Bundesverband
Naturkost Naturwaren

BNN Laboratory Approval System

Performance Test with "Undercover Samples"

Analysis of Pesticides, Nicotine, and Perchlorate in Matcha Tea on a Routine Level

Report

March 2017

Bundesverband
Naturkost Naturwaren

Summary

The laboratory performance assessment was designed and organised by Lach & Bruns in cooperation with PROOF-ACS in March 2017 on behalf of BNN e.V. (Bundesverband Naturkost Naturwaren).

BNN decided to make use of a so-called “undercover test”. The aim of the undercover test is to check the analytical competence of the BNN approved laboratories on a routine level. For that purpose, a sample of Matcha tea was delivered to each of the laboratories by commercial clients of the labs. The test samples were made up like routine samples and were accompanied by common submission forms of the clients. As a consequence, the labs were not aware of the undercover test. The challenge in delivering correct results is significantly increased in undercover tests compared to announced ring tests. No “special care analysis” is applied for undercover samples and thus undercover samples help to identify possible shortcomings, deficiencies and thus areas of improvement.

The test material was prepared of organic Matcha tea with incurred residues of *Nicotine* and *Perchlorate*. In addition to the incurred residues, five pesticides were spiked to the Matcha tea:

Anthraquinone, Bifenthrin, Lindane, Phthalimide, and Trifluralin.

The test material was distributed to six German laboratories and one Dutch laboratory. Depending on the common procedure of the client of the respective lab, the labs received 100 resp. 200 g of the Matcha tea. The tea was either provided in common sampling bags of the clients or in typical stand-up bags.

The labs were asked by their clients to analyse the tea sample applying

- a Pesticide Multi-residue Method (GC and LC modules),
- a single residue method related to Chlorate and Perchlorate, and
- a single residue method related to Nicotine.

The performance assessment considers the following test criteria:

- No false positive results.
- Correct identification of all 7 parameters, thus no false negative results.
- Correct quantification of the multi-method pesticides in terms of 70 to 120 % recovery of the spiked value.
- Correct quantification of Nicotine and Perchlorate in terms of z-scores ≤ |2|.

Bundesverband
Naturkost Naturwaren

Summary of the performance of the laboratories

Criterion	**Criterion passed**
Correct *identification* of all five multi-method pesticides	4 out of 7 laboratories (57 %)
Correct *quantification* of all five multi-method pesticides	2 out of 7 laboratories (29 %)
Correct *quantification* of Nicotine	4 out of 7 laboratories (57 %)
Correct *quantification* of Perchlorate	5 out of 7 laboratories (71 %)
Correct *quantification* of all seven parameters	2 out of 7 laboratories (29 %)
Successful participation according to BNN criteria	4 out of 7 laboratories (57 %)

Assessment of quantification

Analytical results within 70 and 120 % recovery of the spiked levels are considered satisfying for the assessment of the correct quantification of the pesticides Anthraquinone, Bifenthrin, Lindane, Phthalimide, and Trifluralin. Results, which correspond to z-scores ≤ |2| are considered satisfying in terms of the correct quantification of Nicotine and Perchlorate.

Parameter	Spiked level [mg/kg]	Assigned value [mg/kg]	Number of results	Correct quantification
Anthraquinone	0.016	-	4	4 out of 7 laboratories (57 %)
Bifenthrin	0.095	-	6	5 out of 7 laboratories (71 %)
Lindane	0.023	-	5	4 out of 7 laboratories (57 %)
Phthalimide	0.071	-	5	3 out of 7 laboratories 43 %)
Trifluralin	0.019	-	5	4 out of 7 laboratories (57 %)
Nicotine	incurred	0.073	6	4 out of 7 laboratories (57 %)
Perchlorate	incurred	0.27	6	5 out of 7 laboratories (71 %)

Bundesverband
Naturkost Naturwaren

Table of contents

1. Aim of the test and test design

The laboratory performance assessment was designed to verify the analytical competence related to BNN module-combination “A1 (pesticides) – B4 (tea and fruit teas, spices)”.

BNN decided to make use of a so-called “undercover test”. Such tests are an important tool to get knowledge about the daily performance of laboratories under routine conditions. The challenge in delivering correct results is significantly increased in undercover test compared to announced ring tests. No “special care analysis” is applied for undercover samples and thus undercover samples help to identify possible shortcomings, deficiencies and thus areas of improvement.

A sample of Matcha tea was provided to each of the laboratories by one of their commercial clients. The test sample was made up like a routine sample and was accompanied by a common submission form of the respective client. As a consequence, the labs were not aware of the undercover test.

Green tea was selected as matrix. The organic green tea used as test material contained incurred residues of Nicotine and Perchlorate and was additionally spiked with five multi-method pesticides:

Anthraquinone, Bifenthrin, Lindane, Phthalimide, and Trifluralin.

The test material should look like a routine sample in order be analysed by the labs without “special care”. The green tea was milled to a fine powder in order to prepare a homogeneous, high-quality test material. Since green tea powder is not common on the market, the tea sample was labelled as “Matcha tea”, origin “China”.

2. Test material preparation

An external laboratory prepared the test material under supervision of PROOF-ACS.

The organic green tea was ground to a fine powder. A sample of the untreated material was bottled as blank material for analysis.

The green tea powder was stirred in a Stephan cutter. The pesticides, solved in acetone, were added gradually, while the powder was continuously stirred. Stirring was continued after spiking to ensure a homogeneous distribution of all parameters.

Subsamples of 100 resp. 200 g were bottled and stored at -18 °C until shipment.

3. Statistical evaluation of results

3.1. Trueness criterion

The trueness criterion is applied for the evaluation of the multi-method pesticides Anthraquinone, Bifenthrin, Lindane, Phthalimide, and Trifluralin.

The trueness criterion considers the correct quantification of the actual analyte concentration in the sample. The trueness of the results is assessed as the coverage of the spiked level in %. The coverage of the spiked level is calculated of the result of each participant x_i according to the equation below:

$$\text{Coverage of the spiked level [\%]} = \frac{x_i}{spiked\ level} \times 100$$

Accepted range:

Results, which correspond to a recovery of 70 to 120 % of the spiked level, are considered satisfying in this laboratory performance assessment in accordance with the guidelines of the BNN[1]. A non-commercial rounding is applied during the calculation of the accepted ranges (two significant figures).

Examples:

- A recovery of 70 % of the spiked level of Bifenthrin (spiked level: 0.095 mg/kg) corresponds to an arithmetical value of 0.0665 mg/kg, which is rounded to the next lower figure: 0.066 mg/kg (slight increase of the accepted range).
- A recovery of 120 % of the spiked level of Anthraquinone (spiked level: 0.016 mg/kg) corresponds to an arithmetical value of 0.0192 mg/kg. 0.0192 mg/kg is rounded to the next higher figure: 0.020 mg/kg (also slight increase of the accepted range).

3.2. Comparability criterion

The comparability criterion is applied for the evaluation of the incurred residues of Nicotine and Perchlorate. The comparability of results is evaluated according to the z-score model based on an assigned value and a target standard deviation.

[1] BNN, Guidelines for laboratory approval by Bundesverband Naturkost Naturwaren (BNN) e. V. (Federal Association for Natural Foods and Natural Products inc. soc.), http://n-bnn.de/sites/default/dateien/GuidelinesBNNLabApprovalSeptember2015.pdf.

3.2.1. Assigned value

The assigned value x_{pt} is the robust mean, which is derived from the results of the participants according to ISO13528, Algorithm A [2]. The Winsorisation algorithm is applied to minimise the influence of outliers.

The assigned values are subject to commercial rounding and are presented with an accuracy of three significant figures.

3.2.2. z-score

The z-score is derived of the result x_i of each participant, the assigned value x_{pt} and the target standard deviation according to Horwitz σ_H[2,3]:

$$z - score = \frac{x_i - x_{pt}}{\sigma_H}$$

4. Results

4.1. Overall results

The laboratories received the test samples like routine samples of their clients. Thus, the laboratories were not aware of the test. The laboratories were requested to apply a pesticide multi residue method (with GC and LC modules) plus single residue methods for the quantification of Nicotine and Perchlorate.

The laboratories reported the results in a common sample reported to their clients, who forwarded the reports to PROOF-ACS for evaluation.

A summary of the overall performance of the labs is provided in table 1. Four out of seven laboratories (labs 1, 3, 6, 7) quantified at least 5 out of 7 parameters correctly and were thus considered successful according to the criteria for laboratory approval of BNN. Two laboratories (labs 6, 7) quantified all seven parameters correctly.

A more detailed evaluation of the results of the participants is presented in tables 2 to 8 and in figures 1 to 9. Due to a mix-up of the sample by lab 4, the results of lab 4 are not shown in the figures related to the multi-method pesticides (figures 1 to 5).

[2] Statistical methods for use in proficiency testing by interlaboratory comparison. ISO 13528:2015. Corrected version 2016-10-15.
[3] Horwitz W. Evaluation of Analytical Methods Used for Regulation of Foods and Drugs. Anal Chem. 1982;54(1):67A–76A.

4.2. Results per parameter

Anthraquinone

Four labs identified Anthraquinone correctly and quantified Anthraquinone correctly within 70 to 120 % of the spiked level.

Bifenthrin

Six out of seven labs identified Bifenthrin correctly. Five of the reported results are within 70 to 120 % of the spiked level, while the result of lab 3 is too low (58 % of the spiked level).

Lindane

Five out of seven labs identified lindane correctly. Four of them quantified it correctly within 70 to 120 % of the spiked level.

Phthalimide

Background levels of 0.02 mg/kg Phthalimide are common in tea samples[4]. The green tea, which was used as raw material in this test, contained traces of Phthalimide at a concentration level of 0.016 mg/kg.

As a consequence, the target value of Phthalimide of 0.071 mg/kg consists of the spiked level of 0.055 mg/kg and the background level of 0.016 mg/kg in the raw material.

Five labs identified Phthalimide correctly, while three labs reported results close to the target value (94 to 97 % of the target value). The results of labs 1 and 3 are significantly too low (49 % resp. 41 % of the target value).

Trifluralin

Trifluralin was identified correctly by five labs and quantified correctly by four out of seven labs.

Nicotine

Even though the clients of the labs ordered the analysis of Nicotine, lab 4 failed to identify Nicotine.

The result of lab 1 is significantly too high. The result is not considered for the calculation of the assigned value (outlier).

The comparability criterion is applied to evaluate the results of the labs with respect to Nicotine. Due to the limited number of data points, the results of five labs as well as the

[4] relana® Position Paper No. 16-03 “Folpet/Phthalimid” version 2016/07/22.

mean of the homogeneity testing and the mean of the stability testing are considered for the calculation of the assigned value.

The results of the labs 3, 5, 6, and 7 correspond to z-scores ≤ |2| and are considered satisfying.

Perchlorate

Lab 5 failed to identify Perchlorate and reported Chlorate at a concentration level of 0.29 mg/kg instead.

Due to the limited number of data points, the results of the six labs as well as the mean of the homogeneity testing and the mean of the stability testing are considered for the calculation of the assigned value.

Five of the lab (labs 1, 2, 3, 6, and 7) reported results, which correspond to z-scores ≤ |2| and are thus considered satisfying.

The result of lab 4 is significantly too low (z-score -2.5).

4.3. Results per lab

Lab 1

Lab 1 identified all parameters correctly. Five out of seven parameters are quantified correctly. However, the result of Phthalimide is significantly too low and the result of Nicotine is far too high. The lab additionally reported Biphenyl, o-Phenyl phenol and 1-Naphthol, which are not considered false positive as these substances are known to be often present in tea samples as a consequence of process steps during tea production.

The overall performance of lab 1 is considered satisfying according to the criteria for laboratory approval of the BNN.

Lab 2

Lab 2 failed to identify four parameters, which are usually quantified by GC: Anthraquinone, Lindane, Phthalimide, and Trifluralin. The result of Nicotine is too high. Bifenthrin and Perchlorate are quantified correctly.

The lab failed the BNN criteria.

Lab 3

Lab 3 subcontracted the analysis of Nicotine. The lab quantified Anthraquinone, Lindane, Trifluralin, Nicotine and Perchlorate correctly, while the results of Bifenthrin and Phthalimide are too low.

The overall performance of the lab is considered satisfying according to the criteria of BNN.

Lab 4

Lab 4 failed to identify all parameters, which are analysed by the pesticide multi-method approach. After the lab was informed about the undercover test, the lab immediately performed a first investigation. As a result, a mix-up of the test sample with another routine sample (of different matrix, but similar colour) during sample clean-up for multi residue testing was identified.

As a consequence of the mix-up, lab 4 reported results of 0.20 mg/kg Boscalid, 0.10 mg/kg Fluopyram, and 0.014 mg/kg Thiamethoxam. The lab re-analysed the test sample and provided an amendment of the first sample report. The results of the re-analysis are summarised in table 9. However, these results were not considered for evaluation.

Even though the mix-up was relevant for the multi-method pesticides only, the lab also failed to identify Nicotine in the test sample and reported significantly too low results related to Perchlorate.

The overall performance of the lab is dissatisfying. The lab failed the BNN criteria.

Lab 5

Lab 5 failed to identify Anthraquinone and Perchlorate in the test sample. Anthraquinone is not included in the scope of the lab, which was provided for BNN-approval. However, Anthraquinone is a highly relevant parameter especially for organic tea samples. As a consequence, the result of Anthraquinone is considered false negative. The lab should include Anthraquinone to the scope of the analytical method as soon as possible.

In addition, lab 5 reported Chlorate instead of Perchlorate.

The lab quantified Bifenthrin, Phthalimide, and Nicotine within the accepted ranges, while the result related to Lindane is too high (138 % of the spiked level) and the result related to Trifluralin is too low (63 % of the spiked level).

The overall performance of the lab is dissatisfying. The lab failed the BNN criteria.

Lab 6

Lab 6 identified and quantified all parameters correctly.

The performance of the lab is considered highly satisfying.

Lab 7

Lab 7 identified and quantified all parameters correctly.

Lab 7 additionally reported o-Phenyl phenol at a concentration level of 0.025 mg/kg, which was not considered false positive, as this substance is known to be often present in tea samples as a consequence of process steps during tea production.

The performance of the lab is considered highly satisfying.

5. Conclusion

- Four laboratories quantified at least 5 out of 7 parameters and are thus considered successful according to the criteria for laboratory approval of BNN.
- Two of the laboratories identified and quantified all parameters correctly.
- The test revealed significant shortcomings at three laboratories (mix-up of the sample, bad performance for the parameter analysed by GC, mix-up in reporting of Chlorate and Perchlorate).

6. Tables and figures

Table 1. Summary of the overall performance

Lab code	Anthra-quinone	Bifenthrin	Lindane	Phthalimide	Trifluralin	Nicotine	Perchlorate	Participation successful*
1	yes	yes	yes	no	yes	no	yes	yes
2	n.r.	yes	n.r.	n.r.	n.r.	no	yes	no
3	yes	no	yes	no	yes	yes#	yes	yes
4	n.r.	n.r.	n.r.	n.r.	n.r.	no	no	no
5	n.r.	yes	no	yes	no	yes	n.r.	no
6	yes	yes	yes	yes	yes	yes	yes	yes
7	yes	yes	yes	yes	yes	yes	yes	yes
Successful labs in % of total labs	**57**	**71**	**57**	**43**	**57**	**57**	**71**	**57**

yes: correctly quantified; no: quantification dissatisfying; n.r: not reported
The analysis of nicotine was subcontracted.
* A laboratory was considered successfully according to the criteria of the BNN if it identified all parameters and reported at least 5 out of 7 parameters correctly.

Table 2. Results of Anthraquinone

	Anthraquinone		
	Spiked level [mg/kg]		0.016
	Accepted range [mg/kg]	0.011 -	0.020
Laboratory code	**Result [mg/kg]**	**Result in % of the spiked level**	**Trueness criterion passed**
1	0.015	94	yes
2	n.r.	-	no
3	0.016	100	yes
4	n.r.*	-	no
5	n.r.	-	no
6	0.019	119	yes
7	0.019	119	yes

n.r.: not reported
* see table 10

Table 3. Results of Bifenthrin

	Bifenthrin		
	Spiked level [mg/kg]		0.095
	Accepted range [mg/kg]	0.066 -	0.12
Laboratory code	**Result [mg/kg]**	**Result in % of the spiked level**	**Trueness criterion passed**
1	0.085	89	yes
2	0.069	73	yes
3	0.055	58	no
4	n.r.*	-	no
5	0.098	103	yes
6	0.075	79	yes
7	0.075	79	yes

n.r.: not reported
* see table 10

Table 4. Results of Lindane

	Lindane		
	Spiked level [mg/kg]		0.023
	Accepted range [mg/kg]	0.016 -	0.028
Laboratory code	**Result [mg/kg]**	**Result in % of the spiked level**	**Trueness criterion passed**
1	0.020	87	yes
2	n.r.	-	no
3	0.017	74	yes
4	n.r.*	-	no
5	0.032	139	no
6	0.023	100	yes
7	0.028	122	yes

n.r.: not reported
* see table 10

Table 5. Results of Phthalimide

	Phthalimide		
	Target value [mg/kg]		0.071
	Accepted range [mg/kg]	0.049 -	0.086
Laboratory code	**Result [mg/kg]**	**Result in % of the spiked level**	**Trueness criterion passed**
1	0.035	49	no
2	n.r.	-	no
3	0.029	41	no
4	n.r.*	-	no
5	0.067	94	yes
6	0.068	96	yes
7	0.069	97	yes

n.r.: not reported
* see table 10

Table 6. Results of Trifluralin

	Trifluralin		
	Spiked level [mg/kg]		0.019
	Accepted range [mg/kg]	0.013 -	0.023
Laboratory code	**Result [mg/kg]**	**Result in % of the spiked level**	**Trueness criterion passed**
1	0.020	105	yes
2	n.r.	-	no
3	0.014	74	yes
4	n.r.*	-	no
5	0.012	63	no
6	0.014	74	yes
7	0.016	84	yes

n.r.: not reported
* see table 10

Table 7. Results of Nicotine

	Nicotine		
	Spiked level [mg/kg]		incurred
	Assigned value [mg/kg]		0.0729
Laboratory code	**Result [mg/kg]**	**z-score**	**Comparability criterion passed**
1	0.52	27.9	no
2	0.12	2.9	no
3	0.067*	-0.4	yes
4	<0.01	-4.5	no
5	0.095	1.4	yes
6	0.067	-0.4	yes
7	0.063	-0.6	yes

n.r.: not reported
* The analysis of nicotine was subcontracted to another laboratory

Table 8. Results of Perchlorate

	Perchlorate		
	Spiked level [mg/kg]		incurred
	Assigned value [mg/kg]		0.270
Laboratory code	**Result [mg/kg]**	**z-score**	**Comparability criterion passed**
1	0.25	-0.4	yes
2	0.31	0.8	yes
3	0.28	0.2	yes
4	0.14	-2.5	no
5	n.r.	-5.1	no
6	0.28	0.2	yes
7	0.25	-0.4	yes

n.r.: not reported

Table 9. Further results

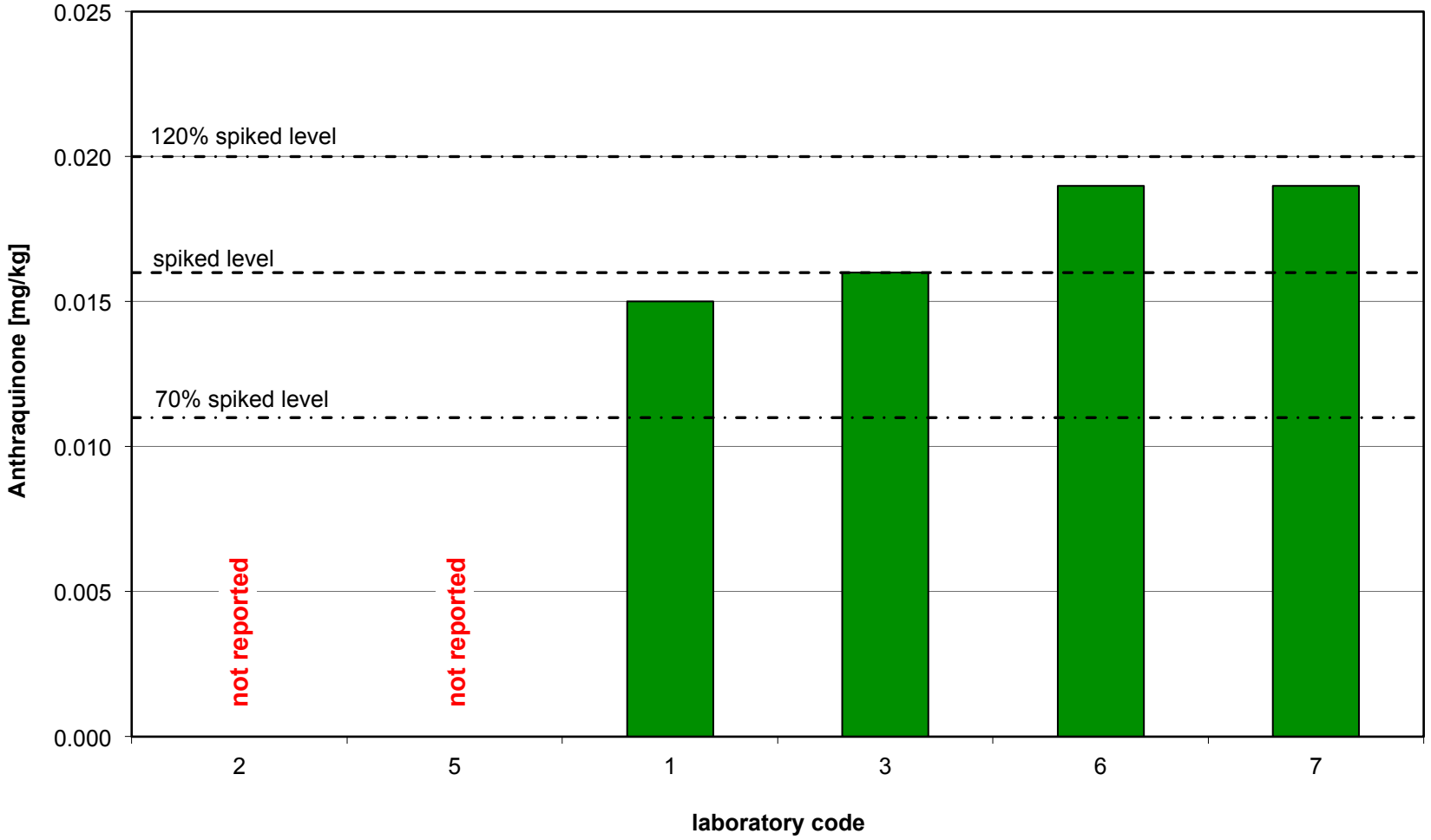
Laboratory code	Further results
1	0.030 mg/kg Biphenyl, 0.010 mg/kg o-Phenyl phenol, 0.035 mg/kg 1-Naphthol
5	0.29 mg/kg Chlorate
7	0.025 mg/kg o-Phenyl phenol

Table 10. Re-analysis of the sample by lab 4

Parameter	Result of analysis 1 (undercover test) [mg/kg]	Result of re-analysis (analysis 2, after uncovering the test) [mg/kg]	Results of analysis 2 in % of the spiked level
Anthraquinone	n.r.	0.014	88
Bifenthrin	n.r.	0.057	60
Lindane	n.r.	0.015	65
Phthalimide	n.r.	Folpet (sum) 0.10	
Trifluralin	n.r.	0.014	74
Nicotine	<0.01	not re-analysed	false negative
Perchlorate	0.14	not re-analysed	z-score -2.5
Further parameters*	0.20 mg/kg Boscalid, 0.10 mg/kg Fluopyram, 0.014 mg/kg Thiamethoxam	not confirmed	-

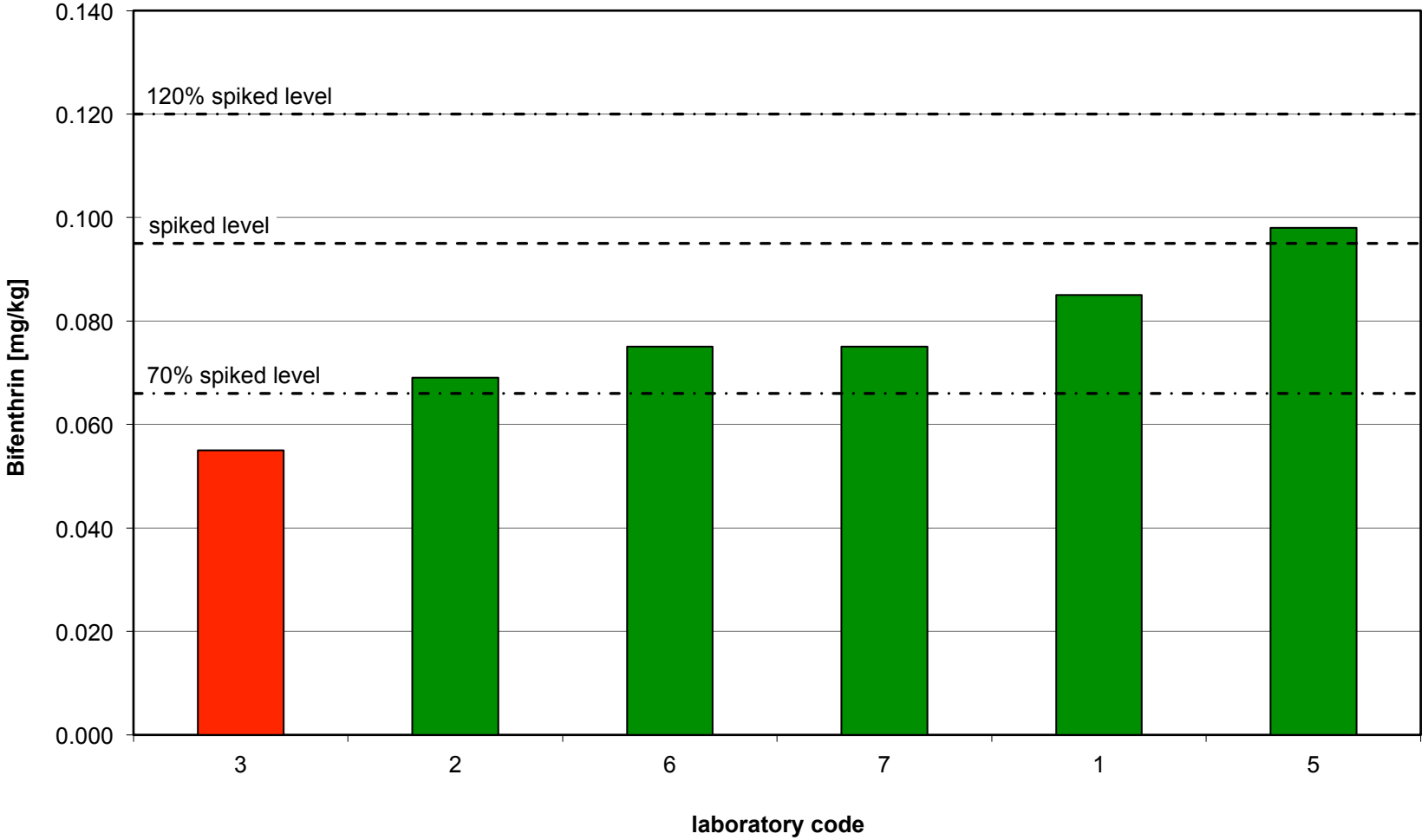
n.r.: not reported
* Due to a mix-up of the test sample with another routine sample.

Figure 1. Assessment of Anthraquinone



Green: satisfactory results, red: dissatisfactory results

Figure 2. Assessment of Bifenthrin

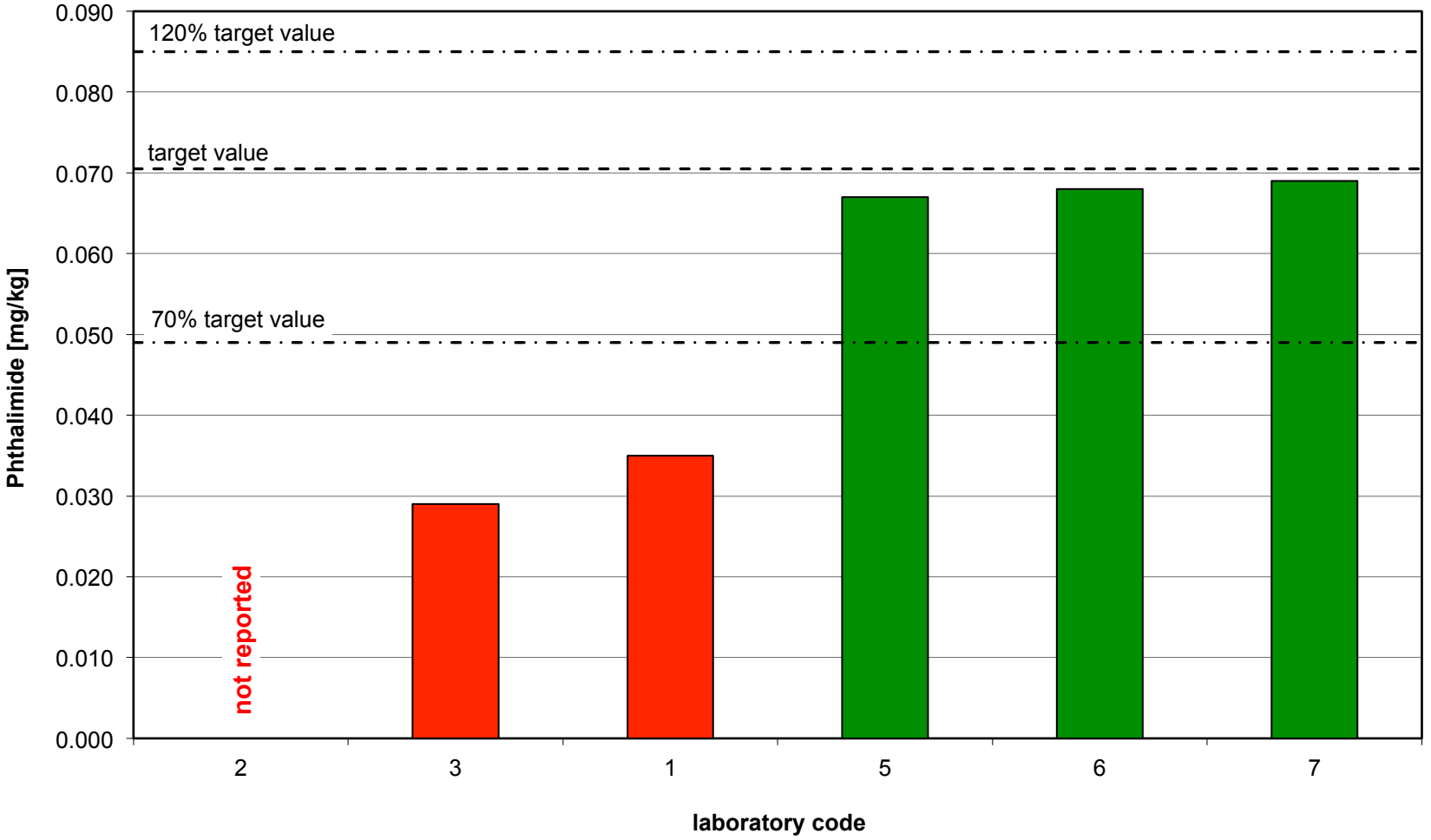


Green: satisfactory results, red: dissatisfactory results

Figure 3. Assessment of Lindane

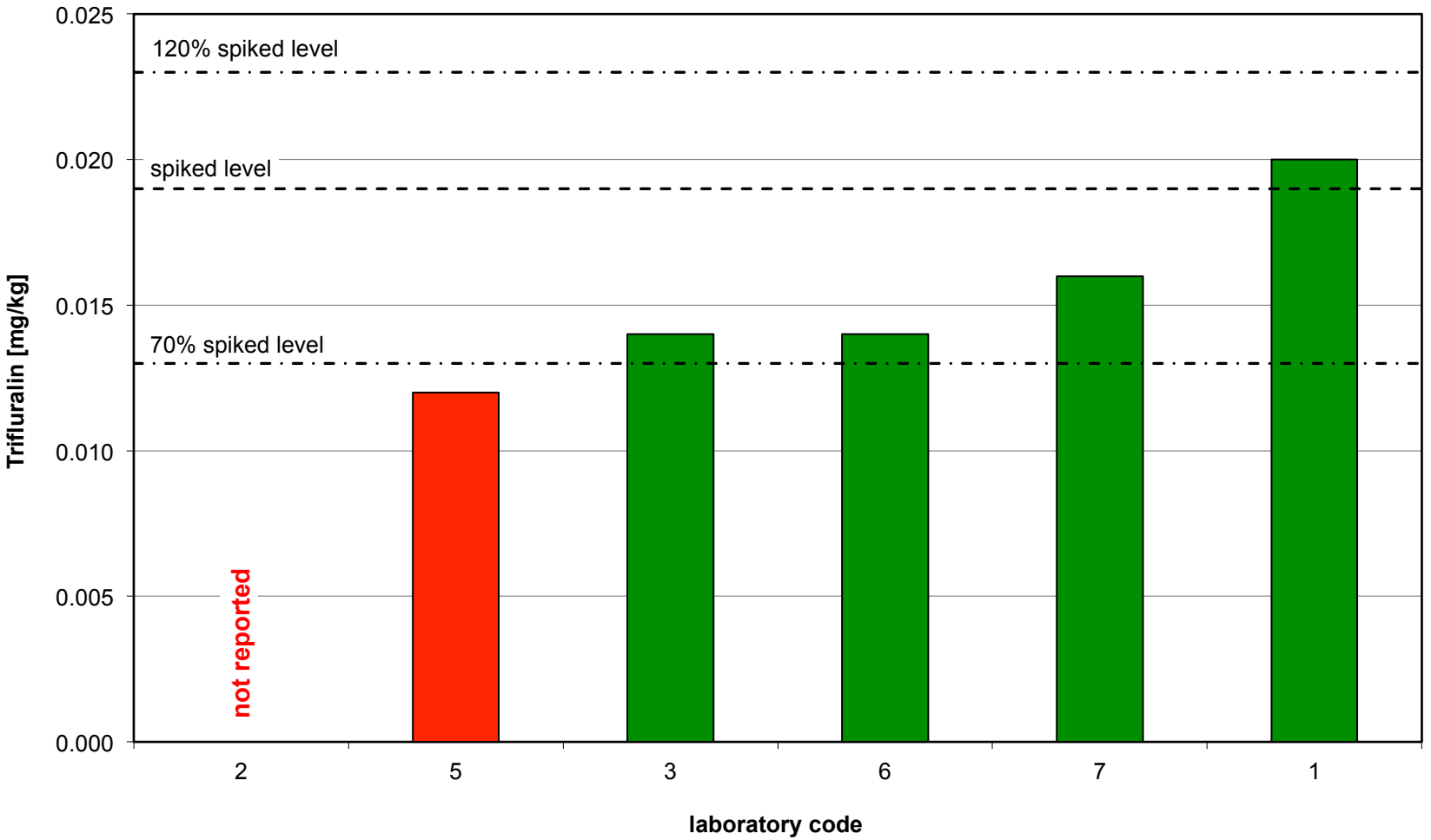
Green: satisfactory results, red: dissatisfactory results

Figure 4. Assessment of Phthalimide



Green: satisfactory results, red: dissatisfactory results

Figure 5. Assessment of Trifluralin

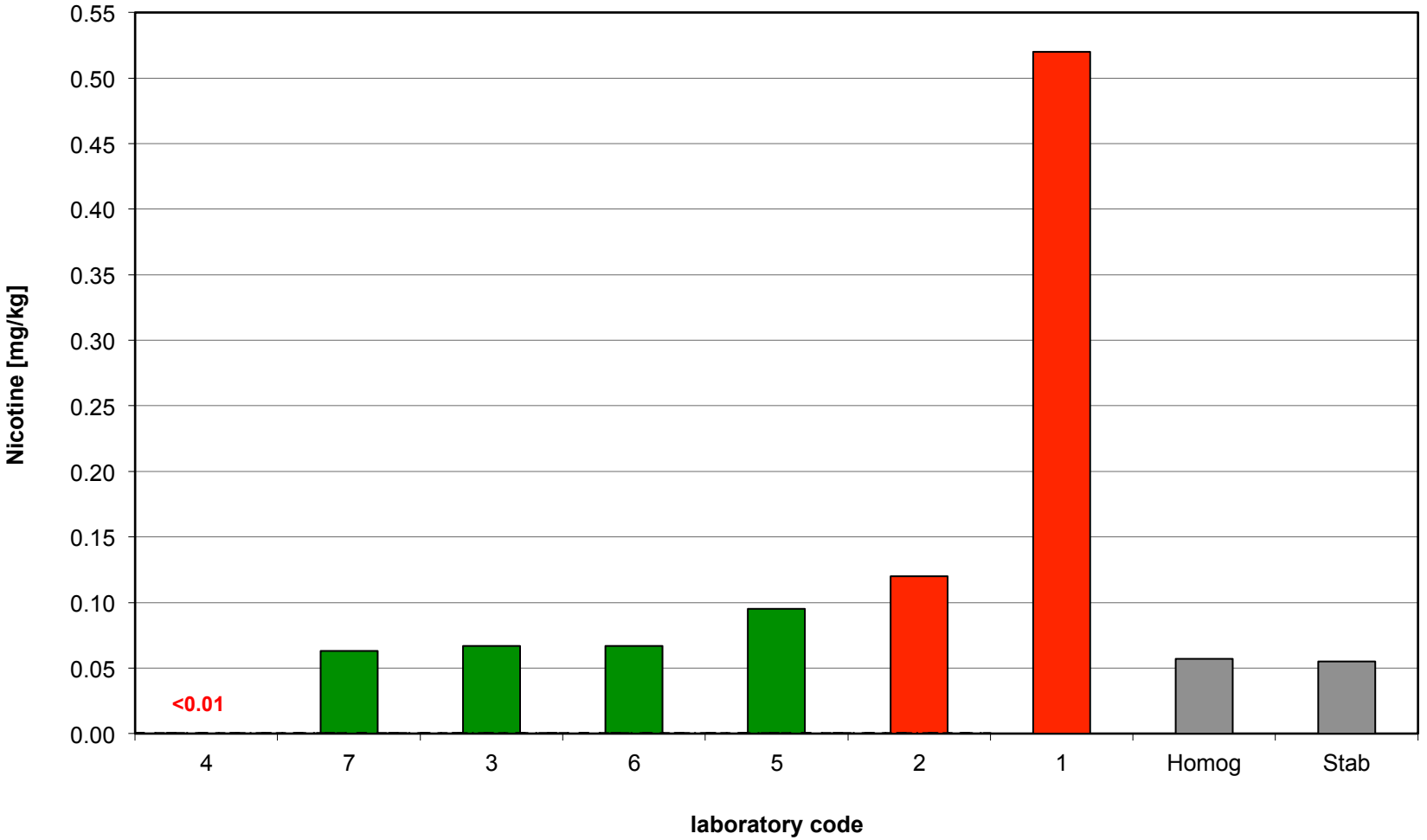


Green: satisfactory results, red: dissatisfactory results

Figure 6. Assessment of Nicotine (incurred residue) - Comparability

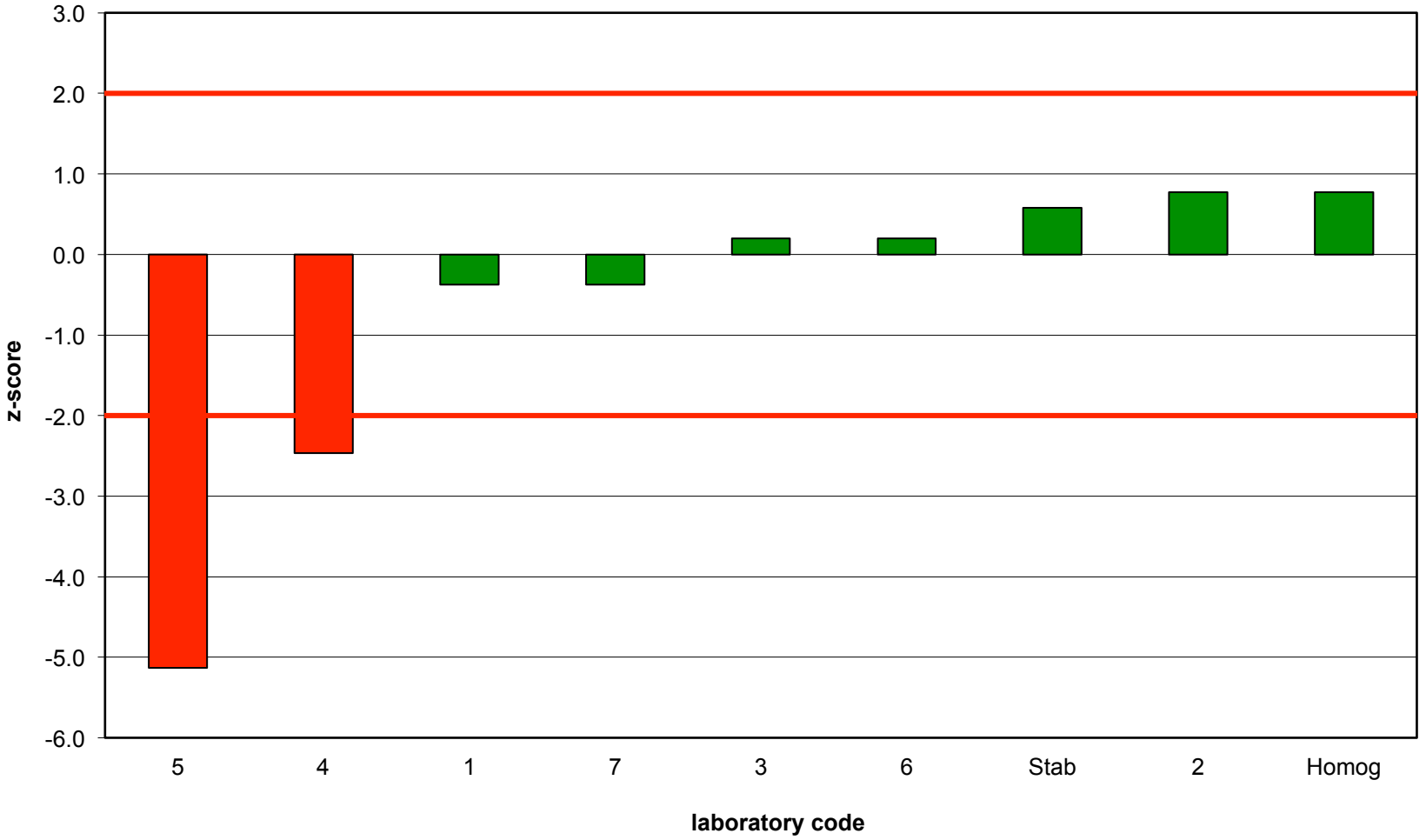
Green: satisfactory results, red: dissatisfactory results

Figure 7. Results related to Nicotine (incurred residue)



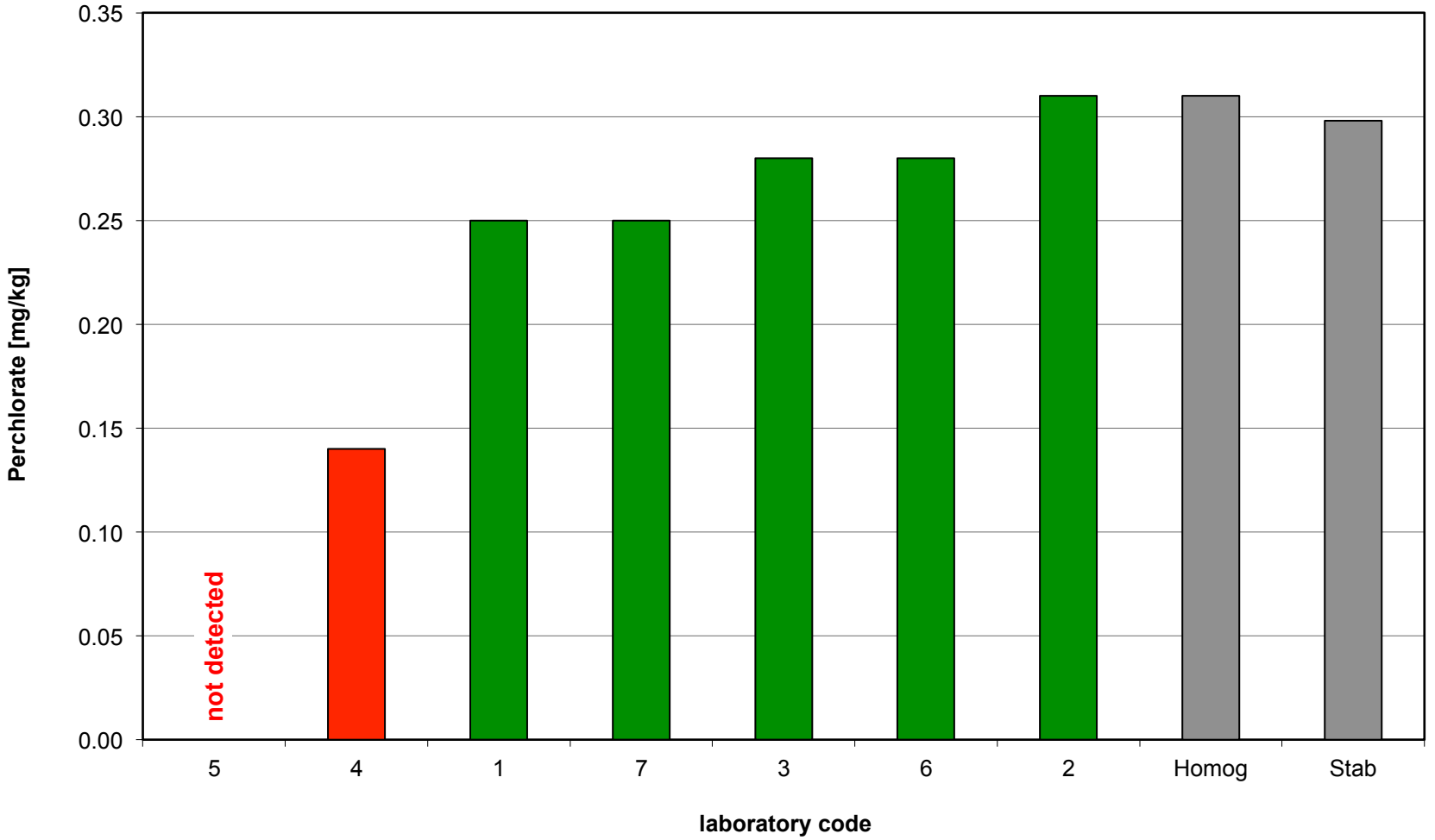
Green: satisfactory results, red: dissatisfactory results

Figure 8. Assessment of Perchlorate (incurred residue) - Comparability



Green: satisfactory results, red: dissatisfactory results

Figure 9. Results related to Perchlorate (incurred residue)



Green: satisfactory results, red: dissatisfactory results

7. Homogeneity testing

Seven samples are randomly chosen for homogeneity testing. Each subsample is analysed for the spiked pesticides in duplicate. One sample is also analysed for Nicotine and Perchlorate in duplicate. The results confirm the homogeneous distribution of the parameters in the test material and the spiked levels (table 11).

Table 11. Results of the homogeneity testing

Subsample No.	Extraction No.	Anthraquinone [mg/kg]	Bifenthrin [mg/kg]	Lindane [mg/kg]	Phthalimide [mg/kg]	Trifluralin [mg/kg]	Nicotine [mg/kg]	Perchlorate [mg/kg]
1	1	0.017	0.061	0.021	0.072	0.015	0.058	0.30
	2	0.010	0.068	0.025	0.082	0.014	0.059	0.31
2	1	0.015	0.060	0.025	0.074	0.014		
	2	0.017	0.066	0.026	0.066	0.014		
3	1	0.016	0.063	0.021	0.072	0.015		
	2	0.015	0.063	0.023	0.073	0.015		
4	1	0.012	0.058	0.024	0.064	0.013		
	2	0.013	0.065	0.026	0.079	0.015		
5	1	0.012	0.063	0.023	0.075	0.013		
	2	0.018	0.077	0.022	0.069	0.015		
6	1	0.014	0.062	0.020	0.076	0.015		
	2	0.019	0.068	0.031	0.080	0.017		
7	1	0.017	0.070	0.029	0.075	0.017		
	2	0.021	0.066	0.034	0.077	0.019		
Mean [mg/kg]		0.015	0.065	0.025	0.074	0.015	0.057	0.31
Standard deviation [mg/kg]		0.0030	0.0048	0.0040	0.0051	0.0016	-	-
Coefficient of variation [%]		20.3	7.4	16.1	6.9	10.9	-	-
Spiked level [mg/kg]		0.016	0.095	0.023	0.071	0.019	incurred	incurred
Recovery of the spiked level [%]		93	68	109	105	79		

8. Stability testing

After the closure of transmission of results the test material is re-analysed to confirm the stability of the parameters over the period of the test. The mean of the re-analysis is compared to the mean result of the homogeneity testing (table 12). The results confirm the stability of all analytes throughout the test (recoveries of 80 to 121 % of the homogeneity testing).

Table 12. Results of the stability testing

Pesticide	Mean result from homogeneity testing [mg/kg]	Stability testing (at closure of transmission of results)			Recovery compared to the mean of the homogeneity testing [%]
		Result 1 [mg/kg]	Result 1 [mg/kg]	Mean of result 1 and 2 [mg/kg]	
Anthraquinone	0.015	0.018	0.018	0.018	120
Bifenthrin	0.065	0.052	0.052	0.052	80
Lindane	0.029	0.020	0.020	0.020	80
Phthalimide	0.074	0.093	0.085	0.089	121
Trifluralin	0.015	0.016	0.015	0.016	103
Nicotine	0.057	0.057	0.053	0.055	94
Perchlorate	0.31	0.30	0.30	0.30	97