Results of Proficiency Test
TCEP in Plastics
February 2015

Organised by: Institute for Interlaboratory Studies
Spijkenisse, the Netherlands

Authors: ing. C.M. Nijssen-Wester
Correctors: ing. N. Boelhouwer & dr. R.G. Visser
Report: iis15P01

April 2015

CONTENTS

Appendices:

1 INTRODUCTION

The substance tris(2-chloroethyl)phosphate (TCEP) is an alkyl phosphate ester used as a flame-retardant, plasticiser and viscosity regulator in polyurethanes, polyester resins, poly acrylic plastic and other polymers. The main industrial branches in which TCEP has been used are the building industry, the furniture and the textile industry.
However, production and use has been in decline since the 1980s, when TCEP has been progressively replaced by other flame retardants. TCEP was comprehensively evaluated under the EU existing substances regulation (EEC) 793/93 in 2009. TCEP is classified under Regulation (EC) No 1272/2008 as a carcinogenic, mutagenic and toxic substance. In March 2012, the European Union decided to lower the limit of TCEP in toys (5 mg/kg) . Regretfully, no certified reference materials (CRMs) for TCEP are available to optimise the determination of TCEP. As an alternative, participation in a proficiency test may enable the laboratories to check their performance and thus to increase this comparability.

Therefore, a proficiency testing scheme (laboratory-evaluating interlaboratory study) for the determination of TCEP was started by the Institute for Interlaboratory Studies in 2014. During the annual proficiency testing program 2014/2015, it was decided to continue the PT for the analysis of TCEP. In the international interlaboratory study of February 2015, 36 laboratories from 16 different countries participated (See appendix 3). In this report the results of the 2015 proficiency test are presented and discussed. This report is also electronically available through the iis internet site www.iisnl.com.

2 SET UP

The Institute for Interlaboratory Studies in Spijkenisse was the organizer of this proficiency test. It was decided to send 1 plastic sample (approximately 3 gram), positive on TCEP, and labelled #15007. Sample analyses for fit-for-use and homogeneity testing were subcontracted to an ISO17025 accredited laboratory. Participants were requested to report rounded and unrounded test results. These unrounded test results were preferably used for statistical evaluation. The participants were asked to report the analytical results using the indicated units on the report form.

2.1 QUALITY SYSTEM

The Institute for Interlaboratory Studies in Spijkenisse, the Netherlands, has implemented a quality system based on ISO/IEC 17043:2010. This ensures strict adherence to protocols for sample preparation and statistical evaluation and 100% confidentiality of participant’s data. Also customer’s satisfaction is measured on a regular basis by sending out questionnaires.

2.2 PROTOCOL

The protocol followed in the organization was the one as described for proficiency testing in the report 'iis Interlaboratory Studies: Protocol for the Organization, Statistics and Evaluation' of April 2014 (iis-protocol, version 3.3). The protocol can be downloaded from iis website http://www.iisnl.com.

2.3 CONFIDENTIALITY STATEMENT

All data presented in this report must be regarded as confidential and for use by the participating companies only. Disclosure of the information in this report is only allowed by means of the entire report. Use of the contents of this report for third parties is only allowed by written permission of the Institute for Interlaboratory Studies. Disclosure of the identity of one or more of the participating companies will be done only after receipt of a written agreement of the companies involved.

2.4 SAMPLES

Plastic (red) granulates, positive on TCEP, were made by a third party. Samples of approx. 3 gram were prepared. Six stratified randomly selected samples were tested using EN71-11 to check the homogeneity of the batch. See the following table for the test results.

	TCEP in mg/kg
sample #15007-1	176.3
sample #15007-2	169.6
sample #15007-3	171.1
sample #15007-4	173.6
sample #15007-5	167.6
sample #15007-6	166.5

table 1: homogeneity test results of subsamples #15007

From the test results of table 1, the repeatability was calculated and compared with 0.3 times the corresponding estimated target reproducibility in agreement with the procedure of ISO 13528, Annex B2 in the next table:

	TCEP in mg/kg
r (observed)	10.4
reference method	EN71-11:2005
0.3 x R (reference method)	11.2

Table 2: repeatabilities of subsamples #15007

The calculated repeatability of the test results was in agreement with 0.3 times the reproducibility mentioned in (or estimated from) the reference method EN71-11. Therefore, homogeneity of the subsamples was assumed.

Approx. 3 grams of sample #15007 was sent to each of the participating laboratories on January 21, 2015.

2.5 ANALYSES

The participants were requested to determine the concentration of TCEP, applying the analysis procedure that is routinely used in the laboratory. To get comparable results a detailed report form, on which the unit was prescribed, was sent together with the sample. Also, a letter of instructions was added to the package. The laboratories were also requested to report some of the test conditions that the laboratory has used.

3 RESULTS

During four weeks after sample despatch, the results of the individual laboratories were gathered. The original data are tabulated in the appendices of this report. The laboratories are presented by their code numbers.

Directly after the deadline, a reminder fax was sent to those laboratories that had not yet reported. Shortly after the deadline, the available results were screened for suspect data. A result was called suspect in case the Huber Elimination Rule (a robust outlier test, see lit.5) found it to be an outlier. The laboratories that produced these suspect data were asked to check the results. Additional or corrected data are placed under 'Remarks' in the result tables in appendix 4. A list of abbreviations used in the tables can be found in appendix 4.

3.1 STATISTICS

The protocol followed in the organisation of this proficiency test is described in the report 'iis Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation' of april 2014 (iis-protocol, version 3.3).

For the statistical evaluation the *unrounded* (when available) figures were used instead of the rounded results. Results reported as '<…' or '>…' were not used in the statistical evaluation.

First, the normality of the distribution of the various data sets per determination was checked by means of the Lilliefors-test, a variant of the Kolmogorov-Smirnov test and by the calculation of skewness and kurtosis. Evaluation of the three normality indicators in combination with the visual evaluation of the graphic Kernel density plot, lead to judgement of the normality being either ‘unknown’, ‘OK’, ‘suspect’ or ‘not OK’. After removal of outliers, this check was repeated. Not all data sets proved to have a normal distribution, in which cases the statistical evaluation of the results should be used with due care.

According to ISO 5725 the original results per determination were submitted to Dixon’s and/or Grubbs' and/or Rosner’s outlier tests. Outliers are marked by D(0.01) for the Dixon’s test, by G(0.01) or DG(0.01) for the Grubbs’ test and by R(0.01) for the Rosner’s test (ref. 14). Stragglers are marked by D(0.05) for the Dixon’s test, by G(0.05) or DG(0.05) for the Grubbs’ test and by R(0.05) for the Rosner’s test. Both outliers and stragglers were not included in the calculations of averages and standard deviations.

For each assigned value the uncertainty was determined in accordance with ISO13528. Subsequently the calculated uncertainty was evaluated against the respective requirement based on the target reproducibility in accordance with ISO13528. When the uncertainty passed the evaluation no remarks are made in the report. However, when the uncertainty failed the evaluation it is mentioned in the report and it will have consequences for the evaluation of the test results.

Finally, the reproducibilities were calculated from the standard deviations by multiplying these with a factor of 2.8.

3.2 GRAPHICS

In order to visualise the data against the reproducibilities from literature, Gauss plots were made, using the sorted data for one determination (see appendix 1). On the Y-axis the reported analysis results are plotted. The corresponding laboratory numbers are under the X-axis.
The straight horizontal line presents the consensus value (a trimmed mean). The four striped lines, parallel to the consensus value line, are the +3s, +2s, -2s and -3s target reproducibility limits of the selected standard. Outliers and other data, which were excluded from the calculations, are represented as a cross. Accepted data are represented as a triangle.
Furthermore, Kernel Density Graphs were made. The Kernel Density Graph is a method for producing a smooth density approximation to a set of data that avoids some problems associated with histograms. Also a normal Gauss curve was projected over the Kernel Density Graph for reference.

3.3 Z-SCORES

To evaluate the performance of the participating laboratories, the z-scores were calculated. As it was decided to evaluate the performance of the participants in this proficiency test (PT) against the literature requirements, e.g. EN reproducibility, the z-scores were calculated using a target standard deviation. This results in an evaluation independent of the spread of this interlaboratory study. The target standard deviation was calculated from the literature reproducibility by division with 2.8.

When a laboratory did use a test method with a reproducibility that is significantly different from the reproducibility of the reference test method used in this report, it is strongly advised to recalculate the z-score, while using the reproducibility of the actual test method used this in order to evaluate whether the reported test result is fit-for-use.

The z-scores were calculated according to:

$z_{(target)}$ = (result - average of PT) / target standard deviation

Absolute values for z<2 are very common and absolute values for z>3 are very rare.

The usual interpretation of z-scores is as follows:

	|z| < 1	good
1 <	|z| < 2	satisfactory
2 <	|z| < 3	questionable
3 <	|z|	unsatisfactory

4 EVALUATION

During the execution of this proficiency test no reporting problems occurred. Thirty-three participants reported a test result of which three participants after the deadline. Three other participants did not report any test results. Finally, the 33 participants did report 32 numerical results. Observed were 2 outlying results, which is 6.3% of the numerical results. In proficiency studies, outlier percentages of 3% - 7.5% are quite normal.
The distribution of the data set of TCEP is not clearly Gaussian.

For the determination of TCEP, there is no standard method available. Most participating laboratories therefore had to perform an in house method. This will consist of a preparation/extraction step and an analytical step. Method EN71-11 describes the analytical determination of TCEP after extraction and has a precision statement for TCEP. That is the reason that in this report EN71-11 is used as reference method (for the analytical determination). It is also possible to use the estimated reproducibility calculated with the Horwitz equation.
Regretfully in EN71-11:2005, no reproducibility requirements for TCEP are mentioned, but only the standard deviation for the repeatability. The target reproducibility is estimated as follows: the standard deviation was multiplied with 2.8 to get the target repeatability. This was multiplied with 3 to get an estimate of the target reproducibility.
For comparison also the Horwitz equation was used to estimate a target reproducibility. This estimated Horwitz reproducibility was equal or smaller than the estimated reproducibility of EN71-11.

4.1 EVALUATION PER COMPONENT

In this section, the results are discussed per sample. All statistical results reported on the sample #15007 are summarised in appendix 1 and analytical details provided by the participants are summarised in appendix 2.

TCEP: This determination was problematic. Two statistical outliers were observed after the exclusion of 18 test results (without the excluded test results the data set showed one outlier). The observed reproducibility after rejection of the suspect data was not in agreement with the estimated target reproducibility of EN71-11:2005 and not in agreement with the estimated reproducibility calculated using the Horwitz equation.

It was decided to use only the data of the laboratories that did reduce the

grain size of the plastic granulate, e.g. by cutting or grinding. Looking at the analytical details, it is remarkable that not all laboratories used a preparation to reduce the size of the granulates of the plastic. In order to extract organic chemicals like TCEP from plastic, a larger surface area would be beneficial.

Furthermore the distribution of the data of all participants was not in line with a Gaussian distribution around the mean of these data. So the consensus value found for the whole group did not appear to be in line with the test results that a large subgroup found. Also when looking at the Gaussian distribution of the laboratories that did cut or grind the sample, the spread in this group was much smaller and the mean was more in line with the mean value based on the distribution. For the Kernel Density graphs, see §5, Discussion.

4.2 PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES

A comparison has been made between the reproducibility as declared by the relevant standard method and the reproducibility as found for the group of participating laboratories.

The number of significant results, the average result, the calculated reproducibility (standard deviation*2.8) and the target reproducibility, derived (or estimated) from the reference test method EN71-11 are presented in the next table.

Parameter	unit	n	Average	2.8 * sd	R(target)
TCEP	mg/kg	12	149.57	52.01	32.67

table 3: reproducibility of TCEP in sample #15007

Without further statistical calculations, it can be concluded that the group of participating laboratories have problems with the analysis of TCEP in plastic at this concentration level. See also the discussion in paragraphs 4.1 and 5.

4.3 COMPARISON OF THE PROFICIENCY TEST OF FEBRUARY 2015 WITH THE PREVIOUS PT

	February 2015	February 2014
Number of reporting labs	33	23
Number of results reported	32	23
Number of statistical outliers	2	1
Percentage outliers	6.3%	4.3%

Table 4: Comparison with previous proficiency test

In proficiency tests, outlier percentages of 3% - 7.5% are quite normal.

The uncertainties in the test results of TCEP in the iis15P01 PT are improved compared to the previous PT, but still not in line with the uncertainties of the method (see table 4).

Parameter	February 2015	February 2014	Est. EN71-11
TCEP	12.4%	23.0%	7.8%

Table 5: Development of relative uncertainties over the years

5 DISCUSSION

The material for this PT was a plastic granulate. In order to extract TCEP from a solid like a polymer, the extraction solvent, the extraction conditions and the surface area will be important variables. The choice of the extraction solvent was the most important variable in the PT of 2014, since TCEP had to be extracted from a low density foam. The conclusion in this 2014 PT was that the use of acetonitrile as a solvent gave a much smaller spread of the test results than the use of other solvents.

Looking at the analytical details for the PT in 2015, solvent and extraction conditions appear to be less important than the surface area. A larger surface area (or smaller grain size) will give the extraction solvent better access to extract the TCEP. This is to be expected as the sample is a hard high density plastic, not foam.
Therefore the total data set was compared to the test results of only the participants that reduced the grain size of the granulate. Statistical evaluation of this latter group showed a higher consensus value for TCEP with a much smaller spread. The consensus value also matched with the peak of the distribution of the data of the whole group. Because of this, it was decided to use only the data of the participants that reduced the grain size for calculation of the z-scores.

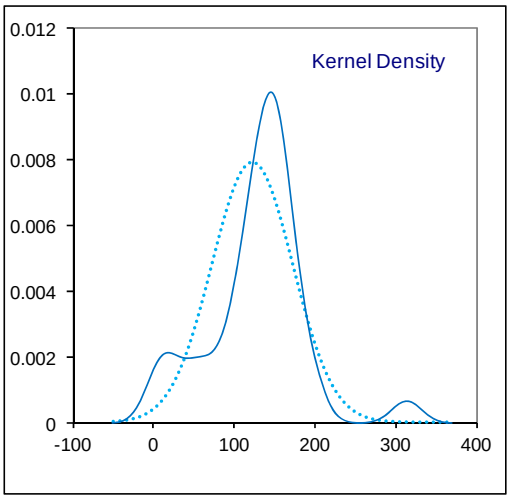


Figure 1: Kernel Density of all data (n=31, no exclusions)

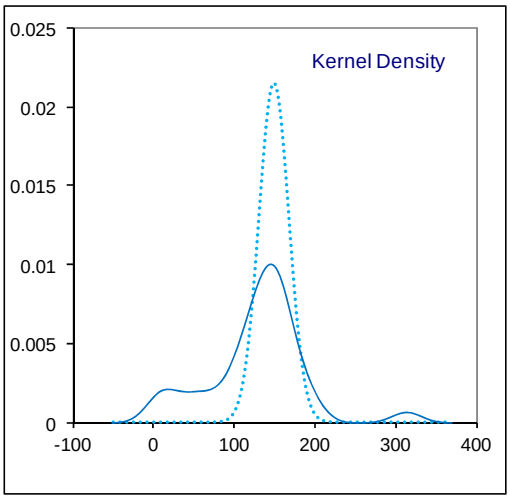


Figure 2: Kernel Density of labs who reduced sample (n=12)

To explain this further, the Gaussian distribution or Kernel Density graphs are used.
Looking at the whole group of data (figure 1), the theoretical Gaussian distribution around the consensus value (maximum at 122.754 mg/kg) using the standard deviation of this group (dotted line) is not in line with the distribution of the data found (continuous line), with a maximum at 145.7 mg/kg.

When using the consensus value of only the group of laboratories, that did reduce the grain size (figure 2, maximum at 149.568 mg/kg), it can be seen that the theoretical Gaussian distribution (dotted line) is in line with the distribution of the data of the whole group (maximum at 145.7 mg/kg).

Looking at the width of the estimated normal distribution graphs (dotted lines), it can also be seen that this is much smaller for the group of labs that reduced the grain size. This means that the mean of this group has a higher certainty to be correct.

Like with many other analyses, in this test extraction is the critical step. This means the combination of solvent, extraction method (time and temperature) and surface area is crucial. Since most participants performed an in-house method, the way of testing is not standardized. This can explain the large spread found for the whole group.

It is most likely that by continuous participation in a PT like this one, the in house procedures will focus more on the same critical steps, thus improving their performance. A similar situation has been described about the already ongoing PTs on metals or phthalates in plastic (see ref. 15, appendix 4).

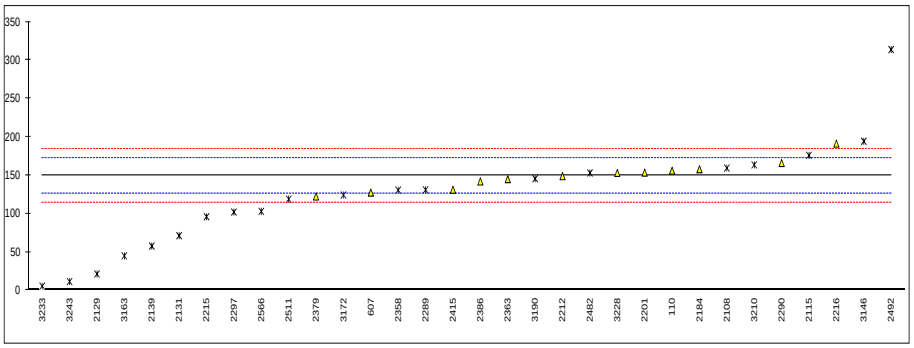
APPENDIX 1

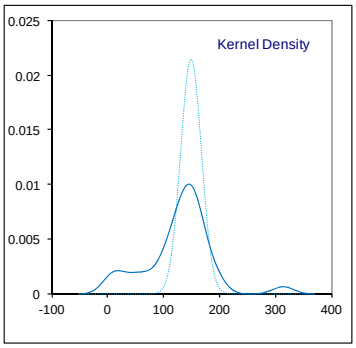
Determination of Tris(2-chloro-ethyl)phosphate (TCEP) in sample #15007; results in mg/kg

lab	method	value	mark	z(targ)	remarks
110	INH-3352	156.01		0.55	
551		-----		-----	
607	in house	127.44		-1.90	
2108	in house	159.4	ex	0.84	
2115	in house	175.98	ex, C	2.26	First reported: 1030.493
2129	ISO17881	21.45	ex	-10.98	
2131	in house	71.28	ex	-6.71	
2139	EN71	57.8	G(0.01)	-7.87	
2184	in house	157.9		0.71	
2201	in house	153.5		0.34	
2212	in house	148.9		-0.06	
2215	in house	96.0	ex	-4.59	
2216	in house	191.33		3.58	
2289	EN71	131	ex	-1.59	
2290	in house	166.012		1.41	
2297	in house	102.1	ex	-4.07	
2358	in house	130.705	ex	-1.62	
2363	INH-780	145.0		-0.39	
2379	in house	122.42		-2.33	
2386	in house	142		-0.65	
2413		-----		-----	
2415	in house	131.2		-1.57	
2482	in house	153.027	ex	0.30	
2492	in house	313.8	ex	14.08	
2493		-----		-----	
2511	GB/T4279	119.0	ex	-2.62	
2566	in house	103	ex	-3.99	
3146	in house	194.3	ex	3.83	
3163	INH-GCMS	45	ex	-8.96	
3172	GB/T24279	124.3	ex	-2.17	
3190	EN71	145.49	ex	-0.35	
3210	in house	163.6	ex	1.20	
3228	in house	153.1		0.30	
3233	in house	5.6	ex, C	-12.34	First reported: 6.53
3238	in house	n.d.		-----	
3243	INH-GC/MS	11.52	C,G(0.05)	-11.83	First reported: 6.91

				All data (no exclusions)	
	normality	suspect		OK	
	n	12		31	
	outliers	2 (+18ex)		1	
	mean (n)	149.568		122.754	
	st.dev. (n)	18.5765		50.2578	
	R(calc.)	52.014		140.722	
	R(EN71-11:05)	32.666	R(Horwitz) = 31.533	26.422	R(Horwitz) = 26.661

Laboratories that did not reduce the sample size by cutting or grinding were excluded, see also §4.1 and §5.





APPENDIX 2

Analytical details

lab	grinded/cut	Size (mm)	extraction solvent	time	detection technique	recovery
110	Cut	1-2 mm	Toluene	1 hr	GC/MS	Yes, 98%
551	-	-	-	-	-	-
607	Cut	<500 µm	Toluene	1 hr	GC/MS	Yes, 90.75%
2108	-	-	THF	-	-	-
2115	No	-	m-Xylene	1 hr	LC-MS/MS	Yes
2129	No	-	Dichloromethane	-	LC-MS	No
2131	No	Yes	THF/Hexane	1 hr	GC/MS	No
2139	Yes	90 µm	Acetonitrile	1 hr	LC-MS/MS	No
2184	Cut	1x1 mm	n-Hexane/Acetone/MTBE (1:1:1)	3 hrs	GC/MS	No
2201	Cut	2x2 mm	Chlorobenzene/Methanol	0.5 hr	LC-MS/MS	No
2212	Cut	1000 µm	Toluene	1 hr	Sonication	Yes, 90%
2215	No	2x2 mm	Toluene	2 hrs	GC/MS	No
2216	Yes	<1.0 mm	THF	1 hr	GC/MS	Yes, 105.60%
2289	-	-	-	-	-	-
2290	Yes	Yes	Chlorobenzene	0.5 hr	GC/MS	Yes, 97%
2297	No	Yes	Toluene	2 hrs	Sonicate	Yes, 85%
2358	No	As received	Ethylacetate/n-Hexane (1:1)	1 hr	Sonication	Yes, 92.70%
2363	Cut	-	-	-	-	-
2379	Grinded	<1x1 mm	Ethylacetate/n-Hexane (1:1)	1 hr	GC/MS	Yes, 84.50%
2386	Yes	<1 mm	Ethylacetate/n-Hexane (1:1)	1 hr	Ultrasonic	Yes, 93%
2413	-	-	-	-	-	-
2415	Yes	1x1 mm	Toluene	1 hr	GC/MSD	Yes, 100%
2482	No	No	Toluene	1 hr	UHPLC-QQQ	No
2492	No	No	Hexane	3 hrs	GC-MS/MS	No
2493	-	-	-	-	-	-
2511	No	-	-	-	-	-
2566	No	As received	Toluene	2 hrs	GC/MS	-
3146	No	2x2 mm	Dichloromethane/Toluene	1 hr, overnight 2 hrs	Ultrasonic	No
3163	No	No	-	-	GC/MS (therm. des.)	-
3172	No	No	n-Hexane/Acetone (7:3 v/v)	0.5 hr	GC-MS/MS	No
3190	No	Yes	Chlorobenzene	0.5 hr	GC/MS	No
3210	No	2-3 mm	THF/Toluene (35/45 v/v)	1.5 hr	GC-MS/MS	Yes, 93%
3228	Cut	1x1 mm	n-Hexane/Acetone/MTBE (1:1:1)	3 hrs	GC/MS	No
3233	No	-	THF/ACN	0.5 hr	GC/MS	No
3238	No	As received	Dichloromethane	15 min.	GC/MS	No
3243	Cut	1 mm	Toluene/Acetonitril 80/20	1 hr	GC/MS	Yes, 60%

APPENDIX 3

Number of participants per country

1 lab in BRAZIL

3 labs in FRANCE

6 labs in GERMANY

4 labs in HONG KONG

1 lab in HUNGARY

1 lab in INDIA

2 labs in ITALY

1 lab in KOREA

1 lab in MALAYSIA

7 labs in P.R. of CHINA

1 lab in SWITZERLAND

1 lab in THAILAND

1 lab in THE NETHERLANDS

1 lab in TUNISIA

3 labs in U.S.A.

2 labs in VIETNAM

APPENDIX 4

Abbreviations:

C = final result after checking of first reported suspect result
D(0.01) = outlier in Dixon's outlier test
D(0.05) = straggler in Dixon's outlier test
G(0.01) = outlier in Grubbs' outlier test
G(0.05) = straggler in Grubbs' outlier test
R(0.01) = outlier in Rosner outlier test
R(0.05) = straggler in Rosner outlier test
DG(0.01) = outlier in Double Grubbs' outlier test
DG(0.05) = straggler in Double Grubbs' outlier test
n.e. = not evaluated
n.d. = not detected

Literature:

1 DIN 53316
2 ISO 17234:2010
3 EN71-11:2005
4 EN71-10:2005
5 iis-Interlaboratory Studies: Protocol for the Organisation, Statistics and Evaluation, April 2014
6 XP G 08-014:97
7 P.L. Davies, Fr Z. Anal. Chem, 351, 513, (1988)
8 W.J. Conover, Practical; Nonparametric Statistics, J. Wiley&Sons, NY, p.302, (1971)
9 ISO 5725, (1986)
10 ISO 5725, parts 1-6, (1994)
11 M. Thompson and R. Wood, J. AOAC Int, 76, 926, (1993)
12 G. Rohm, J. Bohnen & H. Kruessmann, GIT Labor-Fachzeitschrift, p 1080, 11, (1997)
13 OEKO-TEX Std 100, p19, (ed. 04/2013)
14 Bernard Rosner, Percentage Points for a Generalized ESD Many-Outlier Procedure, *Technometrics*, 25(2), pp. 165-172, (1983)
15 R.G. Visser, Reliability of proficiency test results for metals and phthalates in plastics, Accred Qual Assur (2009) 14:29–34