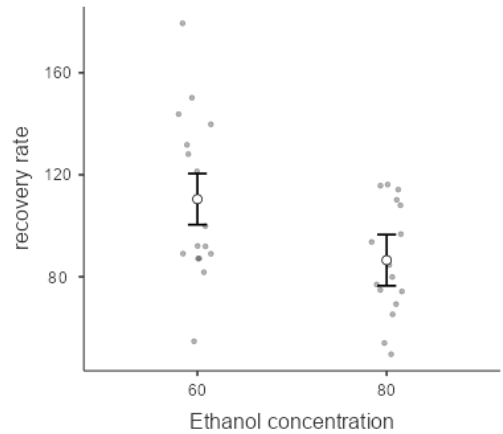


Estimated Marginal Means

Incubation time



Ethanol concentration



ANOVA of DoE: Parameter FUM

ANOVA – recovery rate

	Sum of Squares	df	Mean Square	F	p
Overall model	48700.00	10	4870.00	4.50491	0.002
Extraction time	924.50	1	924.50	0.85519	0.366
concentration sample buffer	528.12	1	528.12	0.48853	0.492
incubation time assay	18915.12	1	18915.12	17.49713	<.001
Ethanol concentration	25538.00	1	25538.00	23.62351	<.001
Ethanol concentration ✻ Extraction time	300.12	1	300.12	0.27763	0.604
Ethanol concentration ✻ concentration sample buffer	450.00	1	450.00	0.41627	0.526
Extraction time ✻ concentration sample buffer	18.00	1	18.00	0.01665	0.899
Ethanol concentration ✻ incubation time assay	1922.00	1	1922.00	1.77791	0.197
Extraction time ✻ incubation time assay	98.00	1	98.00	0.09065	0.766
concentration sample buffer ✻ incubation time assay	6.13	1	6.13	0.00567	0.941
Residuals	22701.87	21	1081.04		

Assumption Checks

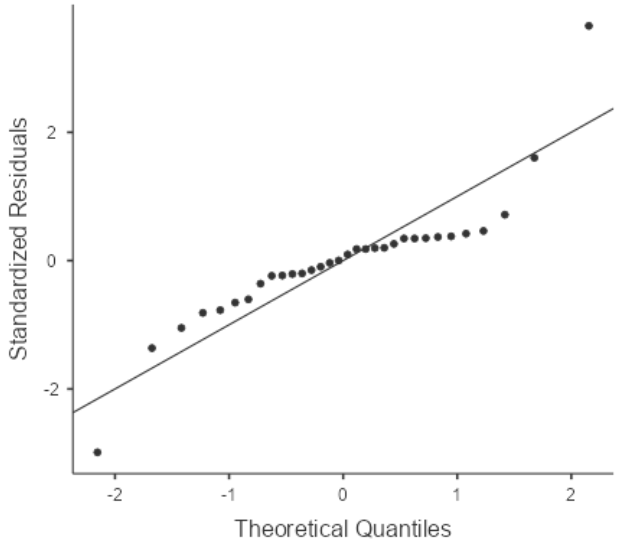
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
21.3	15	16	<.001

Normality Test (Shapiro-Wilk)

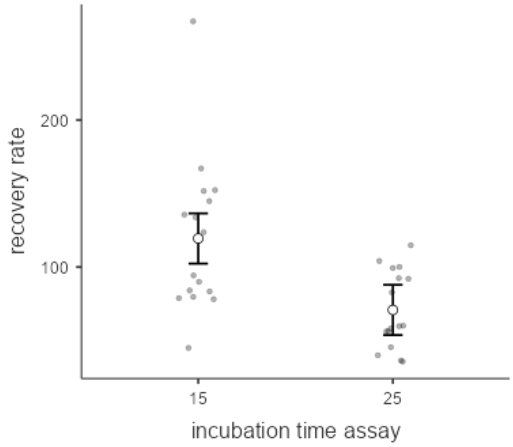
Statistic	p
0.828	<.001

Q-Q Plot

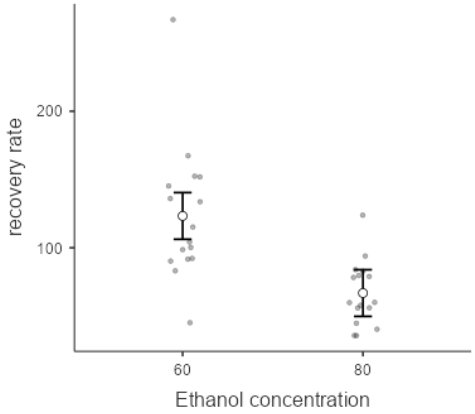


Estimated Marginal Means

incubation time assay



Ethanol concentration



ANOVA of DoE: Parameter OTA

ANOVA – recovery rate

	Sum of Squares	df	Mean Square	F	p
Overall model	100383.00	10	10038.30	7.88028	<.001
Extraction time	1596.12	1	1596.12	1.25299	0.276
concentration sample buffer	1200.50	1	1200.50	0.94242	0.343
incubation time assay	81204.50	1	81204.50	63.74724	<.001
Ethanol concentration	13695.12	1	13695.12	10.75096	0.004
Ethanol concentration ✻ Extraction time	496.13	1	496.13	0.38947	0.539
Ethanol concentration ✻ concentration sample buffer	420.50	1	420.50	0.33010	0.572
Extraction time ✻ concentration sample buffer	8.00	1	8.00	0.00628	0.938
Ethanol concentration ✻ incubation time assay	1682.00	1	1682.00	1.32041	0.263
Extraction time ✻ incubation time assay	2.00	1	2.00	0.00157	0.969
concentration sample buffer ✻ incubation time assay	78.13	1	78.13	0.06133	0.807
Residuals	26750.87	21	1273.85		

Assumption Checks

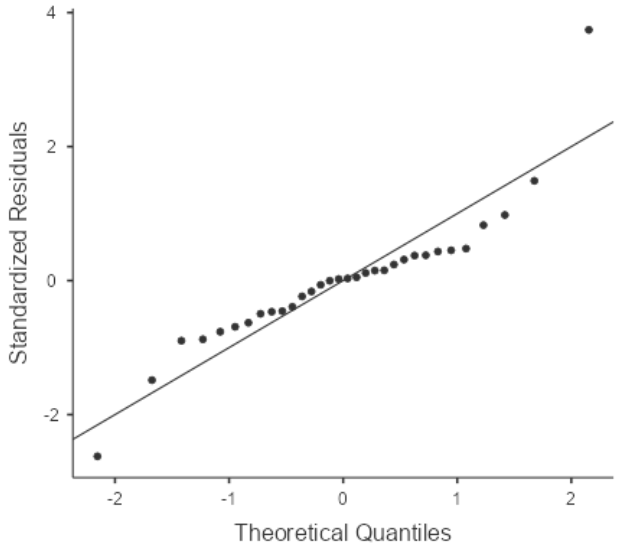
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
11.7	15	16	<.001

Normality Test (Shapiro-Wilk)

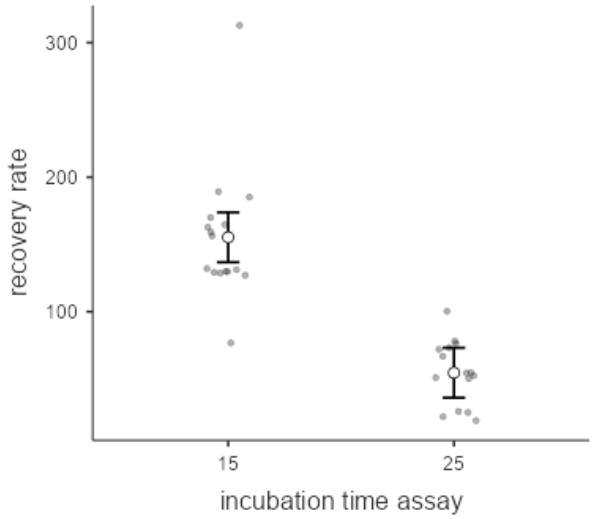
Statistic	p
0.867	0.001

Q-Q Plot

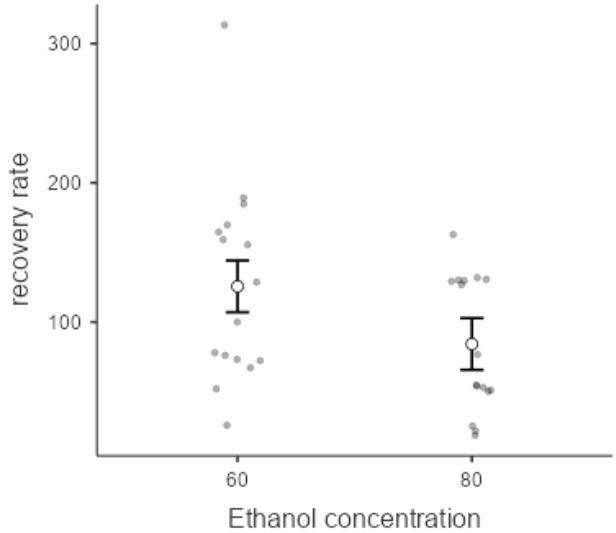


Estimated Marginal Means

incubation time assay



Ethanol concentration



ANOVA of DoE: Parameter T-2/HT-2

ANOVA - recovery rate

	Sum of Squares	df	Mean Square	F	p
Overall model	66965.25	10	6696.53	11.8465	<.001
Extraction time	15.13	1	15.13	0.0268	0.872
concentration sample buffer	15.13	1	15.13	0.0268	0.872
incubation time assay	64620.13	1	64620.13	114.3165	<.001
Ethanol concentration	882.00	1	882.00	1.5603	0.225
Ethanol concentration ✻ Extraction time	45.13	1	45.13	0.0798	0.780
Ethanol concentration ✻ concentration sample buffer	741.12	1	741.12	1.3111	0.265
Extraction time ✻ concentration sample buffer	420.50	1	420.50	0.7439	0.398
Ethanol concentration ✻ incubation time assay	120.13	1	120.13	0.2125	0.650
Extraction time ✻ incubation time assay	98.00	1	98.00	0.1734	0.681
concentration sample buffer ✻ incubation time assay	8.00	1	8.00	0.0142	0.906
Residuals	11870.75	21	565.27		

Assumption Checks

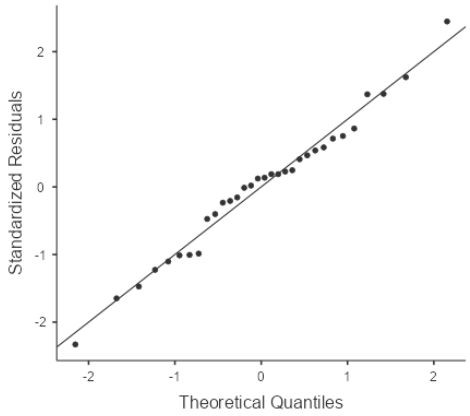
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
3.43	15	16	0.010

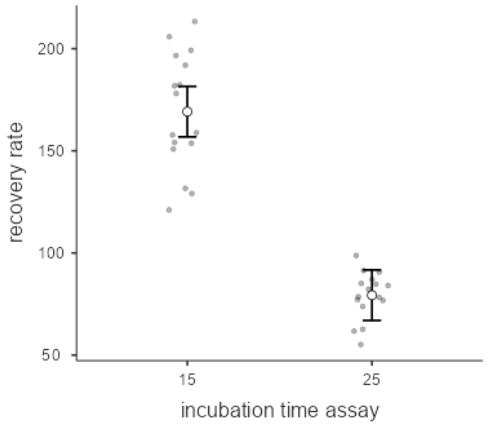
Normality Test (Shapiro-Wilk)

Statistic	p
0.984	0.903

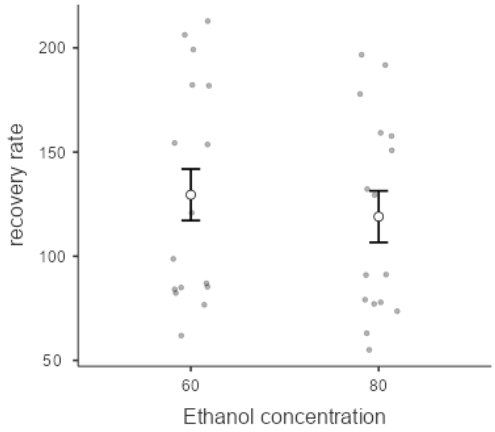
Q-Q Plot



Estimated Marginal Means
incubation time assay



Ethanol concentration



ANOVA of DoE: Parameter ZEN

ANOVA – recovery rate

	Sum of Squares	df	Mean Square	F	p
Overall model	39006.25	10	3900.62	15.52083	<.001
Extraction time	112.50	1	112.50	0.44764	0.511
concentration sample buffer	220.50	1	220.50	0.87738	0.360
incubation time assay	33800.00	1	33800.00	134.49231	<.001
Ethanol concentration	4140.50	1	4140.50	16.47531	<.001
Ethanol concentration ✻ Extraction time	1.13	1	1.13	0.00448	0.947
Ethanol concentration ✻ concentration sample buffer	6.13	1	6.13	0.02437	0.877
Extraction time ✻ concentration sample buffer	1.13	1	1.13	0.00448	0.947
Ethanol concentration ✻ incubation time assay	703.12	1	703.12	2.79778	0.109
Extraction time ✻ incubation time assay	6.13	1	6.13	0.02437	0.877
concentration sample buffer ✻ incubation time assay	15.13	1	15.13	0.06018	0.809
Residuals	5277.63	21	251.32		

Assumption Checks

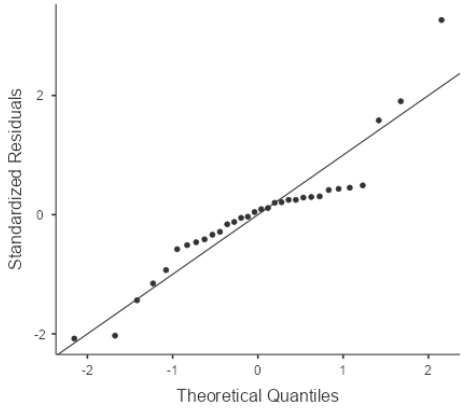
Homogeneity of Variances Test (Levene's)

F	df1	df2	p
4.89	15	16	0.002

Normality Test (Shapiro-Wilk)

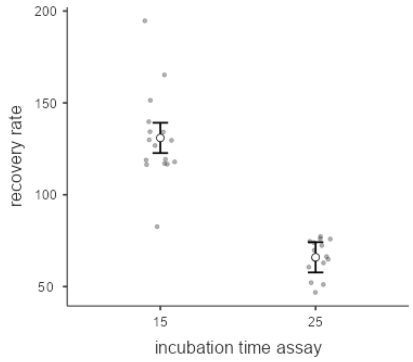
Statistic	p
0.895	0.005

Q-Q Plot

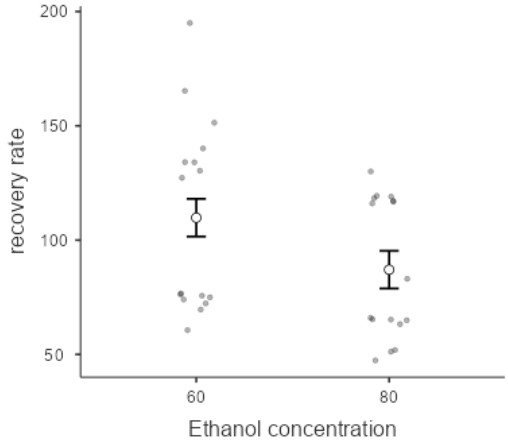


Estimated Marginal Means

incubation time assay



Ethanol concentration



References

[1] The jamovi project (2024). jamovi. (Version 2.6) [Computer Software]. Retrieved from https://www.jamovi.org.

[2] R Core Team (2024). R: A Language and environment for statistical computing. (Version 4.4) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2024-08-07).

[3] Fox, J., & Weisberg, S. (2023). car: Companion to Applied Regression. [R package]. Retrieved from https://cran.r-project.org/package=car.

[4] Lenth, R. (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. [R package]. Retrieved from https://cran.r-project.org/package=emmeans.

Contact

SAFIA Technologies GmbH
c/o Bundesanstalt für Materialforschung und –prüfung (BAM)
Richard-Willstätter-Straße 11
12489 Berlin

T: +49 (0) 162 7323 405

M: info@safia.tech

SAFIA LinkedIn Profile (Link)

SAFIA YouTube Channel (Link)

Your Contact

Peter Carl
Research & Development

T: +49 (0) 172 4260 922

M: carl@safia.tech

Peter Carl LinkedIn Profile (Link)

www.SAFIA.tech