DLA
Dienstleistung
Lebensmittel
Analytik GbR

Evaluation Report
proficiency test

03/2016

Allergens III:

β-Lactoglobulin, Casein and Gluten

in Infant Food

Dienstleistung Lebensmittel Analytik GbR
Waldemar-Bonsels-Weg 170
22926 Ahrensburg, Germany

proficiency-testing@dla-lvu.de www.dla-lvu.de

Coordinator of this PT:
Dr. Matthias Besler

Reprint, also in part, only with written permission from DLA-Ahrensburg

Inhalt / Content

Reprint, also in part, only with written permission from DLA-Ahrensburg

1. Introduction

The participation in proficiency testing schemes is an essential element of the quality-management-system of every laboratory testing food and feed, cosmetics and food contact materials. The implementation of proficiency tests enables the participating laboratories to prove their own analytical competence under realistic conditions. At the same time they receive valuable data regarding the verification and/or validation of the particular testing method [1, 5].
The purpose of DLA is to offer proficiency tests for selected parameters in concentrations with practical relevance.
Realisation and evaluation of the present proficiency test follows the technical requirements of DIN EN ISO/IEC 17043 (2010) and DIN ISO 13528:2009 / ISO 13528:2015 [2, 3].

2. Realisation

2.1 Test material

Two PT-samples for the detection of allergens in the range of mg/kg and one spiking material sample were provided for analysis. The spiking material sample contains the respective allergenic ingredients in the range of 1-10 % and was added to the spiked PT-sample. The results of the spiking material sample should give the possibility of a comparison with the spiked sample in respect to the detectability of the allergens with and without the influence of matrix and / or food processing.

The test material is an infant food mixture of a common in commerce "millet-pap" powder from 4th month and a "rice-pap" powder from 6th month (each labeled "milk and gluten free"). The basic composition of both sample A and sample B was the same (see table 1). After crushing, sieving and homogenisation of the basic mixture an aliquot of it was added stepwise during several homogenisations to the spiking material which contained the allergenic ingredients skimmed milk powder and wheat flour for preparation of sample B.
The composition of the spiking material sample and the amounts of allergens in sample B is given in table 2.

After homogenisation the samples were portioned to approximately 25 g into metallised PET film bags.

Reprint, also in part, only with written permission from DLA-Ahrensburg

Table 1: Composition of DLA-Samples

Ingredients	**Sample A**	**Sample B**
Organic-Millet-Pap with rice, infant pap after 4th month Ingredients: Millet whole flour (75%), rice whole flour (25%), vitamin B1 Nutrients per 100 g: Protein 11 g, carbohydrates 78 g, fat 3,7 g	50,0 g/100 g	48,5 g/100 g
Organic-Rice-Pap with corn and millet, infant pap after 6th month Ingredients: Rice whole flour (70%), corn flour (20%), millet whole flour (10%), vitamin B1 Nutrients per 100 g: Protein 8,6 g, carbohydrates 80 g, fat 2,8 g	50,0 g/100 g	48,5 g/100 g
Spiking material sample	-	2,97 g/100 g

Table 2: Added amounts of allergenic ingredients

Ingredients	**Spiking material sample**	**Amounts in Sample B**
Potato flour Nutrients per 100g: Protein 0 g	93 %	2,76 %
Milk: - as Skimmed Milk Powder - thereof Total Protein - thereof Casein* - *thereof β-Lactoglobulin**	 19600 mg/kg (1,96 %) 5740 mg/kg 4590 mg/kg 574 mg/kg	 582 mg/kg 192 mg/kg 154 mg/kg 19,2 mg/kg
Wheat: - as Wheat flour Type 1050 - thereof total protein* - *thereof gluten***	 15300 mg/kg (1,53 %) 1840 mg/kg 1650 mg/kg	 349 mg/kg 41,8 mg/kg 37,6 mg/kg
Soy flour	2,01 %	0,060 %
Hazelnut spread	1,18 %	0,035 %

* according to labelling and literature data
** Definition of "gluten" from the Gluten Intolerance Labelling Regulation (EU/41/2009) corresponds to 85-91% of wheat protein according to data from the literature

Reprint, also in part, only with written permission from DLA-Ahrensburg

2.1.1 Homogeneity

The **mixture homogeneity before bottling** was examined 8-fold by **microtracer analysis**. It is a standardized method that is part of the international GMP certification system for feed [14].
Before mixing dye coated iron particles of µm size are added to the sample and the number of particles is determined after homogenization in taken aliquots. The evaluation of the mixture homogeneity is based on the Poisson distribution using the chi-square test. A probability of ≥ 5 % is equivalent to a good homogeneous mixture and of ≥ 25% to an excellent mixture [14, 15].
The microtracer analysis of the present PT samples showed a probability of 43% for the spiking material sample and of 49% for the spiked sample B. Additionally particle number results were converted into concentrations, statistically evaluated according to normal distribution and compared to the standard deviation according to Horwitz. This gave a Hor-Rat value of 1,0 and 1,3 respectively. The results of microtracer analysis are given in the documentation.

The **homogeneity of bottled DLA-samples** (spiking material sample and spiked sample B) was checked by ELISA-test for gluten (fig. 1). The resulting standard deviations between the samples of < 15% ensured sufficient homogeneity [16, 17, 20, 21]. In case the criterion for sufficient homogeneity of the test items is not fulfilled the impact on the target standard deviation will be verified. If necessary the evaluation of results will be done considering the standard uncertainty of the assigned value (s. 3.8 and 3.11) [3].

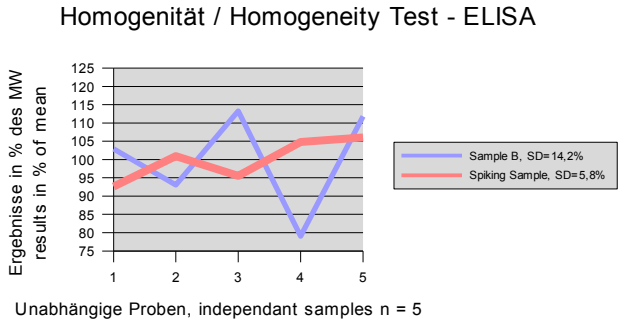

Fig. 1: Testing of homogeneity of DLA-sample B and spiking material sample. Results are given in percent of the arithmetic mean

Reprint, also in part, only with written permission from DLA-Ahrensburg

2.2 Sample shipment and information to the test

The portions of test material (sample A and sample B as well as the spiking material sample) were sent to every participating laboratory in the 7th week of 2016. The testing method was optional. The tests should be finished at April 1st 2016 the latest.

2.3 Submission of results

The participants submitted their results in standard forms, which have been sent by email or were available on our website. On one hand the results given as positive/negative and on the other hand the indicated results of the allergenic ingredients e.g. beta-lactoglobulin, casein or gluten in mg/kg were evaluated.
Queried and documented were the indicated results and details of the test methods like specifity, test kit manufacturer and hints about the procedure.
In case participants submitted several results for the same parameter obtained by different methods these results were evaluated with the same evaluation number with a letter as a suffix and indication of the related method.
One participant submitted the results delayed in agreement with DLA. All other participants submitted their results in time.

Reprint, also in part, only with written permission from DLA-Ahrensburg

3. Evaluation

Different ELISA-methods for the determination of allergens in foods are eventually using different antibodies, are usually calibrated with different reference materials and may utilize differing extraction methods. Among others this can induce different results of the content of the analyte [23, 24, 25, 26]. It is for this reason that we contrast the results of the present proficiency test with several assigned values. Thereby it is possible to evaluate each single result in comparison to the mean of all results and/or in comparison to the mean of results obtained by a single method. For comparison the actually added amount is plotted in the figures of the results.

For quantitative results of the spiking material sample and the spiked sample recovery rates were calculated with respect to the known content of spiked allergens. The recovery rates were given for information only. *No* statistical evaluation was done. The recovery rates should exclusively give an estimation of the matrix- and/or processing influences.

PCR results were valuated qualitatively with respect to the percentages of positive and negative results, respectively. If there are ≥ 75 % positive or negative results, a consensus result is determined for each sample.

3.1 Consensus value from participants (assigned value)

The robust mean of the submitted results was used as assigned value (X_{pt}) („consensus value from participants") providing a normal distribution. The calculation was done according to algorithm A as described in annex C of ISO 13528 [3].
The condition is that the majority of the participants' results show a normal distribution or are distributed unimodal and symmetrically. To this end, an examination of the distribution is carried out, inter alia, using the kernel density estimate [3, 12].

In case there are indications for sources of higher variability such as a bimodal distribution of results, a cause analysis is performed. Frequently different analytical methods may cause an anomaly in results' distribution. If this is the case, separate evaluations with own assigned values (X_{pti}) are made whenever possible.

If possible, this is the standard procedure for the evaluation of ELISA methods for the determination of allergens:

i) ***Robust mean of all results*** - $X_{pt_{ALL}}$
ii) ***Robust mean of single methods*** - $X_{pt_{METHOD\ i}}$
with at least 5 quantitative results given.

Single results giving values outside the measuring range of the participating laboratory or given as „0" are not considered for statistical evaluation (e.g. results given as > 25 mg/kg and < 2,5 mg/kg, respectively) [3].

Reprint, also in part, only with written permission from DLA-Ahrensburg

3.2 Robust standard deviation

For comparison to the target standard deviation σ_{pt} (standard deviation for proficiency assessment) a robust standard deviation (S^x) was calculated. The calculation was done according to algorithm A as described in annex C of ISO 13528 [3].

The following robust standard deviations were considered:

i) ***Robust standard deviation of all results*** - S^x_{ALL}
ii) ***Robust standard deviation of single methods*** - $S^x_{METHOD\ i}$
with at least 5 quantitative results given.

3.3 Exclusion of results and outliers

Before statistical evaluation obvious blunders, such as those with incorrect units, decimal point errors, and results for a another proficiency test item can be removed from the data set [2]. All results should be given at least with 2 significant digits. Specifying 3 significant digits is usually sufficient.

Results obtained by different analytical methods causing an increased variability and/or a bi- or multimodal distribution of results, are treated separately or could be excluded in case of too few numbers of results. For this results are checked by kernel density estimation [3, 12].

Results are identified as outliers by the use of robust statistics. If a value deviates from the robust mean by more than 3 times the robust standard deviation, it is classified as an outlier [3]. Detected outliers are stated for information only, when z-score are < -2 or > 2. Due to the use of robust statistics outliers are not excluded, provided that no other reasons are present [3].

Reprint, also in part, only with written permission from DLA-Ahrensburg

3.4 Target standard deviation (for proficiency assessment)

The target standard deviation of the assigned value σ_{pt} (= standard deviation for proficiency assessment) can be determined according to the following methods.
In the present PT the target standard deviation was determined according to 3.4.3 value by perception.

3.4.1 General model (Horwitz)

Based on statistical characteristics obtained in numerous PTs for different parameters and methods Horwitz has derived a general model for estimating the reproducibility standard deviation σ_R [6]. Later the model was modified by Thompson for certain concentration ranges [10]. The reproducibility standard deviation σ_R can be applied as the relative target standard deviation σ_{pt} in % of the assigned values and calculated according to the following equations [3]. For this the assigned value X_{pt} is used for the concentration c.

Equations	**Range of concentrations**	**corresponds to**
$\sigma_R = 0{,}22c$	$c < 1{,}2 \times 10^{-7}$	< 120 µg/kg
$\sigma_R = 0{,}02c^{0{,}8495}$	$1{,}2 \times 10^{-7} \leq c \leq 0{,}138$	≥ 120 µg/kg
$\sigma_R = 0{,}01c^{0{,}5}$	$c > 0{,}138$	> 13,8 g/100g

with c = mass content of analyte (as relative size, e.g. 1 mg/kg = 1 ppm = 10^{-6} kg/kg)

The target standard deviation according to Horwitz is currently not achievable by ELISA-methods for values in the mg/kg range and was therefore not considered for evaluation.

3.4.2 Value by precision experiment

Using the reproducibility standard deviation σ_R and the repeatability standard deviation σ_r of a precision experiment (collaborative trial or proficiency test) the target standard deviation σ_{pt} can be derived considering the number of replicate measurements m of participants in the present PT [3]:

$$\sigma_{pt} = \sqrt{\sigma_R^2 - \sigma_r^2\left(m - 1/m\right)}$$

Because in the present proficiency test the number of replicate measurements is n = 1, the reproducibility standard deviation σ_R is identical to the target standard deviation σ_{pt}.

Reprint, also in part, only with written permission from DLA-Ahrensburg

The following table shows the relative reproducibility standard deviations from proficiency tests of ELISA-methods from German ASU §64 methods [27, 28, 29]:

Method	Parameter	Matrix	Mean values	Relative σ_R	Literature
ELISA	Soy protein	Sausage	0,36 - 4,07%	14 - 28%	L 06.00-56
ELISA (Manuf. A)	Peanut	Milk chocolate	5,9 - 174 mg/kg	20 - 31%	L 00.00-69
ELISA (Manuf. B)	Peanut	Milk chocolate	10,1 - 216 mg/kg	14 - 32%	L 00.00-69
ELISA (Manuf. A)	Peanut	Dark chocolate	5,7 - 148 mg/kg	22 - 33%	L 00.00-69
ELISA (Manuf. A)	Hazelnut	Dark chocolate	1,6 - 16,3 mg/kg	12 - 33%	L 44.00-7
ELISA (Manuf. A)	Hazelnut	Dark chocolate	2,4 - 21,3 mg/kg	14 - 19%	L 44.00-7

From these precision data of the ASU §64 methods the calculated relative target standard deviations are in the range of 12 - 33%.

Štumr et al. conducted two interlaboratory studies for the validation of commercial ELISA-Test-Kits for the determination of β-lactoglobulin and for the determination of casein [30, 31].
20 food samples with β-lactoglobulin contents in the range of 0 - 33 mg/kg were analyzed by 6 laboratories. Recovery rates ranged between 91 - 118%. Relative repeatability standard deviations ranged from 5,8 - 13% and the relative reproducibility standard deviations ranged from 26 - 49% [30].
Casein was analyzed by 8 laboratories in 10 food samples in the range of 0 - 30 mg/kg and in 3 food samples with contents >30 mg/kg. Recovery rates ranged between 67 - 81%. Relative repeatability standard deviations ranged from 11 - 52% and was for one sample Probe 99% and the relative reproducibility standard deviations ranged from 13 - 61% and were for two samples 96% and 111%, respectively [31].
According to the authors both ELISA-Test-Kits were acceptable for routine control of food samples [30, 31].

The Working Group on Prolamin Analysis and Toxicity (WGPAT) coordinated a collaborative study with two commercial ELISA-Test-Kits for the determination of gluten using the monoclonal R5 antibody [22]. 12 food samples with gliadin in the range of 0 - 168 mg/kg were analyzed by 20 laboratories. Recovery rates ranged between 65 and 110%, relative repeatability deviations ranged from 13 - 25% (method 1) and 11 - 22% (method 2) while the relative reproducibility standard deviations ranged from 23 - 47% (method 1) and 25 - 33% (method 2). According to the authors both ELISA-Test-Kits fulfilled therefore the current validation criteria for ELISA methods [22].

The IRMM (Institute for Reference Materials and Measurements) performed an interlaboratory comparison for five different ELISA-test kits for the quantification of peanut [25]. The mean values for two matrices were in

Reprint, also in part, only with written permission from DLA-Ahrensburg

the concentration range of 0,3 - 16,1 mg/kg and 1,2 - 20,4 mg/kg, respectively. The lowest relative reproducibility standard deviations of the five test kits were for dark chocolate in the range of 20 - 42% and for cookies in the range of 23 – 61%.

3.4.3 Value by perception

The target standard deviation for proficiency assessment can be set at a value that corresponds to the level of performance that the coordinator would wish laboratories to be able to achieve [3].
Criteria for the level of performance of analytical methods for the quantitative determination of allergens in foods were recently elaborated e.g. by the Ministry of Health and Welfare (MHLW) in Japan [20], by the working group 12 „Food Allergens" of the technical committee CEN/TC 275 [17-19], by an international "Food Allergen Working Group" under the advice of the AOAC Presidential Task Force on Food Allergens [21] and by the Codex Alimentarius Committee (CAC/GL 74-2010) [16].

Some of the relevant ELISA and PCR validation criteria of the mentioned panels are listed in tables 3 and 4, respectively.

Table 3: ELISA-Validation

Literature [16-22]	**Recovery rate**	**Repeatability standard deviation**	**Reproducibility standard deviation**
MHLW 2006	50 - 150%		≤ 25%
CEN 2009		≤ 20%	
AOAC 2010	50 - 150%	6,9 - 34,4% (a)	19,5 - 57,2 (a)
CAC 2010	70 - 120%	≤ 25%	≤ 35%

(a) = Example from an hypothetical proficiency scheme in the range of 0,5 - 5 mg/kg

Table 4: PCR-Validation

Literature [16]	**Recovery rate**	**Repeatability standard deviation**	**Reproducibility standard deviation**
CAC 2010	± 25% (a)	≤ 25%	≤ 35%

(a) = Trueness / Richtigkeit

Based on the currently achievable level of performance of ELISA and PCR methods for the quantitative determination of allergens in foods, which could be deduced from the data of precision experiments and from validation criteria, we set a relative target standard deviation σ_{pt} of 25%.
This target standard deviation was applied for the statistical evaluation of the results by z-score and was used for all assigned values mentioned in 3.1.

Reprint, also in part, only with written permission from DLA-Ahrensburg

3.5 z-Score

To assess the results of the participants the z-score is used. It indicates about which multiple of the target standard deviation (σ_{pt}) the result (x_i) of the participant is deviating from the assigned value (X_{pt}) [3].
Participants' z-scores are derived from:

$$z_i = \frac{\left(x_i - x_{pt}\right)}{\sigma_{pt}}$$

The requirements for the analytical performance are generally considered as fulfilled if

$$-2 \leq z \leq 2 \ .$$

For information the z-scores below are calculated with a target standard deviation of 25%:

i) ***z-Score*** - z_{ALL} *(with respect to all methods)*
ii) ***z-Score*** - $z_{METHOD\ i}$ *(with respect to single methods)*

3.5.1 Warning and action signals

In accordance with the norm ISO 13528 it is recommended that a result that gives rise to a *z*-score above 3,0 or below −3,0, shall be considered to give an "action signal" [3]. Likewise, a *z*-score above 2,0 or below −2,0 shall be considered to give a "warning signal". A single "action signal", or "warning signal" in two successive PT-rounds, shall be taken as evidence that an anomaly has occurred which requires investigation. For example a fault isolation or a root cause analysis through the examination of transmission error or an error in the calculation, in the trueness and precision must be performed and if necessary appropriate corrective measures should be applied [3].

In the figures of z-scores DLA gives the limits of warning and action signals as yellow and red lines respectively. According to ISO 13528 the signals are valid only in case of a number of ≥ 10 results [3].

Reprint, also in part, only with written permission from DLA-Ahrensburg

3.6 z'-Score

The z'-score can be used for the valuation of the results of the participants, in cases the standard uncertainty has to be considered (s. 3.8). The z'-score represents the relation of the deviation of the result (x) of the participant from the respective consensus value (X) to the square root of quadrat sum of the target standard deviation ($\hat{\sigma}$) and the standard uncertainty (Ux_{pt}) [3].

The calculation is performed by:

$$z'_i = \frac{x_i - x_{pt}}{\sqrt{\sigma_{pt}^2 + u_{(x_{pt})}^2}}$$

If carried out an evaluation of the results by means of z 'score, we have defined below the expression in the denominator as a target standard deviation σ_{pt}'.

The requirements for the analytical performance are generally considered as fulfilled if

$$-2 \leq z' \leq 2 \ .$$

For warning and action signals see 3.6.1.

3.7 Quotient S*/σpt

Following the HorRat-value the results of a proficiency-test (PT) can be considered convincing, if the quotient of robust standard deviation S* and target standard deviation σ_{pt} does not exceed the value of 2.
A value > 2 means an insufficient precision, i.e. the analytical method is too variable, or the variation between the test participants is higher than estimated. Thus the comparability of the results is not given [3].

3.8 Standard uncertainty of the assigned value

Every assigned value has a standard uncertainty that depends on the analytical method, differences between the analytical methods used, the test material, the number of participating laboratories (P) and on other factors. The standard uncertainty ($U_{(Xpt)}$) for this PT is calculated as follows [3]:

$$u_{(x_{pt})} = 1{,}25 \times \frac{s^*}{\sqrt{p}}$$

If $U_{(Xpt)} \leq 0{,}3\ \sigma_{pt}$ the standard uncertainty of the assigned value needs not to be included in the interpretation of the results of the PT [3].

Reprint, also in part, only with written permission from DLA-Ahrensburg

Values exceeding 0,3 imply, that the target standard deviation could be too low with respect to the standard uncertainty of the assigned value. The Quotient $u(x_{pt})/\sigma_{pt}$ is reported in the characteristics of the test.

3.9 Figures

The assigned values and spiking levels are indicated as coloured lines in the figures of results. This allows the comparison of a single result with different possible target values like the spiked level, the robust mean of all results and the robust mean of a single method.

3.10 Recovery rates: Spiking

For the results of the spiking material sample and the spiked sample recovery rates were calculated with respect to the known content of added allergens. The related values of added allergens are given in 2.1 test material in table 2. As a range of acceptance RA for valuating participant's results the range of 50 - 150% for the recovery rates of allergen-ELISAs proposed by the AOAC was used [21]. For quantitative PCR determinations we use the same range of acceptance.

Reprint, also in part, only with written permission from DLA-Ahrensburg

4. Results

All following tables are anonymized. With the delivering of the evaluation-report the participants are informed about their individual evaluation-number.
The following result sections are structured equally for the allergenic components. First all results for a certain analyte are reported together for sample A and afterwards for sample B.

To ensure the **comparability of quantitative results** DLA harmonized participants' results giving different specifications (e.g. as protein or as allergenic food) as far as possible.

ELISA-results given as gliadin were converted into gluten multiplying the gliadin-content with the factor of 2.

Evaluation was done separately for ELISA and PCR-techniques. The results were grouped according to the applied methods (e.g. test-kits) and sorted chronologically according to the evaluation-number of the participants.

Results were valuated qualitatively with respect to the percentages of positive and negative results, respectively. If there are ≥ 75 % positive or negative results, a consensus result is determined for each sample. Each participant result is valuated qualitatively with respect to the consensus value. The valuation was given as a percentage of results in agreement with the consensus values.

When there are at least 5 quantitative results for all methods or for single methods a statistical evaluation was done.

In cases when a statistical evaluation of the quantitative values was done the result table was given as indicated below:

Evaluation number	Result	Result	z-Score Xpt_{ALL}	z-Score $Xpt_{M\,i}$	Method	Remarks
	pos/neg	[mg/kg]				

Reprint, also in part, only with written permission from DLA-Ahrensburg

The statistical evaluation of results for each parameter was calculated in cases where at least 50% results were positive and at least 5 quantitative values were given:

Characteristics	**All Results** *[mg/kg]*	**Method i** *[mg/kg]*
Assigned value (X_{pt})	$\boldsymbol{X_{pt_{ALL}}}$	$\boldsymbol{X_{pt_{METHOD\ i}}}$
Number of results		
Number of outliers		
Median		
Robust mean (X_{pt})		
Robust standard deviation (S*)		
Target data:		
Target standard deviation σ_{pt}		
lower limit of target range ($X_{pt} - 2\sigma_{pt}$)		
upper limit of target range ($X_{pt} + 2\sigma_{pt}$)		
Quotient S*/σ_{pt}		
Standard uncertainty $U_{(Xpt)}$		
Quotient $U_{(Xpt)}/\sigma_{pt}$		
Number of results in target range		
Percent in target range		

After that the recovery rates of the results for the spiking sample and the spiked sample are reported. The number of results within the range of acceptance of 50-150% is given.

Reprint, also in part, only with written permission from DLA-Ahrensburg

4.1 Proficiency Test Milk

4.1.1 ELISA-Results: β-Lactoglobulin

Qualitative valuation of results: Samples A and B

Evaluation number	Sample A	Sample A	Sample B	Sample B	Qualitative Valuation	Method	Remarks
	pos/neg	[mg/kg]	pos/neg	[mg/kg]	Agreement with consensus value		
9	negative	<5	positive	63	2/2 (100%)	BK	
6	negative	<0,1	positive	22	2/2 (100%)	ES	
13	negative	<0,2	positive	14,1	2/2 (100%)	ES	
16	negative	<0,1	positive	19	2/2 (100%)	ES	
17	negative	<0,05	positive	>1	2/2 (100%)	ES	
14	negative	< 0,01	positive	20	2/2 (100%)	IL	
1	negative	<0.5	positive	5,91	2/2 (100%)	RS	
2	negative	< 0.5	positive	4,3	2/2 (100%)	RS	
3	negative		positive	6,9	2/2 (100%)	RS	
4	negative	<0,5	positive	31	2/2 (100%)	RS	outlier Xpt RS
7	negative		positive	8,6	2/2 (100%)	RS	
8	negative	<1	positive	1,8	2/2 (100%)	RS	
15	negative	<5	positive	8,3	2/2 (100%)	RS	

	Sample A		Sample B	
Number positive	0		13	
Number negative	13		0	
Percent positive	0		100	
Percent negative	100		0	
Consensus value	negative		positive	

Methods:

BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

Comments:

There were 100% negative results for sample A and 100% positive results for sample B by the ELISA-methods.
The consensus values are in agreement with the spiking of sample B.

Reprint, also in part, only with written permission from DLA-Ahrensburg

Quantitative valuation of results: Sample B

Evaluation number	β-Lacto-globulin	z-Score Xpt_{ALL}	z-Score Xpt_{RS}	Method	Remarks
	[mg/kg]				
9	63			BK	
6	22			ES	
13	14,1			ES	
16	19			ES	
17	>1			ES	
14	20			IL	
1	5,91		-0,6	RS	
2	4,3		-1,5	RS	
3	6,9		0,0	RS	
4	31		13,8	RS	outlier Xpt RS
7	8,6		0,9	RS	
8	1,8		-3,0	RS	
15	8,3		0,8	RS	

Methods:

BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

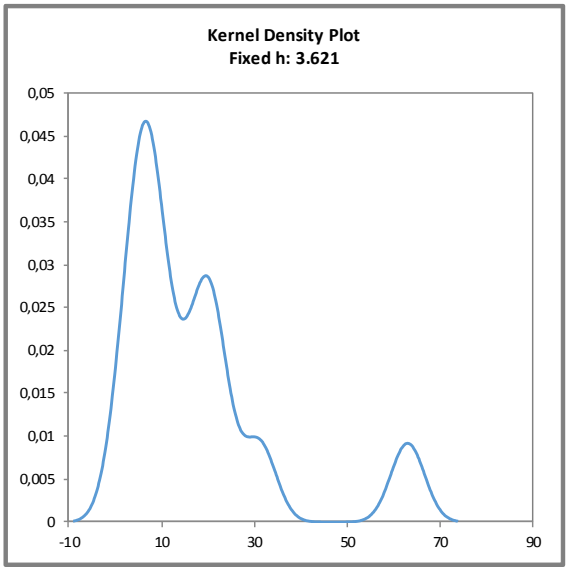

Fig. 2: Kernel Density Plot of all ELISA-results β-Lactoglobulin (with h = σpt of Xpt_{ALL})

Comments:

The kernel density estimation shows a multimodal distribution due to differences of the applied methods: 1. method RS, 2. method ES and IL and 3. method BK. A shoulder at 31 mg/kg is caused by an outlier of method RS (s. fig. 2).

Reprint, also in part, only with written permission from DLA-Ahrensburg

Characteristics: Quantitative evaluation β-Lactoglobulin

Sample B

Characteristics	**All Results** *[mg/kg]*	**Method RS** *[mg/kg]*
Assigned value (X_{pt})	$X_{pt_{ALL}}$	$X_{pt_{METHOD\ RS}}$
Number of results	*13*	*7*
Number of outliers	*-*	*1*
Median	11,4	6,90
Robust mean (X_{pt})	14,5	6,98
Robust standard deviation (S*)	11,6	4,05
Target data:		
Target standard deviation σ_{pt}		1,75
lower limit of target range (X_{pt} - 2σ_{pt})		3,49
upper limit of target range (X_{pt} + 2σ_{pt})		10,5
Quotient S/σ_{pt}*		*2,3*
Standard uncertainty $U_{(Xpt)}$		*1,91*
Quotient $U_{(Xpt)}/\sigma_{pt}$		*1,1*
Number of results in target range		*6*
Percent in target range		*86%*

Method:
RS = R-Biopharm, Ridascreen Fast®

Comments to the statistical characteristics:

The evaluation of all methods showed a multimodal distribution of results depending on testkit methods (see fig. 2). Therefore an evaluation of results across the methods was not performed.
The evaluation of results from method RS showed a slightly increased variability. The quotient S/σ_{pt} was above 2,0. The robust standard deviation is in the range of established values for the reproducibility standard deviation of the applied method (see 3.4.2 value by precision experiments and 3.4.3 value by perception). The comparability of results is given.*

The robust mean of the evaluation of method RS was 32% of the spiking level of β-lactoglobulin to sample B fulfilling the recommendations for the applied methods (s. 3.4.3 and "Recovery rates of β-Lactoglobulin" p. 21).

Reprint, also in part, only with written permission from DLA-Ahrensburg

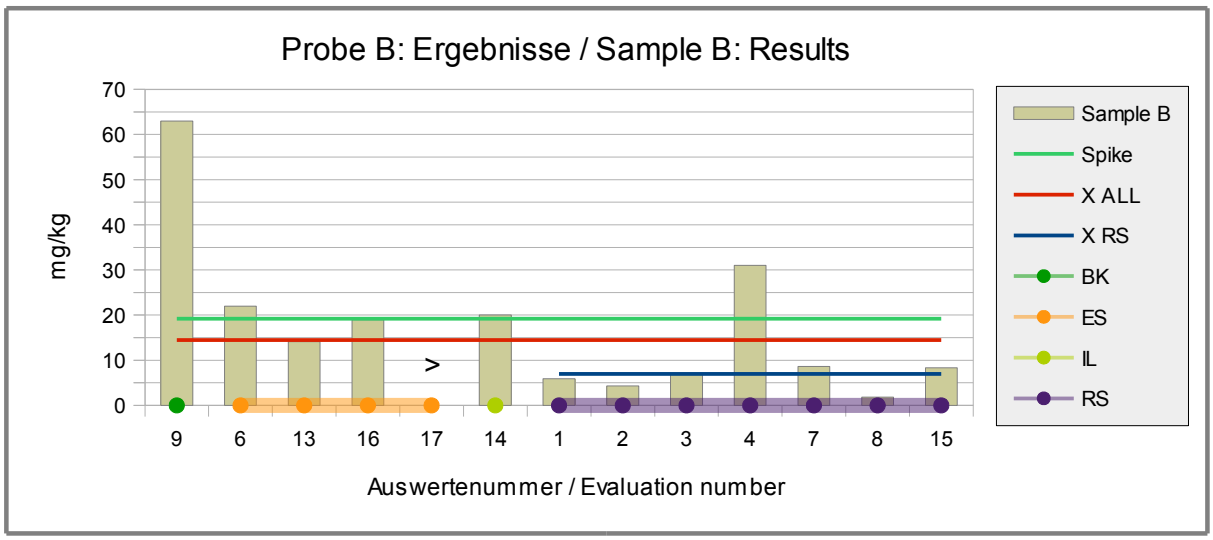

Fig. 3: ELISA-Results β-Lactoglobulin
green line = Spiking level
red line = Assigned value robust mean all results
blue line = Assigned value robust mean results method RS
round symbols = Applied methods (see legend)

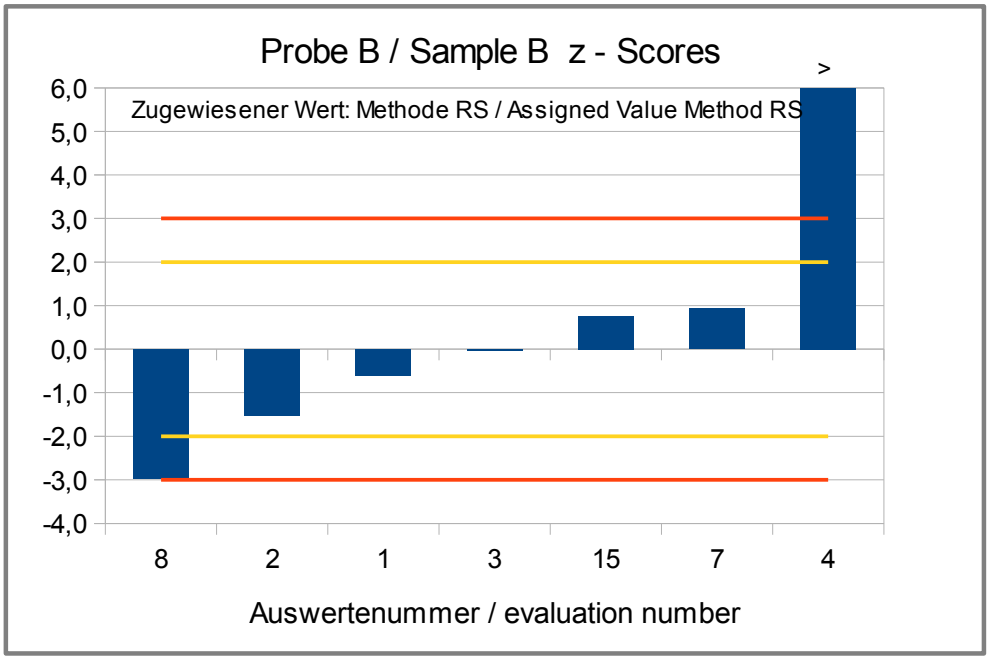

Fig. 4: z-Scores (ELISA-Results as β-Lactoglobulin)
Assigned value robust mean of method RS (R-Biopharm, Ridascreen)

Reprint, also in part, only with written permission from DLA-Ahrensburg

Recovery Rates for β-Lactoglobulin:
Spiking Material Sample and Sample B

Evaluation number	Spiking material	Recovery rate	Sample B	Recovery rate	Method	Remarks
	[mg/kg]	[%]	[mg/kg]	[%]		
9	1150	200	63	328	BK	
6	330	**57**	22	**115**	ES	
13	NA		14,1	**73**	ES	
16	410	**71**	19	**99**	ES	
17			>1		ES	
14	396	**69**	20	**104**	IL	
1	315	**55**	5,91	31	RS	
2			4,3	22	RS	
3	>13.5		6,9	36	RS	
4	1400	244	31	161	RS	outlier Xpt RS
7	280	49	8,6	45	RS	
8			1,8	9	RS	
15	472	**82**	8,3	43	RS	

RA*	50-150 %	RA*	50-150 %
Number in RA	**5**	Number in RA	**4**
Percent in RA	**63**	Percent in RA	**33**

* Range of acceptance of AOAC for allergen ELISAS

Recovery rate
100% relative size:
beta-Lactoglobulin, s. page 4

Methods:
BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

Comments:
For the spiking material sample 63% of the participants obtained a recovery rate within the range of the AOAC-recommendation of 50-150%. For the infant food-sample B produced with the spiking material sample 33% of the recovery rates were in the range of acceptance.

Reprint, also in part, only with written permission from DLA-Ahrensburg

4.1.2 ELISA-Results: Casein

Qualitative valuation of results: Samples A and B

Evaluation number	Sample A	Sample A	Sample B	Sample B	Qualitative Valuation	Method	Remarks
	pos/neg	[mg/kg]	pos/neg	[mg/kg]	Agreement with consensus value		
17	negative	<1	positive	>30	2/2 (100%)	4L	
2	negative	< 0.2	positive		2/2 (100%)	AQ	
9	positive	0,55	positive	313	1/2 (50%)	AQ	
12	negative	<1	positive	435,9	2/2 (100%)	AQ	
16	negative	<0,2	positive	450	2/2 (100%)	AQ	
10	negative	< 2,5	positive	475	2/2 (100%)	BK	
13	negative	<0,6	positive	64,6	2/2 (100%)	ES	
14	negative	< 0,1	positive	250	2/2 (100%)	IL	
1	negative	<2.5	positive	100,58	2/2 (100%)	RS	
3	negative		positive	55,7	2/2 (100%)	RS	
4	negative	<2,5	positive	36	2/2 (100%)	RS	
6	negative	<2,5	positive	155	2/2 (100%)	RS	
7	negative		positive	176,3	2/2 (100%)	RS	
8	negative	<0,5	positive	269,7	2/2 (100%)	RS	
11	negative		positive	254	2/2 (100%)	RS	
15	negative	<5	positive	353,6	2/2 (100%)	RS	

	Sample A		Sample B	
Number positive	1		16	
Number negative	15		0	
Percent positive	6		100	
Percent negative	94		0	
Consensus value	negative		positive	

Methods:

4L = 4LabDiagnostics
AQ = AgraQuant, RomerLabs
BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

Comments:

There were 94% negative results for sample A and 100% positive results for sample B by the ELISA-methods. The positive result for sample A was in the range of the limit of detection and quantification of the respective method.
The consensus values are in agreement with the spiking of sample B.

Reprint, also in part, only with written permission from DLA-Ahrensburg

Quantitative valuation of results: Sample B

Evaluation number	Casein	z-Score Xpt_{ALL}	z'-Score Xpt_{RS}	Method	Remarks
	[mg/kg]				
17	>30			4L	
2				AQ	
9	313			AQ	
12	435,9			AQ	
16	450			AQ	
10	475			BK	
13	64,6			ES	
14	250			IL	
1	100,58		-1,7	RS	
3	55,7		-2,7	RS	
4	36		-3,2	RS	
6	155		-0,5	RS	
7	176,3		0,0	RS	
8	269,7		2,2	RS	
11	254		1,8	RS	
15	353,6		4,1	RS	

Methods:

4L = 4LabDiagnostics
AQ = AgraQuant, RomerLabs
BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

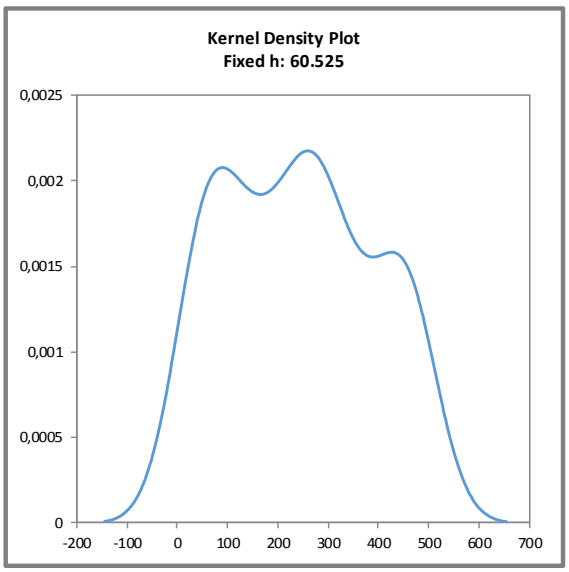

Fig. 5: Kernel Density Plot of all ELISA-results Casein
(with h = σpt of Xpt_{ALL})

Comments:
The kernel density estimation shows a broad distribution of results with modalities, which could not clearly attributed the applied methods: method AQ showed higher results, while the results of method RS were distributed over a broader range. There were only one quantitative result each from the methods BK, ES and IL (s. fig. 5).

Reprint, also in part, only with written permission from DLA-Ahrensburg

Characteristics: Quantitative evaluation Casein

Sample B

Characteristics	**All Results** *[mg/kg]*	**Method RS** *[mg/kg]*
Assigned value (X_{pt})	$X_{pt_{ALL}}$	$X_{pt_{METHOD\ RS}}$
Number of results	*14*	*8*
Number of outliers	*-*	*0*
Median	252	166
Robust mean (X_{pt})	242	175
Robust standard deviation (S*)	171	126
Target data:		
Target standard deviation σ_{pt}'		70,9
lower limit of target range (X_{pt} - $2\sigma_{pt}'$)		33,4
upper limit of target range (X_{pt} + $2\sigma_{pt}$)		317
Quotient S/σ_{pt}'*		*1,8*
Standard uncertainty $U_{(Xpt)}$		*55,7*
Quotient $U_{(Xpt)}/\sigma_{pt}'$		*0,79*
Number of results in target range		7
Percent in target range		*88%*

Method:
RS = R-Biopharm, Ridascreen Fast®

Comments to the statistical characteristics:

The evaluation of all methods showed a multimodal distribution of results depending on testkit methods (see fig. 2). Therefore an evaluation of results across the methods was not performed.
The evaluation of results from method RS showed an increased increased variability. The robust standard deviation is above the range of established values for the reproducibility standard deviation of the applied method (see 3.4.2 value by precision experiments and 3.4.3 value by perception). The quotient S/σ_{pt} was clearly above 2,0. Therefore evaluation was performed by z'-score considering the standard uncertainty. The resulting quotient S*/σ_{pt}' was below 2,0 (see 3.6 to 3.8).*
The robust mean of the evaluation of method RS was approximately 114% of the spiking level of casein to sample B fulfilling the recommendations for the applied methods (s. 3.4.3 and "Recovery rates of Casein" p. 26).

Reprint, also in part, only with written permission from DLA-Ahrensburg

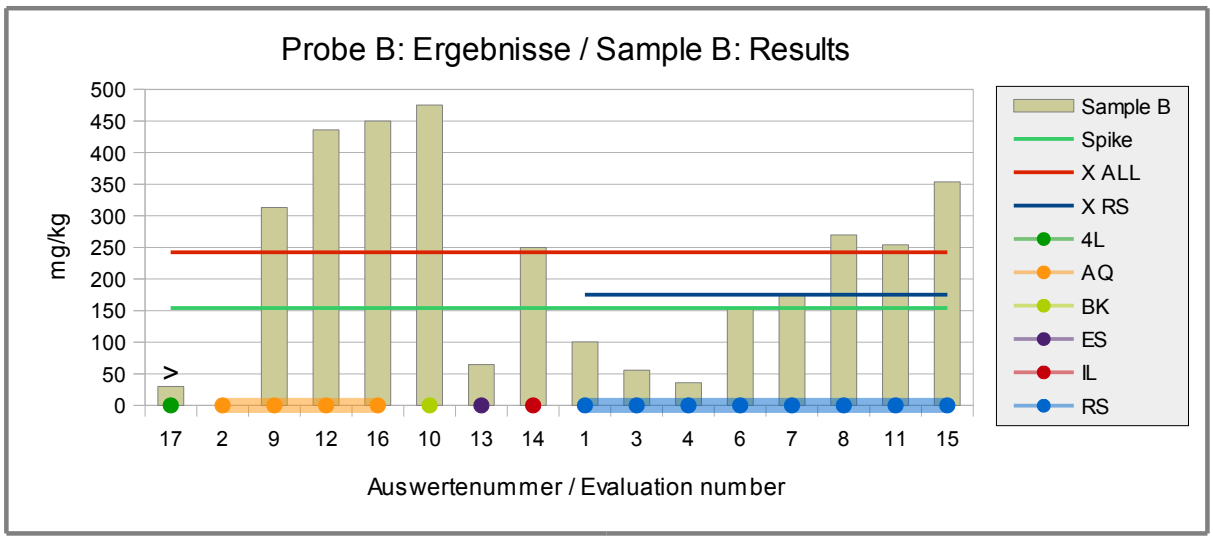

Fig. 6: ELISA-Results Casein
green line = Spiking level
red line = Assigned value robust mean all results
blue line = Assigned value robust mean results method RS
round symbols = Applied methods (see legend)

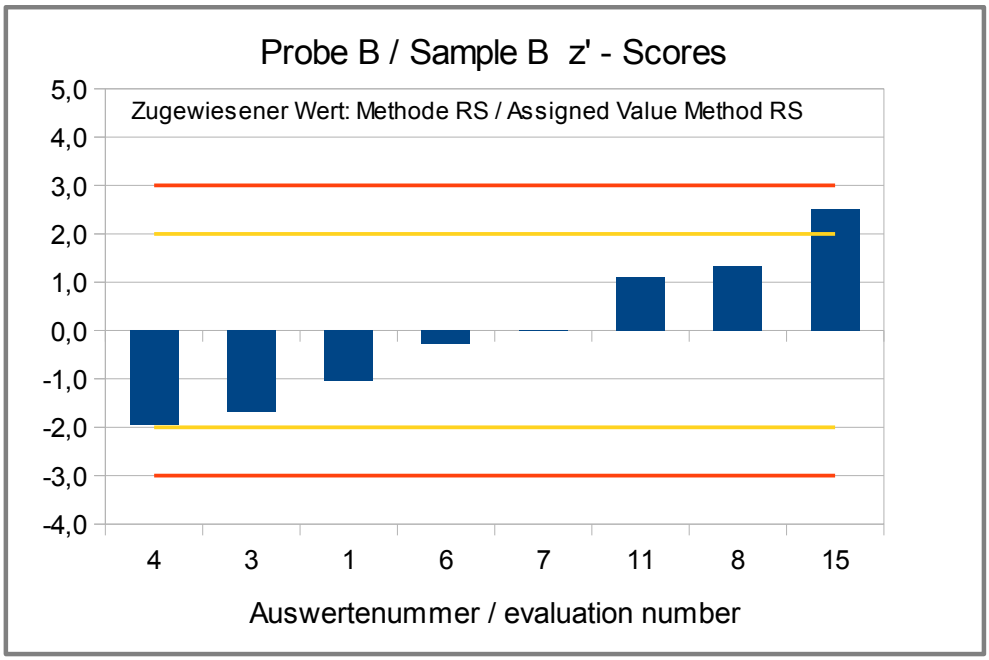

Fig. 7: z'-Scores (ELISA-Results as Casein)
Assigned value robust mean of method RS (R-Biopharm, Ridascreen)

Reprint, also in part, only with written permission from DLA-Ahrensburg

Recovery Rates for Casein:
Spiking Material Sample and Sample B

Evaluation number	Spiking material	Recovery rate	Sample B	Recovery rate	Method	Remarks
	[mg/kg]	[%]	[mg/kg]	[%]		
17			>30		4L	
2					AQ	
9	8873	193	313	203	AQ	
12	>3750		435,9	283	AQ	
16	13000	283	450	292	AQ	
10	19000	414	475	308	BK	
13	NA		64,6	42	ES	
14	5900	**129**	250	162	IL	
1	4194	**91**	100,58	**65**	RS	
3	>67.5		55,7	36	RS	
4	4300	**94**	36	23	RS	
6	8500	185	155	**101**	RS	
7	8324	181	176,3	**114**	RS	
8			269,7	175	RS	
11	6460	**141**	254	165	RS	
15	8462,5	184	353,6	230	RS	

RA*	50-150 %	RA*	50-150 %
Number in RA	**4**	Number in RA	**3**
Percent in RA	**40**	Percent in RA	**21**

* Range of acceptance of AOAC for allergen ELISAS

Recovery rate
100% relative size:
Casein, s. page 4

Methods:

BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

Comments:
For the spiking material sample 40% (4) of the participants obtained a recovery rate within the range of the AOAC-recommendation of 50-150%. For the infant food-sample B produced with the spiking material sample 21% (3) of the recovery rates were in the range of acceptance.

Reprint, also in part, only with written permission from DLA-Ahrensburg

4.2 Proficiency Test Gluten

4.2.1 ELISA-Results: Gluten

Qualitative valuation of results: Samples A and B

Methods:

Evaluation number	Sample A	Sample A	Sample B	Sample B	Qualitative Valuation	Method	Remarks
	pos/neg	[mg/kg]	pos/neg	[mg/kg]	Agreement with consensus value		
7	positive	19,2	positive	78,8	1/2 (50%)	AQ	
14	negative	< 1	positive	100	2/2 (100%)	IL	Result converted *
1	negative	<5	positive	43,77	2/2 (100%)	RS	
2	negative	< 5	positive	45,2	2/2 (100%)	RS	
3	negative		positive	40,8	2/2 (100%)	RS	
4	negative	<5	positive	33	2/2 (100%)	RS	
6	negative	<7	positive	35	2/2 (100%)	RS	
8	negative	<5	positive	69	2/2 (100%)	RS	
9	negative	<5	positive	17	2/2 (100%)	RS	
12	negative	<5	positive	26,22	2/2 (100%)	RS	
13	negative	<5	positive	38,1	2/2 (100%)	RS	
15	negative	<0,25	positive	44,1	2/2 (100%)	RS	
16	negative	<5	positive	41	2/2 (100%)	RS	
17	negative	<3	positive	21,9	2/2 (100%)	RS	
10	negative	< 5	positive	61,4	2/2 (100%)	VT	
16	negative	<10	positive	67	2/2 (100%)	VT	

* calculation see p. 15

	Sample A		Sample B	
Number positive	1		16	
Number negative	15		0	
Percent positive	6		100	
Percent negative	94		0	
Consensus value	negative		positive	

AQ = AgraQuant, RomerLabs
IL = Immunolab
RS = Ridascreen®, R-Biopharm
VT = Veratox, Neogen

Comments:

There were 94% negative results for sample A and 100% positive results for sample B by the ELISA-methods. One positive result for sample A was obtained by method AQ (AgraQuant G12). According to the manufacturer the testkit detects prolamines from oat (in contrast to the other methods). The two PCR results of method Sure Food Gluten, which detects oat as well according to the testkit manual, are positive and negative for sample A. The presence of oat traces in the sample could not be excluded.

The consensus values are in qualitative agreement with the spiking of sample B.

Reprint, also in part, only with written permission from DLA-Ahrensburg

Quantitative valuation of results: Sample B

Evaluation number	Gluten	z-Score Xpt_{ALL}	z-Score Xpt_{RS}	Method	Remarks
	[mg/kg]				
7	78,8	2,8		AQ	
14	100	4,6		IL	Result converted *
1	43,77	-0,2	0,7	RS	
2	45,2	-0,1	0,9	RS	
3	40,8	-0,5	0,4	RS	
4	33	-1,2	-0,4	RS	
6	35	-1,0	-0,2	RS	
8	69	2,0	3,5	RS	
9	17	-2,5	-2,2	RS	
12	26,22	-1,7	-1,2	RS	
13	38,1	-0,7	0,1	RS	
15	44,1	-0,2	0,8	RS	
16	41	-0,5	0,4	RS	
17	21,9	-2,1	-1,6	RS	
10	61,4	1,3		VT	
16	67	1,8		VT	

* calculation see p. 15

Methods:

AQ = AgraQuant, RomerLabs
IL = Immunolab
RS = Ridascreen®, R-Biopharm
VT = Veratox, Neogen

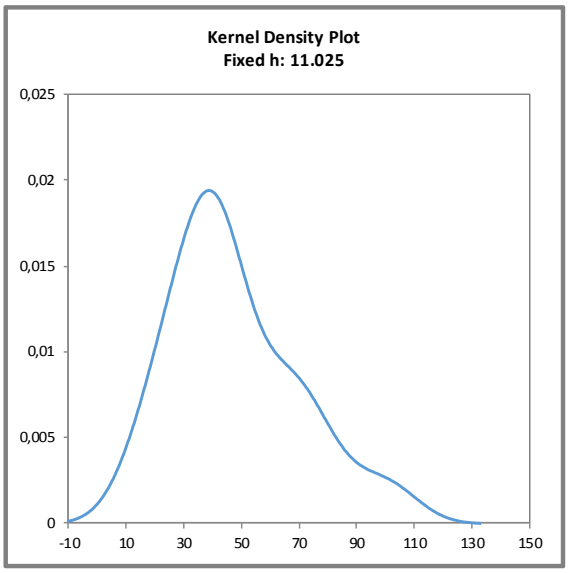

Fig. 8: Kernel Density Plot of all ELISA-results gluten
(with h = σpt of Xpt_{ALL})

Comment:
The kernel density estimation shows a normal distribution of results with two slight shoulders at 60-80 mg/kg (methods AQ a. VT) and 100 mg/kg (method IL) (s. fig. 8).

Reprint, also in part, only with written permission from DLA-Ahrensburg

Characteristics: Quantitative evaluation Gluten

Sample B

Characteristics	**All Results** *[mg/kg]*	**Method RS** *[mg/kg]*
Assigned value (X_{pt})	$X_{pt_{ALL}}$	$X_{pt_{METHOD\ RS}}$
Number of results	*16*	*12*
Number of outliers	*0*	*0*
Median	42,4	39,5
Robust mean (X_{pt})	46,3	36,9
Robust standard deviation (S*)	21,9	11,7
Target data:		
Target standard deviation σ_{pt}	11,6	9,23
lower limit of target range ($X_{pt} - 2\sigma_{pt}$)	23,2	18,5
upper limit of target range ($X_{pt} + 2\sigma_{pt}$)	69,5	55,4
Quotient S/σ_{pt}*	*1,9*	*1,3*
Standard uncertainty $U_{(Xpt)}$	*6,86*	*4,24*
Quotient $U_{(Xpt)}/\sigma_{pt}$	*0,59*	*0,46*
Number of results in target range	*11*	*11*
Percent in target range	*69%*	*92%*

Method:
RS = R-Biopharm, Ridascreen®

Comments to the statistical characteristics:

The evaluation of all methods and the evaluation of results from method RS showed a normal to low variability of results, respectively. The quotients S/σ_{pt} were below 2,0. The robust standard deviation is in the range of established values for the reproducibility standard deviation of the applied methods (see 3.4.2 value by precision experiments and 3.4.3 value by perception). The comparability of results is given. This conclusion is limited for the evaluation across the methods, because there are only a few results for the methods AQ, IL and VT. All of these results were above the robust mean of all methods.*

The robust means of the evaluations were 123% and 98% of the spiking level of gluten to sample B and within the recommendations for the applied methods (s. 3.4.3 and "Recovery rates of Gluten" p.32).

Reprint, also in part, only with written permission from DLA-Ahrensburg

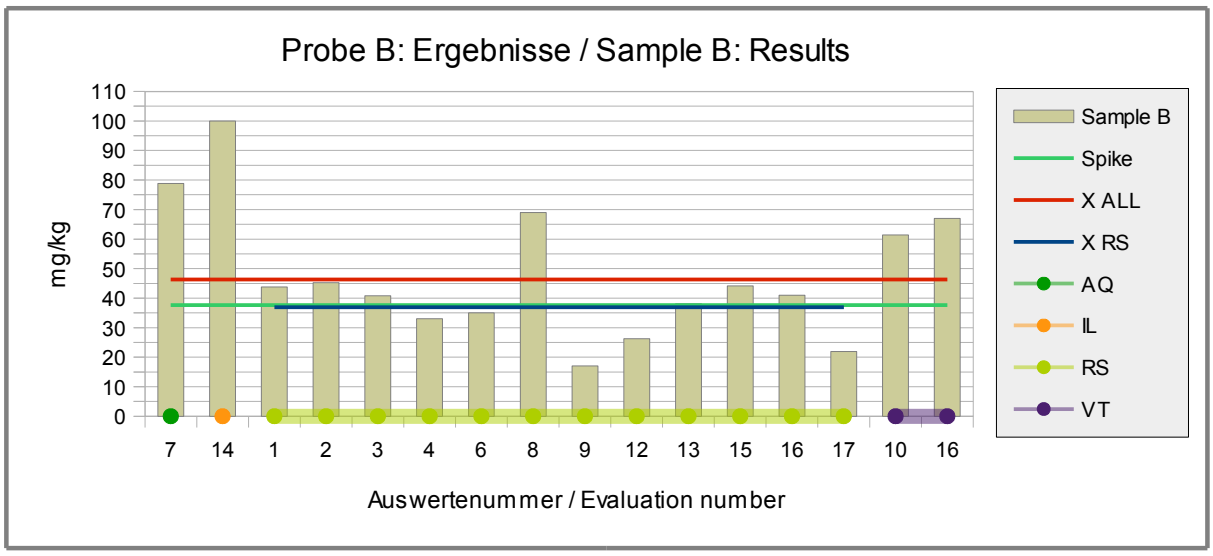

Fig. 9: ELISA-Results Gluten
green line = Spiking level
red line = Assigned value robust mean all results
blue line = Assigned value robust mean results method RS
round symbols = Applied methods (see legend)

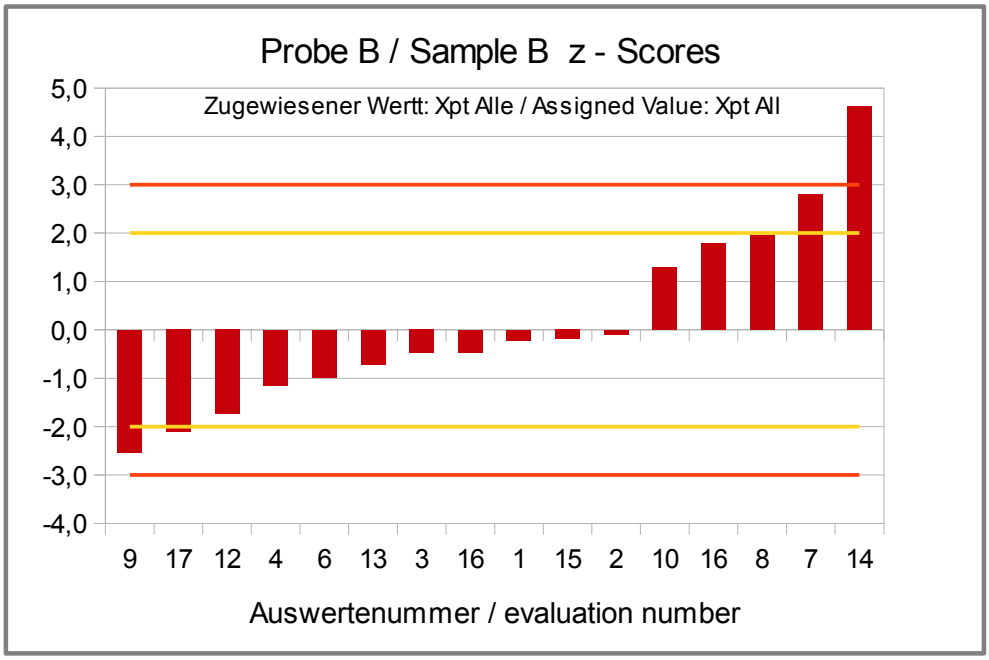

Fig. 10: z-Scores (ELISA-Results as Gluten)
Assigned value robust mean of all results

Reprint, also in part, only with written permission from DLA-Ahrensburg

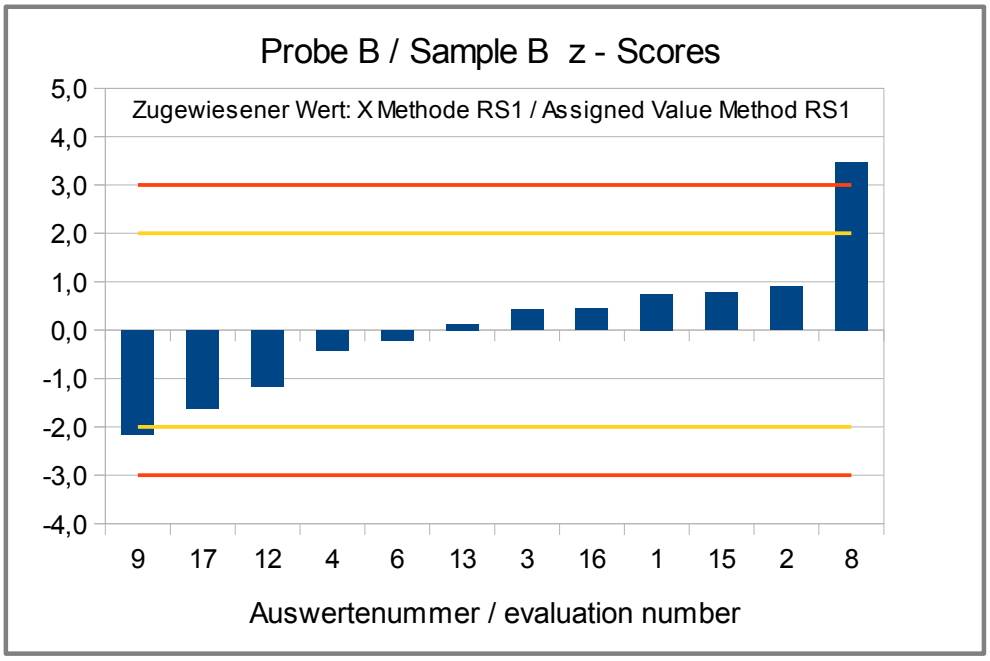

Fig. 11: z-Scores (ELISA-Results as Gluten)
Assigned value robust mean of method RS
(R-Biopharm, Ridascreen)

Reprint, also in part, only with written permission from DLA-Ahrensburg

Recovery Rates for Gluten:
Spiking Material Sample and Sample B

Evaluation number	Spiking material	Recovery rate	Sample B	Recovery rate	Method	Remarks
	[mg/kg]	[%]	[mg/kg]	[%]		
7	1035	**63**	78,8	210	AQ	
14	2940	178	100	266	IL	Result converted *
1	934	**57**	43,77	**116**	RS	
2			45,2	**120**	RS	
3	>320		40,8	**109**	RS	
4	11021	668	33	**88**	RS	
6	785	48	35	**93**	RS	
8			69	184	RS	
9	375	23	17	45	RS	
12	735,34	45	26,22	**70**	RS	
13	NA		38,1	**101**	RS	
15	1501,3	**91**	44,1	**117**	RS	
16	750	45	41	**109**	RS	
17			21,9	**58**	RS	
10			61,4	163	VT	
16	-		67	178	VT	

* calculation see p. 15

RA*	50-150 %	RA*	50-150 %
Number in RA	**3**	Number in RA	**10**
Percent in RA	**33**	Percent in RA	**63**

* Range of acceptance of AOAC for allergen ELISAS

Recovery rate
100% relative size:
Gluten, s. page 4

Methods:

AQ = AgraQuant, RomerLabs
IL = Immunolab
RS = Ridascreen®, R-Biopharm
VT = Veratox, Neogen

Comments:
For the spiking material sample 33% (3) of the participants obtained a recovery rate within the range of the AOAC-recommendation of 50-150%. For the infant food-sample B produced with the spiking material sample 63% of the recovery rates were in the range of acceptance.

Reprint, also in part, only with written permission from DLA-Ahrensburg

4.2.2 PCR-Results: Gluten-containing Cereals

Evaluation number	Result Sample A	Result Sample A	Result Sample B	Result Sample B	Qualitative Valuation	Method	Remarks
	pos / neg	mg/kg	pos / neg	mg/kg	Agreement with Consensus Value		
3	positive		positive		-	SFA	
17	negative		positive		-	SFA	
16	negative		positive		-	div	

Methods:

SFA = Sure Food Allergen, R-Biopharm / Congen

div = not indicated / other method

Comments:

There were 3 positive results for sample B for gluten-containing cereal-DNA by the PCR-methods. The results are in agreement with the spiking of sample B.

For sample A there were one positive and two negative results. Among the ELISA results was one positive result as well (see p.27).

Reprint, also in part, only with written permission from DLA-Ahrensburg

5. Documentation

Details by the participants

5.1 Details by the participants

5.1.1 ELISA: β-Lactoglobulin

Primary data

Evaluation number	Result Sample A		Result Sample B		Result Spiking Sample		quantitative Result given as	Meth. Abr.	Method
	qualitative	mg/kg	qualitative	mg/kg	qualitative	mg/kg	e.g. food / food protein		Test-Kit + Manufacturer
9	negative	<5	positive	63	positive	1150	beta-Lactoglobulin	BK	BioKits β-Lactoglobulin Assay Kit, Neogen
6	negative	<0,1	positive	22	positive	330	beta-Lactoglobulin	ES	ELISA-Systems β-Lactoglobulin Residue Detection ELISA
13	negative	<0,2	positive	14,1	NA	NA	beta-Lactoglobulin	ES	ELISA-Systems β-Lactoglobulin Residue Detection ELISA
16	negative	<0,1	positive	19	positive	410	beta-Lactoglobulin	ES	ELISA-Systems β-Lactoglobulin Residue Detection ELISA
17	negative	<0,05	positive	>1	-		Given as	ES	ELISA-Systems β-Lactoglobulin Residue Detection ELISA
14	negative	< 0,01	positive	20	positive	396	beta-Lactoglobulin	IL	Immunolab β-Lactoglobulin ELISA
1	negative	<0.5	positive	5,91	positive	315	beta-Lactoglobulin	RS	Ridascreen Fast β-Lactoglobulin (R4902), r-Biopharm
2	negative	< 0.5	positive	4,3	-		beta-Lactoglobulin	RS	Ridascreen Fast β-Lactoglobulin (R4902), r-Biopharm
3	negative		positive	6,9	positive	>13.5	beta-Lactoglobulin	RS	Ridascreen Fast β-Lactoglobulin (R4902), r-Biopharm
4	negative	<0,5	positive	31	positive	1400	beta-Lactoglobulin	RS	Ridascreen Fast β-Lactoglobulin (R4902), r-Biopharm
7	negative		positive	8,6	positive	280	beta-Lactoglobulin	RS	Ridascreen Fast β-Lactoglobulin (R4902), r-Biopharm
8	-	<1	-	1,8	-		beta-Lactoglobulin	RS	Ridascreen Fast β-Lactoglobulin (R4902), r-Biopharm
15	-	<5	-	8,3	-	472	ß-Lactoglobulin	RS	r-biopharm, RIDASCREEN®FAST ß-Lactoglobulin (R4902)

Methods:

BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

Reprint, also in part, only with written permission from DLA-Ahrensburg

Other details to the Methods

Evaluation number	Meth. Abk.	Specifity	Remarks to the Method (Extraction and Determination)	Further Remarks
		Antibody	e.g. Extraction Solution / Time / Temperature	
9	BK			
6	ES			
13	ES	β-Lactoglobulin	Manufacturer's extraction solution / 15 min / 60C	
16	ES			
17	ES			
14	IL			
1	RS	As stated in the kit	As per kit instructions	Sample solidified upon addition of primary extraction solution. To produce a liquid mix, half the amount of secondary extraction solution w as added 1 stage early and the remaining amount of solution w as added at the seconday stage.
2	RS			
3	RS			
4	RS			
7	RS			
8	RS	Monoclonal específico de beta-lactoglobulina	Solución de Extracción 2/10min/100°C +A-AEP/10min/60°C	
15	RS		As per kit instructions	

Reprint, also in part, only with written permission from DLA-Ahrensburg

5.1.2 ELISA: Casein

Primary data

Evaluation number	Result Sample A		Result Sample B		Result Spiking Sample		quantitative Result given as	Meth. Abr.	Method
	qualitative	mg/kg	qualitative	mg/kg	qualitative	mg/kg	e.g. food / food protein		Test-Kit + Manufacturer
17	negative	<1	positive	>30	-		Ergebnis als	4L	Kit 4LAB MILK ALERT II
2	negative	< 0.2	positive		-		Casein	AQ	AgraQuant Casein (CO-KAL1200), RomerLabs
9	positive	0,55	positive	313	positive	8873	Casein	AQ	AgraQuant Casein (CO-KAL1200), RomerLabs
12	negative	<1	positive	435,9	positive	>3750	Casein	AQ	AgraQuant Casein (CO-KAL1200), RomerLabs
16	negative	<0,2	positive	450	positive	13000	Casein	AQ	AgraQuant Casein (CO-KAL1200), RomerLabs
10	negative	< 2,5 ppm	positive	475 ppm	positive	1,90%	Casein	BK	BioKits Casein Assay Kit, Neogen
13	negative	<0,6	positive	64,6	NA	NA	Casein	ES	ELISA-Systems Casein Residue Detection ELISA
14	negative	< 0,1	positive	250	positive	5900	Casein	IL	Immunolab Casein ELISA
1	negative	<2.5	positive	100,58	positive	4194	Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
3	negative		positive	55,7	positive	>67.5	Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
4	negative	<2,5	positive	36	positive	4300	Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
6	negative	<2,5	positive	155	positive	8500	Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
7	negative		positive	176,3	positive	8324	Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
8	-	<0,5	-	269,7	-		Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
11	negative		-	254	-	6460	Casein	RS	Ridascreen Fast Casein (R4612), r-Biopharm
15	-	<5	-	353,6	-	8462,5	Casein	RS	r-biopharm, RIDASCREEN®FAST Casein (R4612)

Methods:

4L = 4LabDiagnostics
AQ = AgraQuant, RomerLabs
BK = BioKits, Neogen
ES = ELISA Systems
IL = Immunolab
RS = Ridascreen®, R-Biopharm

Reprint, also in part, only with written permission from DLA-Ahrensburg

Other details to the methods

Evaluation number	Meth. Abr.	Specifity	Remarks to the Method (Extraction and Determination)	Further Remarks
		Antibody	e.g. Extraction Solution / Time / Temperature	
17	4L			
2	AQ			
9	AQ			
12	AQ	Casein	Extraction solution / 15 min / 60°C	Spiking sample: 1:4000; sample A/B: 1:500;
16	AQ			
10	BK	Casein		applied Veratox Casein Allergen
13	ES	Casein	Manufacturer's extraction solution / 15 min / 60C	
14	IL			
1	RS	As stated in the kit	As per kit instructions	Sample solidified upon addition of primary extraction solution. To produce a liquid mix, half the amount of secondary extraction solution w as added 1 stage early and the remaining amount of solution w as added at the seconday stage.
3	RS			
4	RS			
6	RS			
7	RS		sample preparation for infant food (w ith AEP), LOQ 0,5ppm	
8	RS	specific for casein	Allergen extraction buffer diluted/10min/60°C	
11	RS	Casein		
15	RS		As per kit instructions	

Reprint, also in part, only with written permission from DLA-Ahrensburg

5.1.3 ELISA: Gluten

Primary data

Evaluation number	Result Sample A		Result Sample B		Result Spiking Sample		quantitative Result given as	Meth. Abr.	Method
	qualitative	mg/kg	qualitative	mg/kg	qualitative	mg/kg	e.g. food / food protein		Test-Kit + Manufacturer
7	positive	19,2	positive	78,8	positive	1035	Gluten	AQ	AgraQuant Gluten G12, RomerLabs
14	negative	< 1	positive	50	positive	1470	Gliadin	IL	Immunolab Gliadin GLU-E02
1	negative	<5	positive	43,77	positive	934	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
2	negative	< 5	positive	45,2	positive		Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
3	negative		positive	40,8	positive	>320	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
4	negative	<5	positive	33	positive	11021	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
6	negative	<7	positive	35	positive	785	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
8	-	<5	-	69	-		Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
9	negative	<5	positive	17	positive	375	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
12	negative	<5	positive	26,22	positive	735,34	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
13	negative	<5	positive	38,1	NA	NA	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
15	-	<0,25	-	44,1	-	1501,3	Gluten	RS	r-biopharm, RIDASCREEN®FAST Gliadin (R7001)
16	negative	<5	positive	41	positive	750	Gluten	RS	Ridascreen Gluten (R7001), r-Biopharm
17	negative	<3	positive	21,9	-		Given as	RS	Ridascreen Gluten (R7001), r-Biopharm
10	negative	< 5 ppm	positive	61,4 ppm	positive		Gluten	VT	Veratox Gliadin R5, Neogen
16	negative	<10	positive	67	-	-	Gluten	VT	Veratox Gliadin, Neogen

Methods:

AQ = AgraQuant, RomerLabs
IL = Immunolab
RS = Ridascreen®, R-Biopharm
VT = Veratox, Neogen

Reprint, also in part, only with written permission from DLA-Ahrensburg

Other details to the methods

Evaluation number	Meth. Abr.	Specifity	Remarks to the Method (Extraction and Determination)	Further Remarks
		Antibody	e.g. Extraction Solution / Time / Temperature	
7	AQ			
14	IL			
1	RS	As stated in the kit	As per kit instructions	Sample solidified upon addition of primary extraction solution. To produce a liquid mix, half the amount of secondary extraction solution w as added 1 stage early and the remaining amount of solution w as added at the se-conday stage.
2	RS		Cocktail solution	
3	RS			
4	RS	R5		
6	RS			
8	RS	R5-Mendez	Cocktail solution/40min/50°C+ Ethanol80%/1 hour/room temperature	
9	RS			
12	RS	Gliadin	Cocktail-Lösung Art. Nr. R7006 / 60 min / 25°C	Spiking sample: 1:4000; sample A/B: 1:500;
13	RS	Gluten (R5)	Cocktail solution / 40 min / 50C	
15	RS		As per kit instructions	
16	RS			
17	RS			
10	VT	Gliadin		
16	VT			

Reprint, also in part, only with written permission from DLA-Ahrensburg

5.1.4 PCR: Gluten-containing Cereals

Primary data

Evaluation number	Result Sample A		Result Sample B		Result Spiking Sample		quantitative Result given as	Meth. Abr.	Method
	qualitative	mg/kg	qualitative	mg/kg	qualitative	mg/kg	e.g. food / food protein		Test-Kit + Manufacturer
3	positive		positive		positive		Gluten	SFA	Sure Food Allergen, Congen / r-Biopharm
17	negative		positive					SFA	Sure Food Allergen, Congen / r-Biopharm
16	negative	-	positive	-	positive	-	Wheat-DNA, Rye-DNA, Barley-DNA	div	internal method

Method:

SFA = Sure Food Allergen, R-Biopharm / Congen

div = not indicated / other method

Other Remarks to the Methods

Evaluation number	Meth. Abr.	Specifity	Remarks to the Method (Extraction and Determination)	Further Remarks
		Antibody	e.g. Extraction / Enzymes / Clean-Up / Real Time PCR / Gel electrophoresis / Cycles	
3	SFA			
17	SFA			
16	div		CTAB / Protease K / Chloroform + Promega Wizard/ End Point PCR/ 4% Agarose gel / 45 Cycles	

Reprint, also in part, only with written permission from DLA-Ahrensburg

5.2 Homogeneity

5.2.1 Mixture homogeneity before bottling

Microtracer Homogeneity Test
Spiking Material Sample

Weight whole sample	1,21	kg
Microtracer	FSS-rot lake	
Particle size	75 – 300	µm
Weight per particle	2,0	µg
Addition of tracer	35,8	mg/kg

Result of analysis

Sample	Weight [g]	Particle number	Particles [mg/kg]
1	6,41	123	38,4
2	5,33	101	37,9
3	5,99	92	30,7
4	6,16	97	31,5
5	5,96	92	30,9
6	5,55	106	38,2
7	5,77	102	35,4
8	5,80	108	37,2

Poisson distribution		
Number of samples	8	
Degree of freedom	7	
Mean	102,8	Partikel
Standard deviation	10,1	Partikel
χ^2 (CHI-Quadrat)	6,95	
Probability	**43**	%
Recovery rate	98	%

Normal distribution		
Number of samples	8	
Mean	35,0	mg/kg
Standard deviation	3,44	mg/kg
rel. Standard deviaton	9,82	%
Horwitz standard deviation	9,37	%
HorRat-value	**1,0**	
Recovery rate	98	%

Microtracer Homogeneity Test
DLA 03-2016 Sample B

Weight whole sample	2,07	kg
Microtracer	FSS-rot lake	
Particle size	75 – 300	µm
Weight per particle	2,0	µg
Addition of tracer	19,6	mg/kg

Result of analysis

Sample	Weight [g]	Particle number	Particles [mg/kg]
1	6,6	67	20,5
2	5,9	55	18,7
3	5,7	45	15,8
4	6,8	49	14,5
5	6,1	53	17,4
6	6,8	55	16,1
7	6,2	59	19,1
8	5,5	39	14,2

Poisson distribution		
Number of samples	8	
Degree of freedom	7	
Mean	52,7	Partikel
Standard deviation	6,97	Partikel
χ^2 (CHI-Quadrat)	6,46	
Probability	**49**	%
Recovery rate	87	%

Normal distribution		
Number of samples	8	
Mean	17,0	mg/kg
Standard deviation	2,25	mg/kg
rel. Standard deviaton	13,2	%
Horwitz standard deviation	10,4	%
HorRat-value	**1,3**	
Recovery rate	87	%

Reprint, also in part, only with written permission from DLA-Ahrensburg

6. Index of participant laboratories

Teilnehmer / Participant	Ort / Town	Land / Country
		SPAIN
		FRANCE
		CANADA
		ITALY
		Germany
		Germany
		Germany
		Germany
		SWEDEN
		Germany
		SPAIN
		BELGIUM
		GREAT BRITAIN
		Germany
		GREAT BRITAIN
		Germany
		Germany

[Die Adressdaten der Teilnehmer wurden für die allgemeine Veröffentlichung des Auswerte-Berichts nicht angegeben.]

[The address data of the participants were deleted for publication of the evaluation report.]

Reprint, also in part, only with written permission from DLA-Ahrensburg

7. Index of references

1. DIN EN ISO/IEC 17025:2005; Allgemeine Anforderungen an die Kompetenz von Prüf- und Kalibrierlaboratorien / General requirements for the competence of testing and calibration laboratories
2. DIN EN ISO/IEC 17043:2010; Konformitätsbewertung – Allgemeine Anforderungen an Eignungsprüfungen / Conformity assessment – General requirements for proficiency testing
3. ISO 13528:2015 & DIN ISO 13528:2009; Statistische Verfahren für Eignungsprüfungen durch Ringversuche / Statistical methods for use in proficiency testing by interlaboratory comparisons
4. ASU §64 LFGB: Planung und statistische Auswertung von Ringversuchen zur Methodenvalidierung / DIN ISO 5725 series part 1, 2 and 6 Accuracy (trueness and precision) of measurement methods and results
5. Verordnung / Regulation 882/2004/EU; Verordnung über über amtliche Kontrollen zur Überprüfung der Einhaltung des Lebensmittel- und Futtermittelrechts sowie der Bestimmungen über Tiergesundheit und Tierschutz / Regulation on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules
6. Evaluation of analytical methods used for regulation of food and drugs; W. Horwitz; Analytical Chemistry, 54, 67-76 (1982)
7. The International Harmonised Protocol for the Proficiency Testing of Anan-lytical Laboratories ; J.AOAC Int., 76(4), 926 – 940 (1993)
8. A Horwitz-like funktion describes precision in proficiency test; M. Thompson, P.J. Lowthian; Analyst, 120, 271-272 (1995)
9. Protocol for the design, conduct and interpretation of method performance studies; W. Horwitz; Pure & Applied Chemistry, 67, 331-343 (1995)
10. Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing; M. Thompson; Analyst, 125, 385-386 (2000)
11. The International Harmonised Protocol for the Proficiency Testing of Analytical Chemistry Laboratories; Pure Appl Chem, 78, 145 – 196 (2006)
12. AMC Kernel Density - Representing data distributions with kernel density estimates, amc technical brief, Editor M Thompson, Analytical Methods Committee, AMCTB No 4, Revised March 2006 and Excel Add-in Kernel.xla 1.0e by Royal Society of Chemistry
13. EURACHEM/CITAC Leitfaden, Ermittlung der Messunsicherheit bei analytischen Messungen (2003); Quantifying Uncertainty in Analytical Measurement (1999)
14. GMP+ Feed Certification scheme, Module: Feed Safety Assurance, chapter 5.7 Checking procedure for the process accuracy of compound feed with micro tracers in GMP+ BA2 Control of residues, Version: 1st of January 2015 GMP+ International B.V.
15. MTSE SOP No. 010.01 (2014): Quantitative measurement of mixing uniformity and carry-over in powder mixtures with the rotary detector technique, MTSE Micro Tracers Services Europe GmbH
16. Codex Alimentarius Commission (2010) - Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific protiens in foods, CAC/GL 74-2010
17. DIN EN ISO 15633-1:2009; Nachweis von Lebensmittelallergenen mit immunologischen Verfahren - Teil 1: Allgemeine Betrachtungen / Foodstuffs - Detection of food allergens by immunological methods - Part 1: General considerations
18. DIN EN ISO 15634-1:2009; Nachweis von Lebensmittelallergenen mit molekularbiologischen Verfahren - Teil 1: Allgemeine Betrachtungen / Foodstuffs - Detection of food allergens by molecular biological methods - Part 1: General considerations
19. DIN EN ISO 15842:2010 Lebensmittel – Nachweis von Lebensmittelallergenen – Allgemeine Betrachtungen und Validierung von Verfahren / Foodstuffs - Detection of food allergens - General considerations and validation of

Reprint, also in part, only with written permission from DLA-Ahrensburg

methods
20. Ministry of Health and Welfare, JSM, Japan 2006
21. Working Group Food Allergens, Abbott et al., Validation Procedures for Quantitative Food Allergen ELISA Methods: Community Guidance and Best Practices JAOAC Int. 93:442-50 (2010)
22. Working Group on Prolamin Analysis and Toxicity (WGPAT): Méndez et al. Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. Eur J Gastroenterol Hepatol. 17:1053-63 (2005)
23. DLA Publikation: Performance of ELISA and PCR methods for the determination of allergens in food: an evaluation of six years of proficiency testing for soy (Glycine max L.) and wheat gluten (Triticum aestivum L.); Scharf et al.; J Agric Food Chem. 61(43):10261-72 (2013)
24. EFSA (2014) Scientific Opinion on the evaluation of allergenic foods and food ingredients for labelling purposes1, EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), European Food Safety Authority (EFSA), Parma, Italy, EFSA Journal 2014;12(11):3894
25. IRMM, Poms et al.; Inter-laboratory validation study of five different commercial ELISA test kits for determination of peanut residues in cookie and dark chocolate; European Commission, Joint Research Centre, Belgium; GE/R/FSQ/D08/05/2004
26. Jayasena et al. (2015) Comparison of six commercial ELISA kits for their specificity and sensitivity in detecting different major peanut allergens. J Agric Food Chem. 2015 Feb 18;63(6):1849-55
27. ASU §64 LFGB L 06.00-56 Bestimmung von Sojaprotein in Fleisch und Fleischerzeugnissen Enzymimmunologisches Verfahren (2007)
28. ASU §64 LFGB L 00.00-69 Bestimmung von Erdnuss-Kontaminationen in Lebensmitteln mittels ELISA im Mikrotiterplattensystem (2003)
29. ASU §64 LFGB L 44.00-7 Bestimmung von Haselnuss-Kontaminationen in Schokolade und Schokoladenwaren mittels ELISA im Mikrotiterplattensystem (2006)
30. Štumr et al.; Enzyme-Linked Immunosorbent Assay Kit for Beta-Lactoglobulin Determination: Interlaboratory Study, JAOAC 92:1519-25 (2009)
31. Štumr et al.; ELISA Kit for Casein Determination: Interlaboratory Study, JAOAC 93:676-682 (2010)

Reprint, also in part, only with written permission from DLA-Ahrensburg