Asia Pacific Laboratory Accreditation Cooperation

APLAC T079 Phthalates in Plastic materials Proficiency Testing Program

Final Report

November 2011

ACKNOWLEDGEMENTS

Taiwan Accreditation Foundation (TAF) wishes to acknowledge the funding supports by Asia Pacific Laboratory Accreditation Cooperation (APLAC), and also thanks to APLAC members and the laboratories they nominated for their participation and cooperation.
TAF would like to appreciate Mr. Sheng-nan Liu and Ms. Chyong-rong Chen from BSMI and relevant experts for their professional works on this program, preparing samples and technical advises.
TAF would be grateful for PIDC to be as the reference laboratory conducted the homogeneity tests.

CONTENTS

APPENDICES

1. Introduction

This report summarizes the results of **APLAC T079 Phthalates in Plastic materials** Proficiency Testing Program involving testing laboratories for the Asia Pacific Laboratory Accreditation Cooperation (APLAC).

2. Program Feature

(a) APLAC, EA and ILAC members were invited to participate in this proficiency testing program. There were in total 31 laboratory participants nominated by 11 accreditation bodies (ABs) in this program. All samples were sent to the ABs and then transferred to the participating laboratories.

Table 1 Summary of Economies and Participants

Economy	No. of laboratories nominated	No. of laboratories returning result
Hong Kong, China	3	3
Japan	4	4
Thailand (TISI)	3	3
Thailand (BLA-DSS)	1	1
United States (L-A-B)	1	1
KOREA	3	3
United States (A2LA)	4	4
Taiwan	4	4
India	2	2
United States (ACLASS)	3	3
China	3	3
Total	**31**	**31**

(b) Participating laboratories were supplied two amber sample bottles. Each bottle contains about 7 grams of PVC pellets.

(c) Each participating laboratory was required to analysis the phthalate concentrations of **DEHP and DNOP**.

(d) Laboratories were requested to complete the test procedures for phthalates in plastic according to their copy of the Instructions to Participants and to record their results on the accompanying Results Sheet, both of which were distributed to participants with the samples. (The Instructions and the Results Sheet are shown in Appendix C.)

(e) Each laboratory was randomly allocated a unique code number for the program to enable confidentiality of results. Its code number makes reference to each laboratory in this report.

3. Content of Appendices

APPENDIX A contains:

(a) The results and measurement uncertainties of test samples reported by the participants and the calculated Z-Score result for each test on each sample.

(b) A table of robust statistics for each test, such as number of results, median, normalized IQR, Robust CV, minimum, maximum and range.

(c) Z-Scores and ordered Z-Score charts calculated for laboratories for each test.

(d) Youdon Plots of Sample A and Sample B

(e) The information of test method reported for each laboratory.

APPENDIX B describes the sample preparation details and the homogeneity/ stability testing procedure of test samples.

APPENDIX C contains a copy of the *Instructions to Participants* and the *Results Sheet*, as supplied to participants.

APPENDIX D shows the details of statistical procedures and the calculations and formulae of statistic parameters.

4. Statistical design of the program

A split level statistical design was used in this program. The two samples labeled Sample A and Sample B (sample pair) were assumed to be identical.

The robust statistical procedures were used to generate the z-scores and a summary of statistics for the sample-number of results, original relative standard deviation, final relative standard deviation, robust average (x^*) and robust standard deviation (s^*).

For each laboratory, the between-laboratories z-score and the within-laboratories z-score were calculated for the sample pair. These between-laboratories z-scores were based on the (standardized) sum of the calculated mean of the duplicate or triplicate results reported for each sample. The within-laboratories z-scores were based on the (standardized) difference between the mean of the duplicate or triplicate results reported for each sample.

For each laboratory, a single robust Z-score was calculated for each returned result. For each item, we can use the Z-score of each laboratory to understand the degree how the result of that laboratory deviates from the population. If the absolute Z score of the testing result is larger than three ($|Z-Score| \geq 3$), it represents that the testing result of that laboratory is an outlier and the laboratory will be denoted by "※". For the testing result of the laboratory with $|Z-score| \geq 3$, one should make further assessments for the test procedure.

5. Assigned values and uncertainty

Assigned values were robust average (x^*) calculated from participating laboratories results based on robust statistics in ISO 13528 annex C robust analysis: algorithm A.

Assigned values (x^*) of uncertainty were estimated from robust standard deviation(s^*) of robust statistics based on $1.25s^*/\sqrt{n}$, n was 31.

	DNEP (%)		DNOP (%)	
	Sample A	Sample B	Sample A	Sample B
Assigned value(x^*)	0.768	0.782	0.731	0.745
Standard deviation(s^*)	0.066	0.074	0.121	0.132
Uncertainty (95 %)	0.015	0.017	0.027	0.030

6. Outlier Results

A Z-Score tells how a single data point compares to normal data. A Z-Score says not only whether a point was above or below average, but how unusual the measurement is. The outliers for the feedback results of the participating laboratories are shown in Table 2 as below; and the detailed information of the test results, statistics and test procedures were referred to the Summary of Results in Appendix A.

Table 2 List of Outliers

$\lvert Z-score\rvert \geq 3$	Lab Code	The Proportion of Outliers (%)
Sample A		
DEHP	12, 26, 31	9.8% (3/31)
DNOP	26, 27	6.5% (2/31)
Sample B		
DEHP	12, 27, 31	9.8% (3/31)
DNOP	27	3.2% (1/31)

7. Technical Comments

In this phthalate proficiency test program, there were 31 laboratories registered to participate. The number of participating laboratories and their nationalities were listed in Table 1. After completing the registration procedure, each participant had received two test samples in this proficiency test program. In each sample, there were two kinds of di-(2-ethylhexyl) phthalate (DEHP) and Di-n-octyl phthalate (DNOP) to be added. The both concentrations of DNOP and DEHP were analyzed by participating laboratories according to CPSC test method (CPSC-CH-C1001-09.3) and the analysis data were reported from participant laboratories. A summary of test results for all participating laboratories was listed in Table 3. The concentrations of DEHP and DNOP phthalate in two test samples were shown in Table 4 separately and their statistics parameters for two kinds of phthalate were also listed in Table 5. The Z-score histograms and Z-score bar chart of test results for two kinds of phthalate were drawn in figure 1 and figure 2.

Each participating laboratory received two test samples assumed to be identical. From table 4 in appendix A, we calculated the difference between concentrations of phthalate (DEHP and DNOP) in two test samples and then compared with the reported measurement uncertainties from the participating laboratories. It can be found that there are 8 laboratories in DEHP results when the absolute difference between test sample A and B is larger than the measurement uncertainties reported by the participants and there are 7 laboratories in DNOP results when the absolute difference between test sample A and B is larger than the measurement uncertainties reported by the participants. All these laboratories should investigate for the test procedure and uncertainty budget.

A Z-Score is a statistical measure. When test population is in a normal distribution, the distribution ratio is 99.73% which absolute value of Z score is below 3. The Z score could be estimated the deviation degree between test results and the mean of the sample.

In this proficiency test program, both results of a split level of two samples are calculated to generate the between-laboratory z-score (ZB) and within-laboratory z-score (ZW). Youden plots are used to illustrate the data. These plots consist of a rectangular plot on which laboratory is represented by a point. The horizontal component of each point is the result of laboratory on the Sample A of a pair while the vertical component is its result on Sample B. A 95% confidence ellipse is drawn based on an assumption of variaty normal distribution of the results. Points that are

outside this ellipse are statistically significantly different than the others. If the result is outside the ellipse in the upper left or lower right quadrant, the error is likely to be caused by poor repeatability (random error); if a result is outside the ellipse in the upper right or lower left quadrant, the error is systematic, or bias.

From outlier results in section 5, there were three laboratories results with outliers in DEHP analysis of two samples. The DEHP Z score of laboratory code 12 was less than -3 and the DEHP Z score of laboratory code 31 was more than 3 in both test samples. The DEHP Z score of laboratory code 27 was less than -3 in sample B and the DEHP Z score of laboratory code 26 was more than 3 in sample A. There are two laboratories results with outliers in DNOP analysis of sample A and one laboratory result with outlier in DNOP analysis of sample B. The DNOP Z score of laboratory code 27 was more than 3 in both test samples. The DNOP Z score of laboratory code 26 was equal to 3 in sample A.

In DEHP concentration analysis, the between-laboratories z-score (ZB) showed two laboratories. The ZB score of laboratory code 12 was less than -3 and the ZB score of laboratory code 31 was more than 3. In DEHP analysis, the within-laboratory z-score (ZW) showed no laboratories with z-score more than 3. In DNOP concentration analysis, the between-laboratories z-score (ZB) showed one laboratory with z-score greater than three. The ZB score of laboratory code 27 was more than 3. The within-laboratory z-score (ZW) showed three laboratories with z-score greater than three. The ZW score of laboratory code 14, 26 and 27 were more than 3.

From the above analysis, the laboratory codes 12 and 31 showed that there would be serious systematic errors for DEHP test in these laboratories. The laboratory codes 26 and 27 showed that there would be random errors for DEHP test in these laboratories. Besides, in the part of DNOP test, there would be random errors in these laboratory codes 14 and 26. The systematic error and random error existed simultaneously in the laboratory code 27.

Appendix A

Summary of Results

Information of Test

Table 3 Laboratories' Test Results for DEHP of Sample A and Sample B

Item	Sample A		Sample B	
Lab Code	DEHP (%)	Measurement Uncertainty	DEHP (%)	Measurement Uncertainty
1	0.729	0.0048	0.721	0.0048
2	0.917	0.053	0.846	0.046
3	0.724	0.0666	0.738	0.0679
4	0.822	0.041	0.797	0.043
5	0.772	-	0.763	-
6	0.765	±0.056 (K=2)	0.781	±0.075 (K=2)
7	0.652	±0.0200	0.635	±0.0200
8	0.762	0.042	0.772	0.042
9	0.753	0.060	0.758	0.063
10	0.705	-	0.655	-
11	0.790	-	0.862	-
12	0.453	0.006	0.530	0.006
13	0.927	0.044	0.925	0.033
14	0.832	±0.013	0.968	±0.013
15	0.800	0.190	0.808	0.190
16	0.792	±0.015	0.794	±0.015
17	0.819	0.257	0.823	0.258
18	0.765	0.064	0.791	0.064
19	0.742	-	0.736	-
20	0.671	0.193	0.720	0.193
21	0.828	0.118	0.839	0.116
22	0.739	±0.111	0.814	±0.120
23	0.821	0.055 (95% confidence, k=2.16)	0.867	0.071 (95% confidence, k=2.20)
24	0.752	±0.057%	0.816	±0.061%
25	0.746	-	0.742	-
26	1.000	±0.045	0.865	±0.039
27	0.630	2% of reported value	0.530	2% of reported value
28	0.691	0.0096	0.625	0.0096
29	0.769	0.093	0.773	0.094
30	0.760	0.046%, k=2	0.755	0.045%, k=2
31	1.173	±15%	1.050	±15%

Note:"-" denotes result not provide.

Table 3 (Cont'd) Laboratories' Test Results for DNOP of Sample A and Sample B

Item	Sample A		Sample B	
Lab Code	DNOP (%)	Measurement Uncertainty	DNOP (%)	Measurement Uncertainty
1	0.648	0.0043	0.657	0.0043
2	0.856	0.053	0.829	0.047
3	0.799	0.0696	0.825	0.0719
4	0.980	0.050	0.949	0.051
5	0.606	-	0.629	-
6	0.666	±0.048 (K=2)	0.677	±0.068 (K=2)
7	0.648	±0.0205	0.656	±0.0205
8	0.662	0.218	0.679	0.022
9	0.685	0.052	0.685	0.057
10	0.614	-	0.571	-
11	0.757	-	0.846	-
12	0.460	0.005	0.533	0.005
13	0.891	0.042	0.891	0.033
14	0.739	±0.020	0.892	±0.020
15	0.752	0.190	0.749	0.190
16	0.761	±0.015	0.763	±0.015
17	0.858	0.200	0.876	0.204
18	0.648	0.076	0.687	0.076
19	0.655	-	0.642	-
20	0.474	0.184	0.505	0.184
21	0.748	0.173	0.754	0.174
22	0.663	±0.100	0.708	±0.106
23	0.736	0.066 (95% confidence, k=2.20)	0.765	0.065 (95% confidence, k=2.22)
24	0.700	±0.053%	0.764	±0.058%
25	0.648	-	0.662	-
26	1.090	±0.014	0.888	±0.012
27	1.720	2% of reported value	1.470	2% of reported value
28	0.670	0.0139	0.609	0.0139
29	0.755	0.115	0.751	0.114
30	0.745	0.037%, k=2	0.741	0.037%, k=2
31	0.955	±15%	0.923	±15%

Note:"-" denotes result not provide.

Table 4 Laboratories' DEHP Result for Two Samples

Lab Code	DEHP (%)		
	Sample A	Sample B	Absolute Difference between A and B
1	0.729	0.721	0.008*
2	0.917	0.846	0.071*
3	0.724	0.738	0.014
4	0.822	0.797	0.025
5	0.772	0.763	0.009
6	0.765	0.781	0.016
7	0.652	0.635	0.017
8	0.762	0.772	0.010
9	0.753	0.758	0.005
10	0.705	0.655	0.050
11	0.790	0.862	0.072
12	0.453	0.530	0.077*
13	0.927	0.925	0.002
14	0.832	0.968	0.136*
15	0.800	0.808	0.008
16	0.792	0.794	0.002
17	0.819	0.823	0.004
18	0.765	0.791	0.026
19	0.742	0.736	0.006
20	0.671	0.720	0.049
21	0.828	0.839	0.011
22	0.739	0.814	0.075
23	0.821	0.867	0.046
24	0.752	0.816	0.064*
25	0.746	0.742	0.004
26	1.000	0.865	0.135*
27	0.630	0.530	0.100*
28	0.691	0.625	0.066*
29	0.769	0.773	0.004
30	0.760	0.755	0.005
31	1.173	1.050	0.123

Note: " * " shows that the absolute difference between test sample A and B is larger than the measurement uncertainties reported by the participants.

Table 4 (Cont'd) Laboratories' DNOP Result for Two Samples

Lab Code	DNOP (%)		
	Sample A	Sample B	Absolute Difference between A and B
1	0.648	0.657	0.009*
2	0.856	0.829	0.027
3	0.799	0.825	0.026
4	0.980	0.949	0.041
5	0.606	0.629	0.023
6	0.666	0.677	0.011
7	0.648	0.656	0.008
8	0.662	0.679	0.017
9	0.685	0.685	0.000
10	0.614	0.571	0.043
11	0.757	0.846	0.089
12	0.460	0.533	0.073*
13	0.891	0.891	0.000
14	0.739	0.892	0.153*
15	0.752	0.749	0.003
16	0.761	0.763	0.002
17	0.858	0.876	0.018
18	0.648	0.687	0.039
19	0.655	0.642	0.013
20	0.474	0.505	0.031
21	0.748	0.754	0.006
22	0.663	0.708	0.045
23	0.736	0.765	0.029
24	0.700	0.764	0.064*
25	0.648	0.662	0.014
26	1.090	0.888	0.202*
27	1.720	1.470	0.250*
28	0.670	0.609	0.061*
29	0.755	0.751	0.004
30	0.745	0.741	0.004
31	0.955	0.923	0.032

Note: " * " shows that the absolute difference between test sample A and B is larger than the measurement uncertainties reported by the participants.

Table 5 Z-Scores of Laboratories for DEHP

Lab Code	Sample A	Sample B	AB-SUM	AB-DIFF
	Z-Score	Z-Score	ZB-Score	ZW-Score
1	-0.6	-0.8	-0.7	-0.2
2	2.3	0.9	1.7	1.2
3	-0.7	-0.6	-0.6	-0.1
4	0.8	0.2	0.6	0.2
5	0.1	-0.3	-0.1	-0.2
6	0.0	0.0	0.0	0.0
7	-1.8	-2.0	-2.0	0.0
8	-0.1	-0.1	-0.1	-0.2
9	-0.2	-0.3	-0.3	-0.3
10	-1.0	-1.7	-1.4	0.7
11	0.3	1.1	0.8	1.2
12	-4.8 ※	-3.4 ※	-4.3 ※	1.3
13	2.4	1.9	2.3	-0.3
14	1.0	2.5	1.9	2.6
15	0.5	0.4	0.5	-0.2
16	0.4	0.2	0.3	-0.3
17	0.8	0.6	0.7	-0.3
18	0.0	0.1	0.1	0.2
19	-0.4	-0.6	-0.5	-0.2
20	-1.5	-0.8	-1.2	0.7
21	0.9	0.8	0.9	-0.1
22	-0.4	0.4	0.1	1.3
23	0.8	1.2	1.1	0.6
24	-0.2	0.5	0.2	1.0
25	-0.3	-0.5	-0.4	-0.3
26	3.5 ※	1.1	2.4	2.6
27	-2.1	-3.4 ※	-2.9	1.8
28	-1.2	-2.1	-1.8	1.1
29	0.0	-0.1	0.0	-0.3
30	-0.1	-0.4	-0.2	-0.3
31	6.2 ※	3.6 ※	5.2 ※	2.3

| Z—Score | ≧3 are marked by "※"

Note:"-" not applicable

DEHP	Sample A	Sample B
Original RSD (%)	15.5	14.2
Final RSD (%)	8.6	9.5
x^*	0.768	0.782
s^*	0.066	0.074

DNOP	Sample A	Sample B
Original RSD (%)	29.2	22.9
Final RSD (%)	16.6	17.7
x^*	0.731	0.745
s^*	0.121	0.132

RSD denotes Relative Standard Deviation.

" x^* " denotes Robust average.

" s^* " denotes Robust standard deviation.

	AB-SUM	AB-DIFF
No. of result	31	31
Median (%)	1.093	0.012
Normalized IQR (%)	0.093	0.032
Robust CV (%)	8.5	268.2
Maximum (%)	1.572	0.096
Minimum (%)	0.695	0.001
Range (%)	0.877	0.095

Table 5 (Cont'd) Z-Scores of Laboratories for DNOP

Lab Code	Sample A	Sample B	AB-SUM	AB-DIFF
	Z-Score	Z-Score	ZB-Score	ZW-Score
1	-0.7	-0.7	-0.8	-0.6
2	1.0	0.6	0.8	0.0
3	0.6	0.6	0.6	0.0
4	2.1	1.5	1.9	0.2
5	-1.0	-0.9	-1.1	-0.1
6	-0.5	-0.5	-0.6	-0.6
7	-0.7	-0.7	-0.8	-0.7
8	-0.6	-0.5	-0.6	-0.3
9	-0.4	-0.5	-0.5	-1.0
10	-1.0	-1.3	-1.3	0.6
11	0.2	0.8	0.5	2.4
12	-2.2	-1.6	-2.1	1.8
13	1.3	1.1	1.2	-1.0
14	0.1	1.1	0.6	4.8 ※
15	0.2	0.0	0.1	-0.9
16	0.2	0.1	0.2	-0.9
17	1.0	1.0	1.0	-0.3
18	-0.7	-0.4	-0.6	0.5
19	-0.6	-0.8	-0.8	-0.5
20	-2.1	-1.8	-2.1	0.2
21	0.1	0.1	0.1	-0.8
22	-0.6	-0.3	-0.5	0.7
23	0.0	0.2	0.1	0.1
24	-0.3	0.1	-0.1	1.4
25	-0.7	-0.6	-0.7	-0.5
26	3.0 ※	1.1	2.1	6.7
27	8.2 ※	5.5 ※	7.2 ※	8.5 ※
28	-0.5	-1.0	-0.9	1.3
29	0.2	0.0	0.1	-0.8
30	0.1	0.0	0.0	-0.8
31	1.8	1.3	1.7	0.2

| Z−Score | ≧3 are marked by "※"

Note:"-" not applicable

DEHP	Sample A	Sample B
Original RSD (%)	15.5	14.2
Final RSD (%)	8.6	9.5
x^*	0.768	0.782
s^*	0.066	0.074

DNOP	Sample A	Sample B
Original RSD (%)	29.2	22.9
Final RSD (%)	16.6	17.7
x^*	0.731	0.745
s^*	0.121	0.132

RSD denotes Relative Standard Deviation.

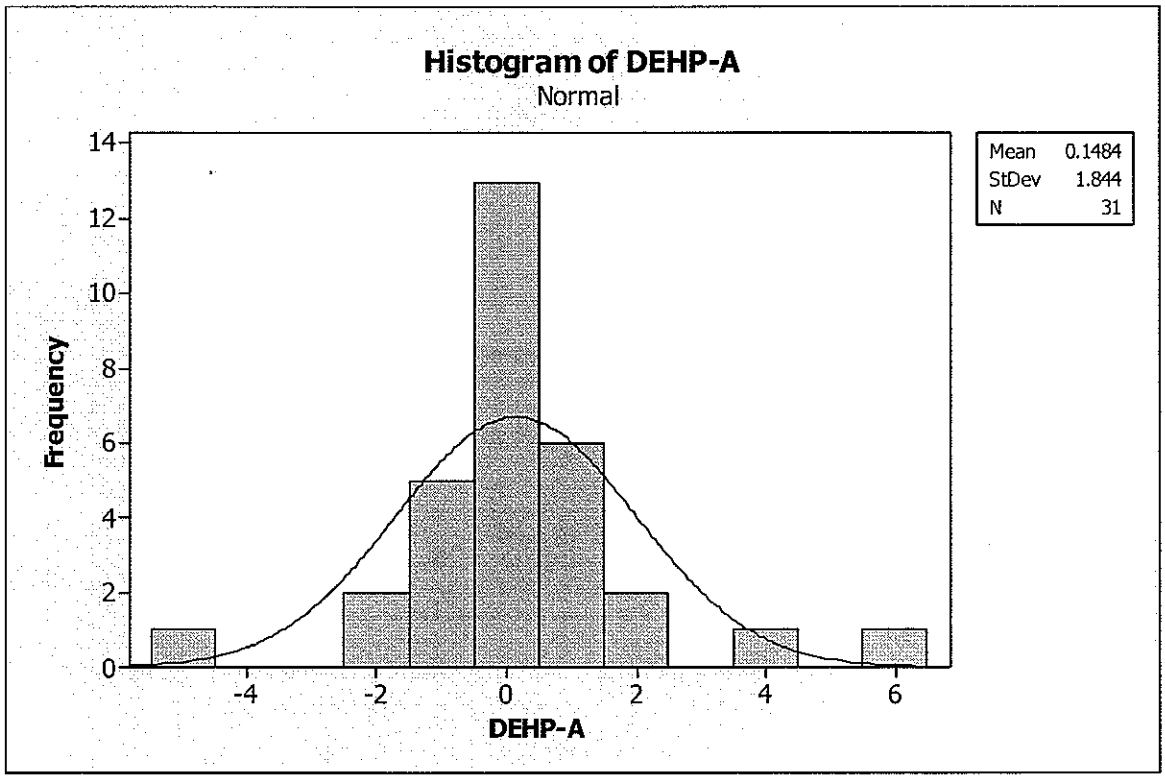
" x^* " denotes Robust average.

" s^* " denotes Robust standard deviation.

	AB-SUM	AB-DIFF
No. of result	31	31
Median (%)	1.051	0.018
Normalized IQR (%)	0.168	0.019
Robust CV (%)	16.0	101.2
Maximum (%)	2.256	0.177
Minimum (%)	0.692	0.000
Range (%)	1.563	0.177

Figure 1 Histogram of Z-Score

Sample A- DEHP



Sample B- DEHP

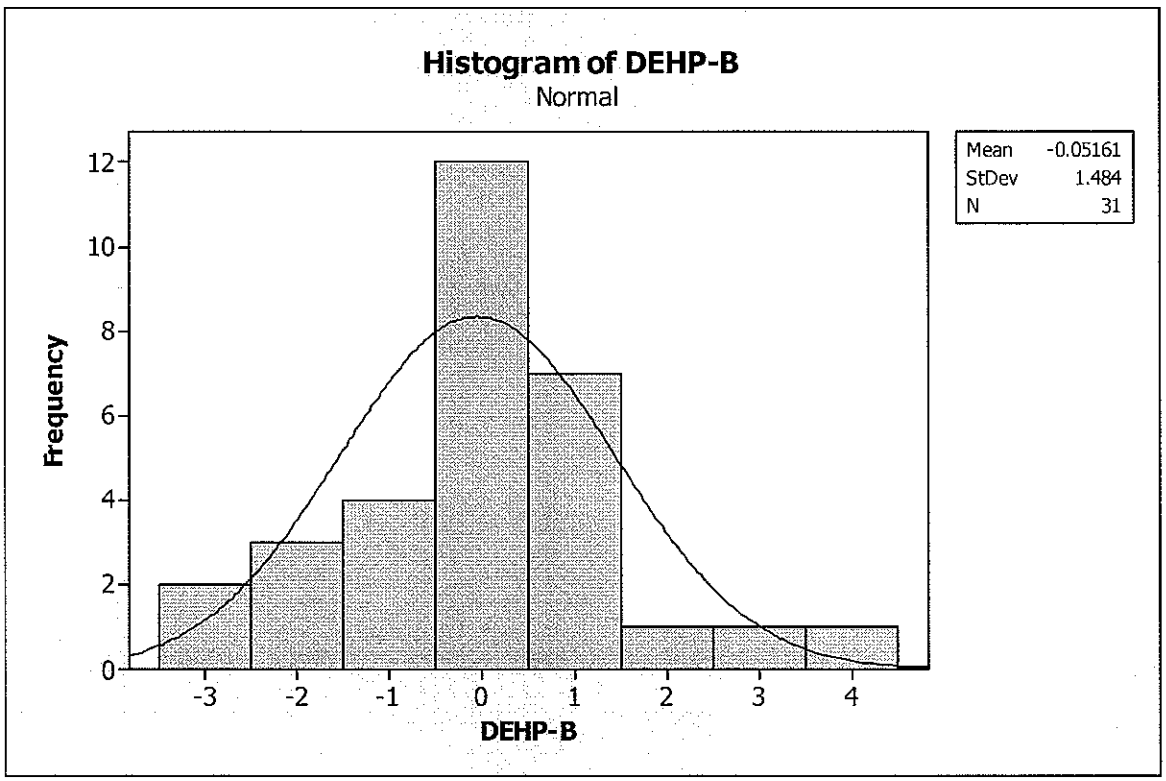
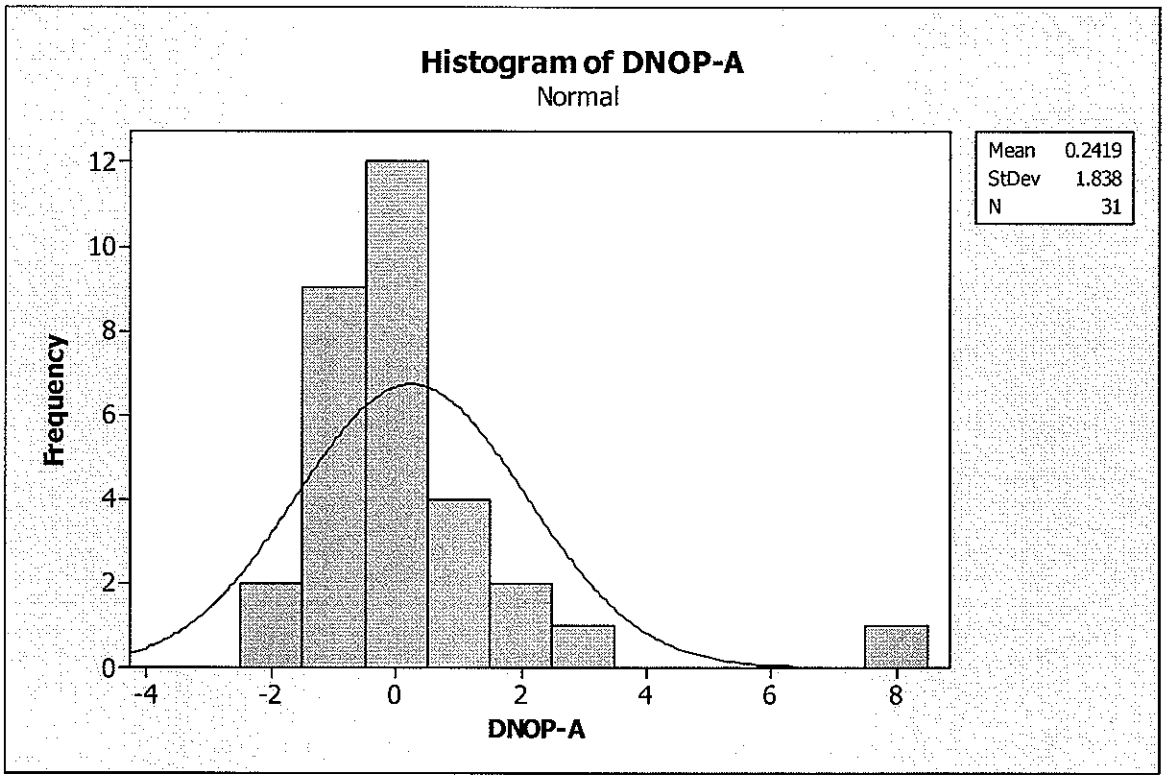


Figure 1(Cont'd) Histogram of Z-Score

Sample A- DNOP



Sample B- DNOP

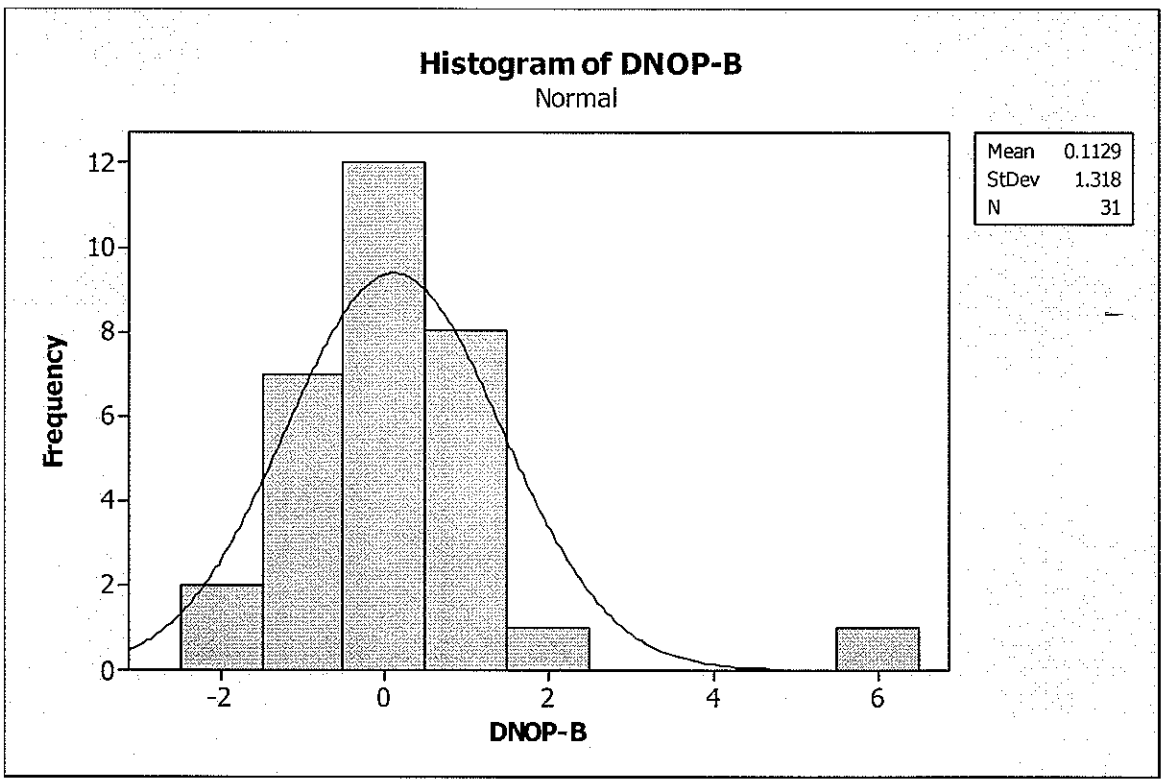
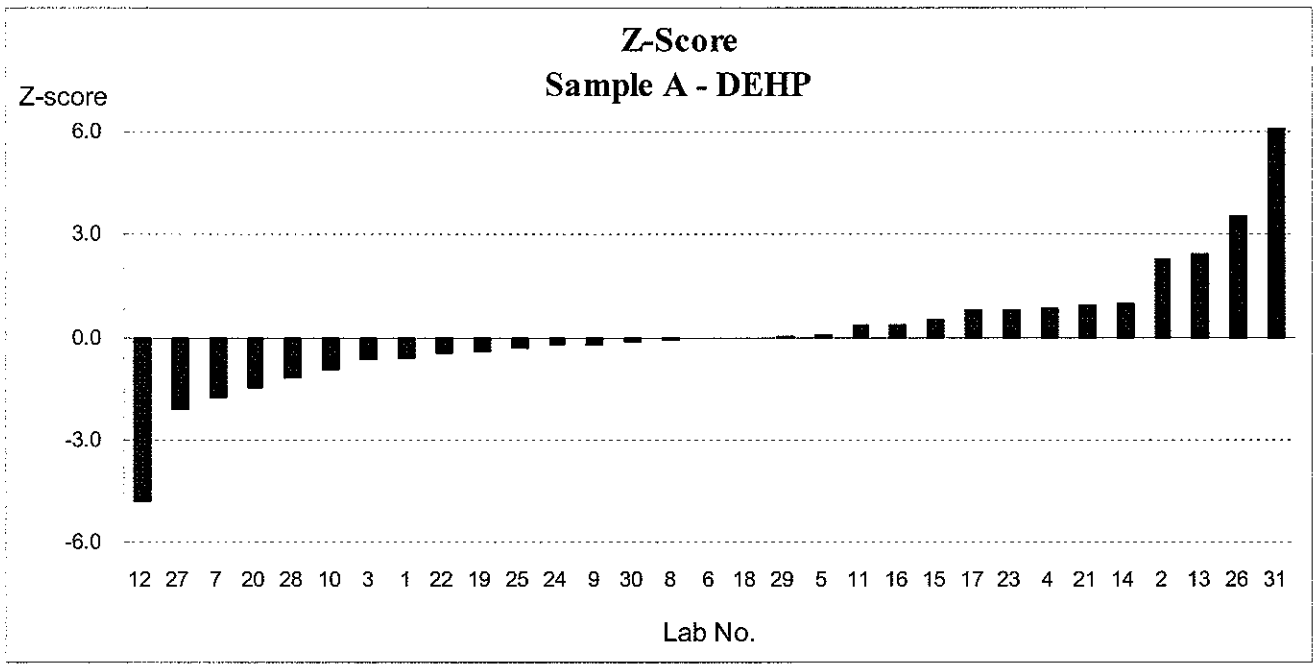


Figure 2 Bar Chart of Z-Score

Sample A- DEHP



Sample B- DEHP

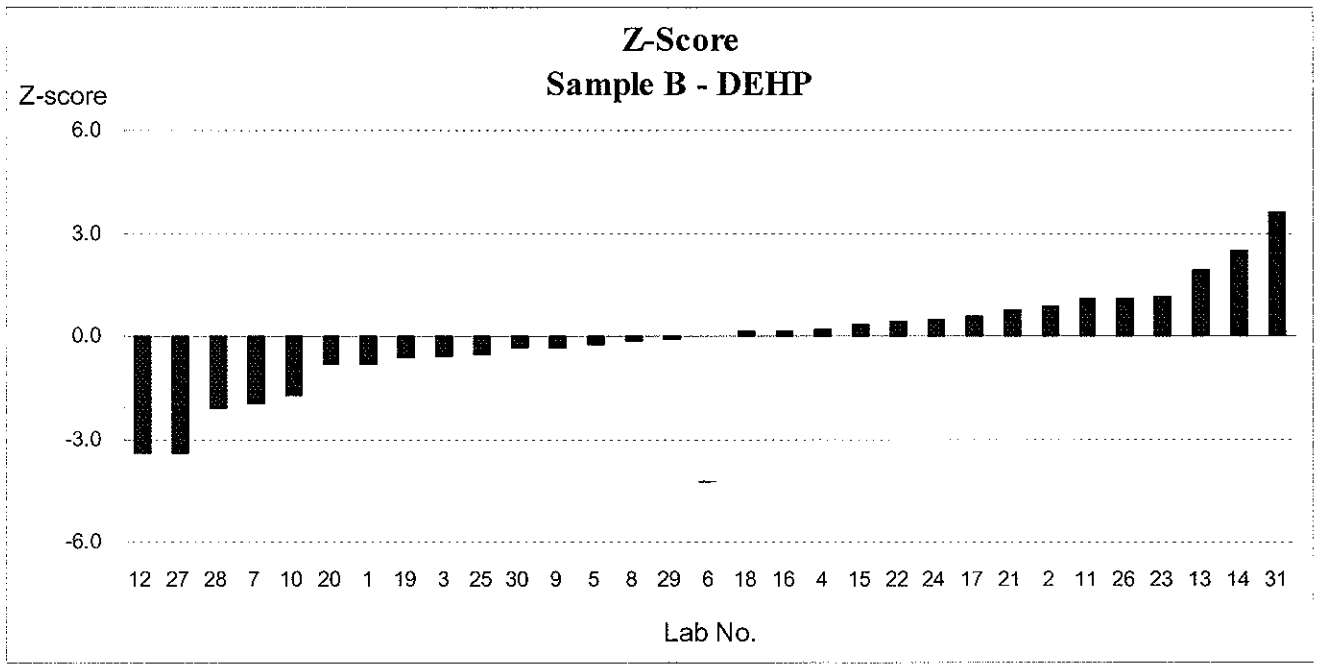
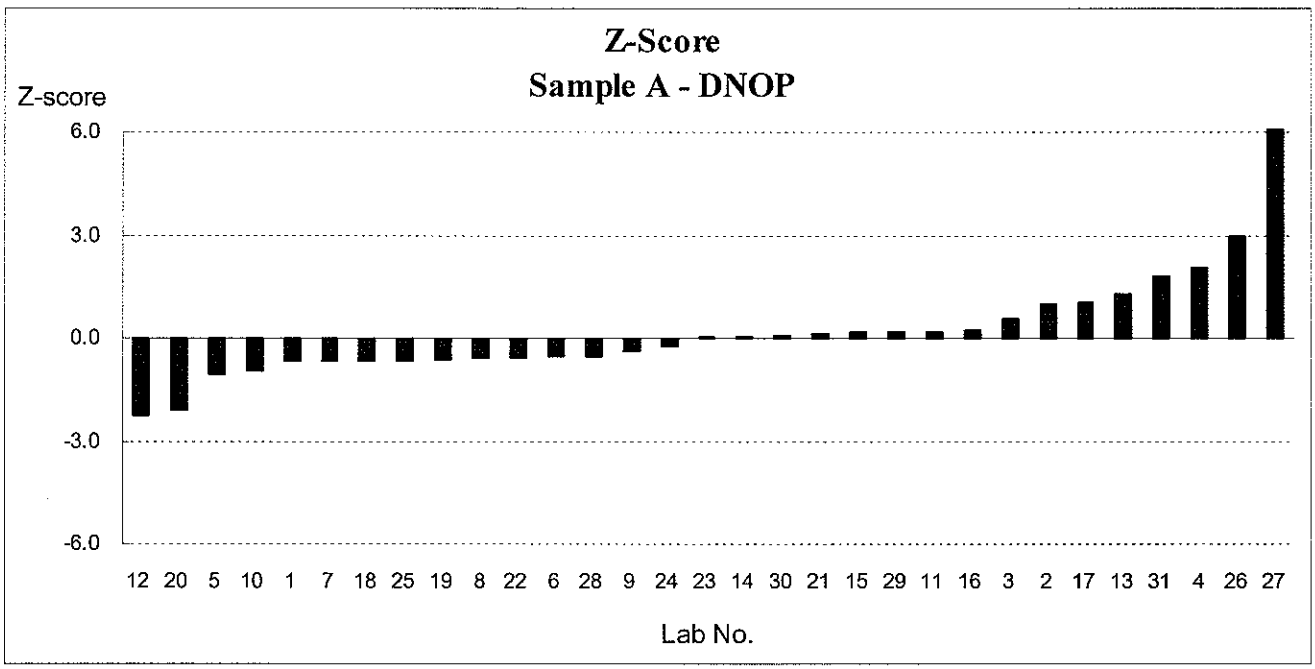


Figure 2 (Cont'd) Bar Chart of Z-Score

Sample A- DNOP



Sample B- DNOP

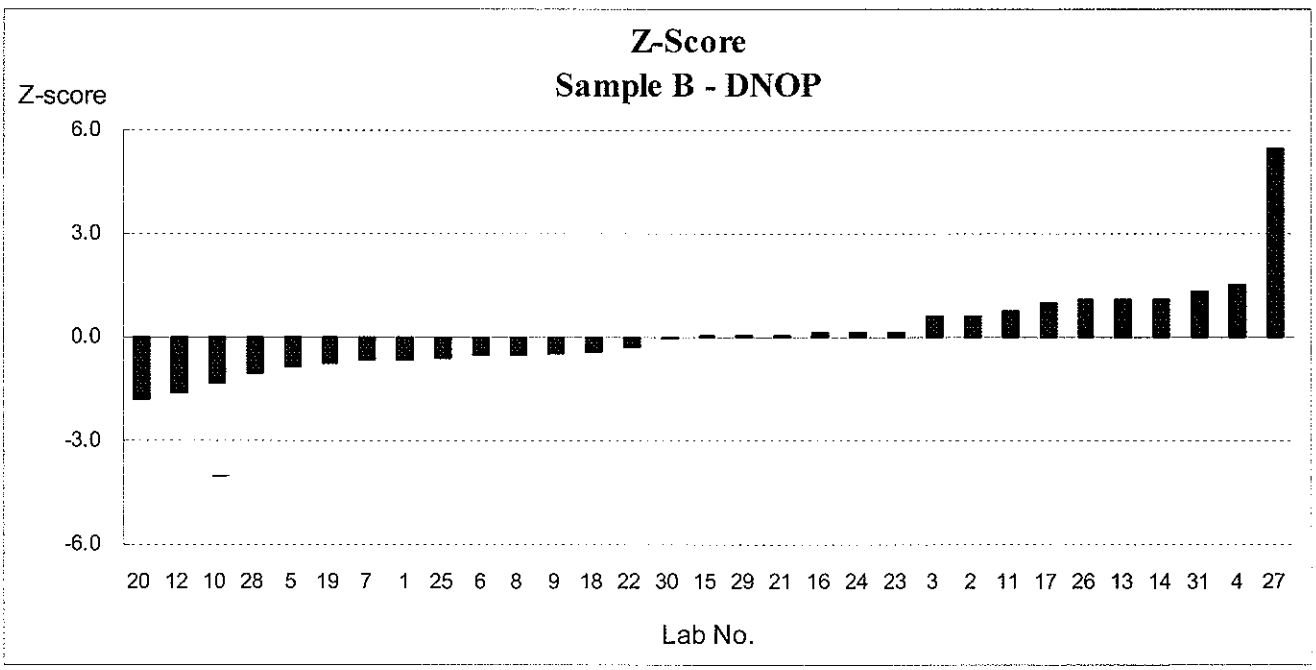
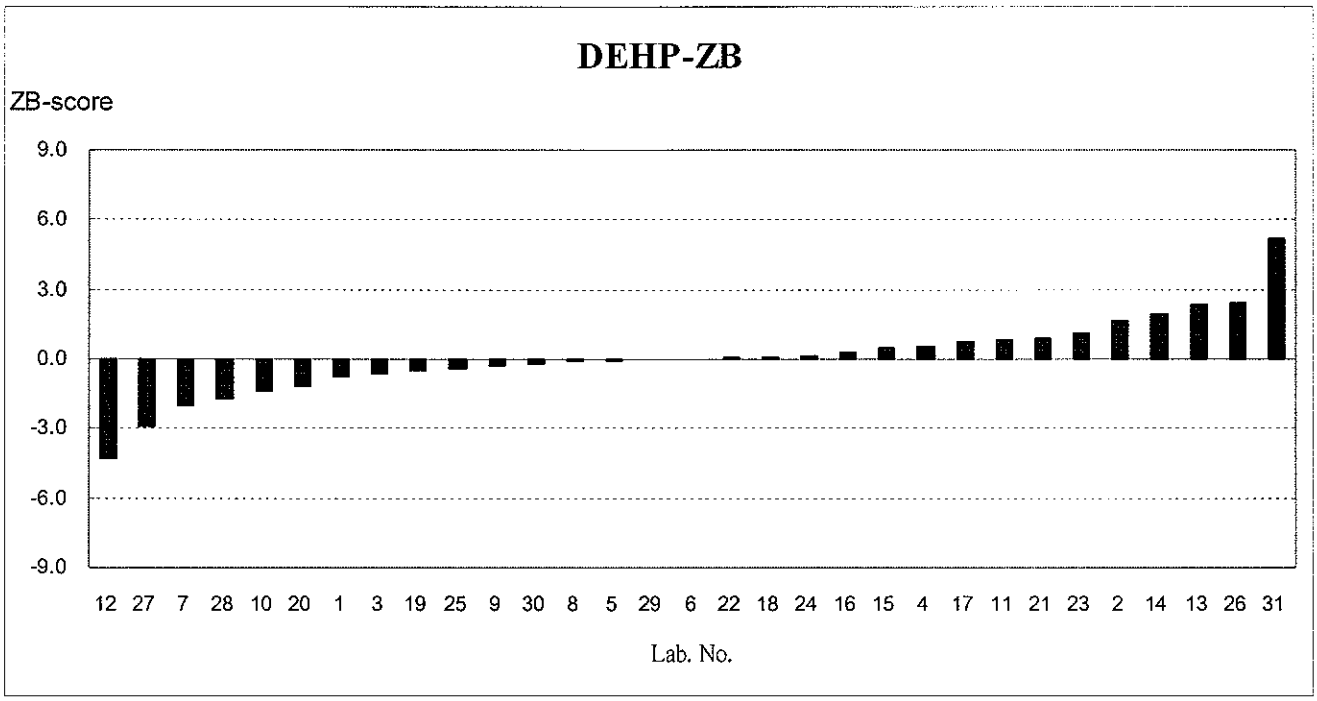


Figure 3 Bar Chart of Z-Score for between-laboratories (ZB) and within-laboratories(ZW)

DEHP



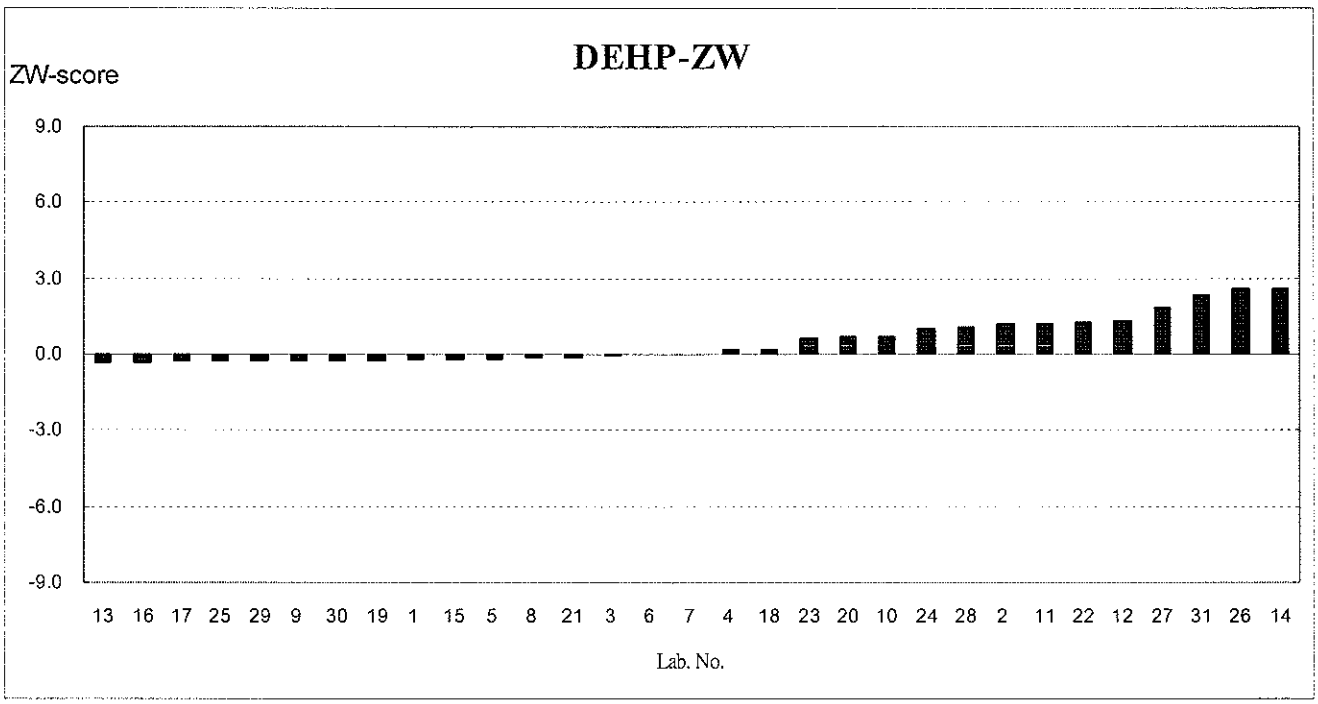
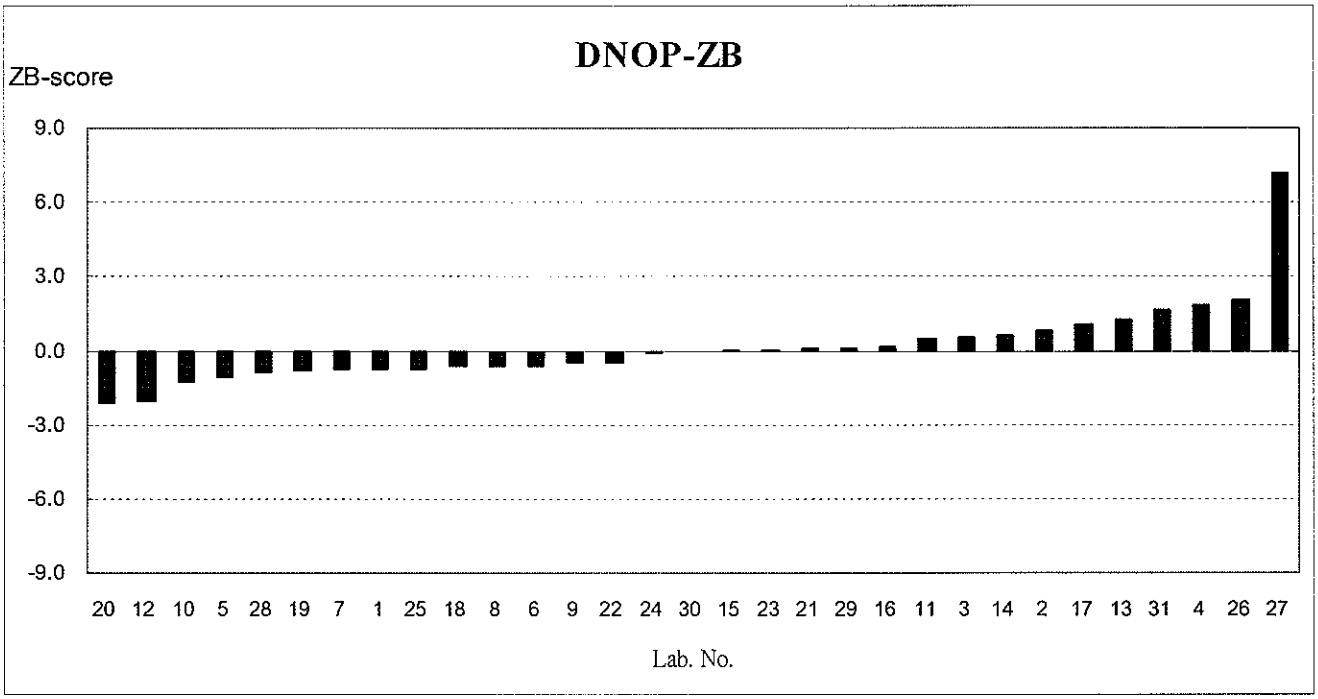


Figure 3 (Cont'd) Bar Chart of Z-Score for between-laboratories (ZB) and within-laboratories (ZW)

DNOP



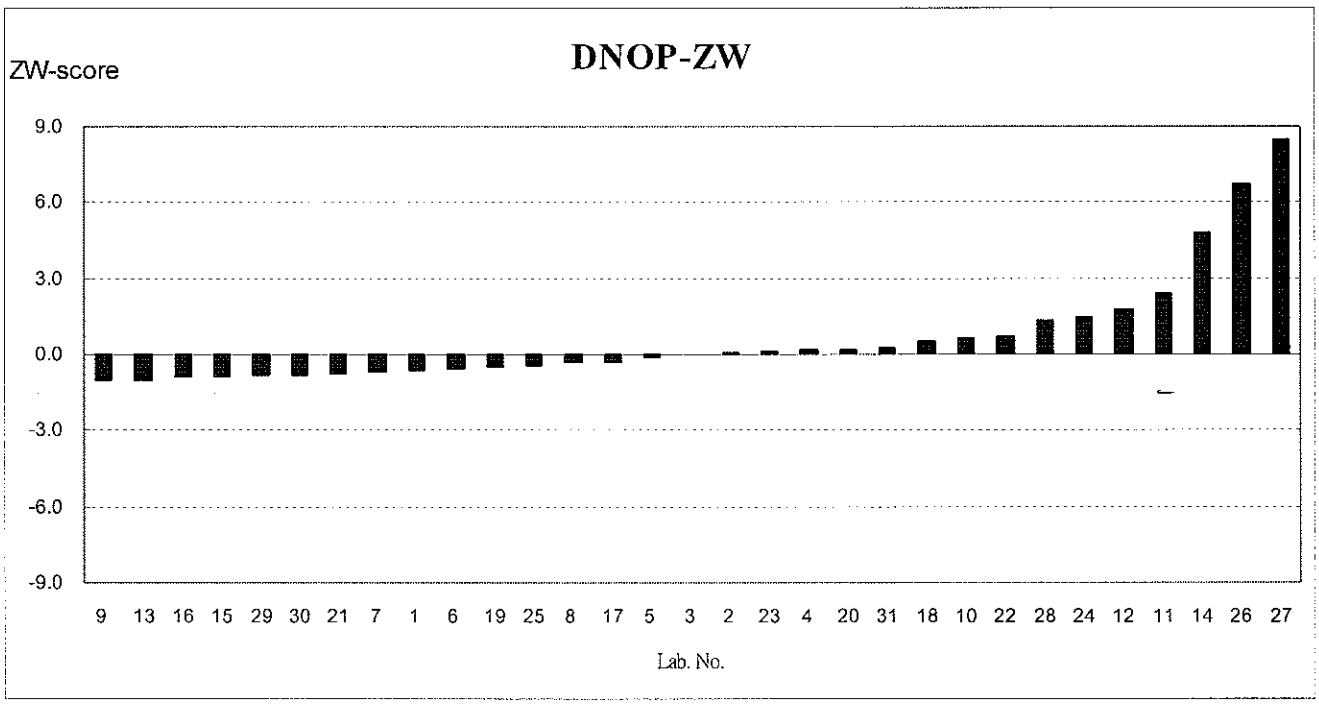


Figure 4 Youdon Plot of Sample A and Sample B

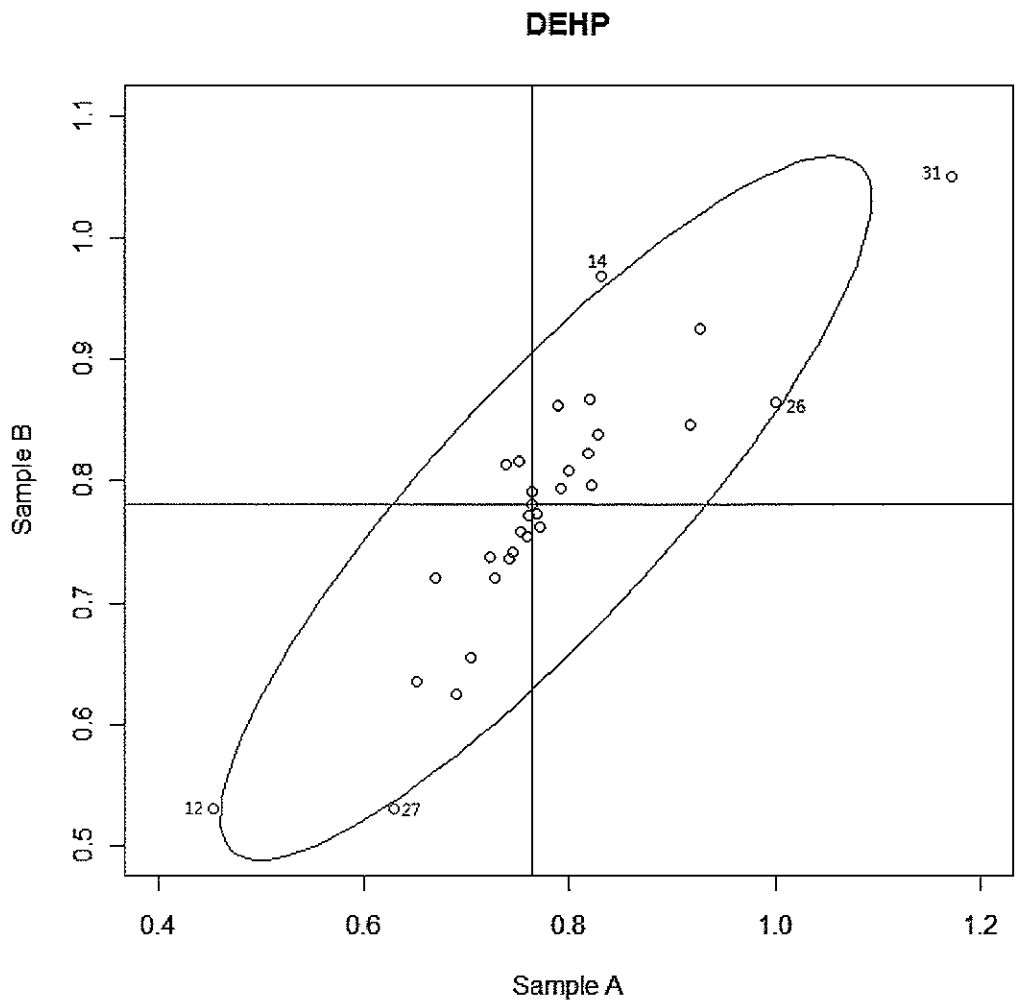


Figure 4 (Cont'd) Youdon Plot of Sample A and Sample B

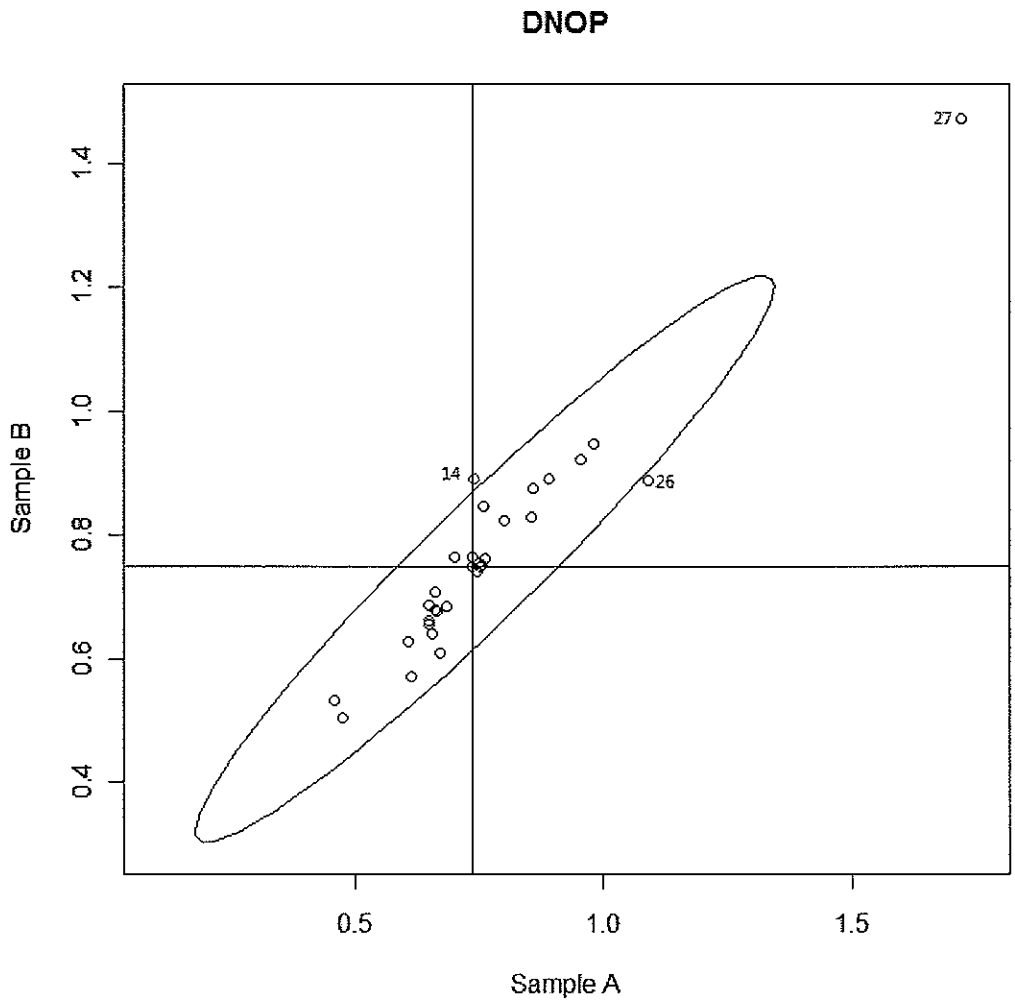


Table 5 Information of test methods in participating laboratories

Lab Code	Extraction Method	Analysis Method	Method Accreditation	Other additional information
1	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
2	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
3	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	Will be certified.
4	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	-
5	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	-
6	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	Because AREA of internal standard benzyl benzoate by a matrix effect, I used the absolute calibration method.
7	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
8	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
9	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	I undergo the conditions of extraction method and the analysis method in CPSC-CH-C1001-09.3. But to get the results, I used the isotope dilution methods with 13C-labelled Phthalates.
10	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
11	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
12	-	-		-
13	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	-
14	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
15	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
16	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	NA

Lab Code	Extraction Method	Analysis Method	Method Accreditation	Other additional information
17	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
18	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	None
19	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
20	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	Nil
21	EN 14372:2004	EN 14372:2004	Yes	In-house method base on EN 174372:2004
22	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
23	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
24	-	-	No	We used CPSC-CH-C1001-09.3/CPSC -CH-C1001-09.2 method in which we are not accredited.
25	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
26	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
27	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
28	CPSC-CH-C1001-09.3	-	Yes	Not sufficient data for calculation of uncertainty. Values derived from method detection limit study
29	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	No	The samples were ground to pass through a 500 µm sieve.
30	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-
31	CPSC-CH-C1001-09.3	CPSC-CH-C1001-09.3	Yes	-

Appendix B

Preparation of the samples

Homogeneity test

1、Production of proficiency testing samples

1.1、Material and reagent

(1) Property of material

Material of the proficiency testing sample is high stability and processability of PVC pellet. Polyvinyl chloride, commonly abbreviated as PVC, is a thermoplastic polymer. Polyvinyl chloride is the third most widely produced plastic, after polyethylene and polypropylene. It can be made softer and more flexible by the addition of plasticizers, the most widely being used "phthalates". In addition to add external plasticizers, PVC thermal stability is poor. In order to avoid PVC products from cracking when under high temperature in the process, stabilizer is usually added to reduce thermal cracking during the process, and make sure the quality of products.

Table 1.1 Basic properties of PVC (extrusion grade)

Properties	Standard	unit	3MQB000VXC0F3
Specific gravity	ASTM D792	--	1.377
Durometer Hardness	ASTM D2240	degree	84
Tensile strength	ASTM D638	Psi	6622
Impact strength	ASTM D256	ft-lb/in	1.97
Modulus of Elasticity in tension	ASTM D638	Psi	397330
Heat Deflection Temp	ASTM D648	℃	67.5

(2) Reagent of proficiency testing

The"phthalate"of proficiency testing samples is di-(2-ethylhexyl) phthalate (DEHP) and Di-n-octyl phthalate (DNOP), the formula as follows, please check table 1.2. DEHP and DnOP are high boiling point (about 386 ℃and 340 ℃) chemical substances, and PVC processing temperature is about 150 ℃~ 170 ℃ from MSDS information. The processing temperature of banbury mixer shall be controlled at 150 ℃which is much lower than the boiling point of plasticizers (DEHP and DNOP) and PVC cracking temperature. Thus, the PT-sample would not occur pyrolysis reaction under this operating temperature, enhance the stability of product by add heat stabilizer.

Table 1.2 Formula of proficiency testing samples

Sample A & B		
	Plasticizer contents	
Components	Di-n-octyl phthalate (DNOP) + PVC	di-(2-ethylhexyl) phthalate (DEHP) + PVC
Concentration	0.5~1%(w/w)	0.5~1%(w/w)

1.2 Equipment and Method

(1) Equipment

A. Equipment of pre-treatment

- Precision balance：with capability to measure to 0.001 g.

B. Compounding and processing equipment

- Brabender plastograph：A small amount of polymer mixing equipment and operating temperature up to above 250 ℃. The maximum of torque value is 100 N. Therefore, this equipment can reach the best mixing conditions quickly.
- Banbury mixer：The Banbury mixer is a brand of internal mixer batch mixer. The mixer consists of two rotating spiral shaped blades which encased in segments of cylindrical housings, so as to leave a ridge between the blades.
- Pulverizers：To smash the compounding products after compounding by Banbury mixer.

C. Analysis equipment

- Analysis equipment of phthalates：Gas Chromatography-Mass Spectrophotometer(GC/MS)
- Analysis equipment of phthalates：Gas Chromatography-Flame Ionization Detector(GC-FID)

(2) Experimental Methods

A. Pre-treatment

PVC pellets dryed at 60 ℃-70 ℃ for 1-2 h

B. Compounding and processing procedure

Before set the compounding conditions, the physical properties of the material must be confirmed, such as melting point of plastic material (Tm), pyrolysis temperature and effusion temperature ... etc. To determine the best compounding conditions by Brabender on small-scale process(compounding temperature, Screw speed... etc). This compounding condition, applied to the Banbury mixer on large-scale process, to obtain a more stable and uniform output.

1.3 Analysis

(1) Analysis of DEHP and DNOP

- Sample preparation：

Sample preparation and analytic method will reference to **CPSC-CH-C1001-09.2** of the U.S.Consumer Product Safety Commission's (CPSC).

1. Weigh out 0.20±0.005 g of sample into a sealable glass vial.
2. Add 5 ml of THF to the sample. Shake, stir, or otherwise mix sample for at least 30 minutes to allow dissolution.
3. Precipitate any PVC polymer with 10 ml of hexane for every 5 ml of THF . Allow at least 5 minutes for polymer to settle.

4. Filter THF/hexane solution through a 0.45 μm PTFE syringe filter. Combine 0.3 ml of the filtrate (THF/hexane solution) with 0.2 ml of internal standard (BB,150 μg/ml) in a GC vial, and dilute to 1.5 ml with cyclohexane.
5. A GC-MS system with an auto-sampler is suggested for the sample analysis.

- Preparation of calibration standard：

A、Solutions containing the phthalate(s) of interest in cyclohexane. Each standard should contain 20 μg/ml of internal standard BB. A minimum of four calibration standards are used. Concentration of calibration standards is 5μg/ml-40μg/ml.

B、Analyze a CRM to ensure a proper calibration. The analyzed value should be within ±15 % of the expected value. If not, prepare new standards and re-run calibration.

1.4 Sampling method

(1) Proficiency testing samples

Proficiency testing samples were divided into the 362 group sample. During the homogeneity and stability analysis, 10 samples were selected from 362 group sample.

The data of homogeneity test of proficiency testing sample were calculated by using one-way ANOVA of single variance. The test method is according to CPSC-CH-C1001-09.2 for analysis. The following table is a random sampling of the group :

Table 1.3 Table of random sampling phthalate sample

Number of analysis	Number of random sampling
1	23
2	217
3	310
4	165
5	46
6	62
7	8
8	167
9	138
10	284

2. Homogeneity test

2.1 Result of homogeneity test of proficiency testing samples

Homogeneity check of plasticizers proficiency sample is according to using one-way ANOVA of single variance. The number of extra samples needed mainly depends on the between-bottle homogeneity study. The minimum number of bottles selected at random is between 10 and 30, but should generally not be smaller than 10. Analytical method refer to CPSC-CH-C1001-09.2-Standard Operating Procedure for Determination of Phthalates of the U.S. Consumer Product Safety Commission's (CPSC) testing laboratory (LSC) for the analysis of phthalate content in child care items and toys.

The data of homogeneous test collected were analyzed by using one-way ANOVA of single variance. Certified Reference Materials ERM®-EC680k and ERM®-EC681k certification report and CNAS—GL03 Guidance on Evaluating the Homogeneity and Stability of Samples Used for Proficiency Testing are referenced. Homogeneous test for each sample tested 2 replicates under same conditions. Calculate the total average $\overline{x}$, Sum of square, between-sample and within-sample mean square of measurements of Proficiency Testing sample. The F value is calculated based on ANOVA analysis. This value is less than the critical value. The level of significant is 95%. ">0.05" shows variation between samples. "<0.05" shows unvarying between samples. Proficiency testing sample is uniform. Representation of F critical value is $F\ 0.05(f_1,f_2)$.

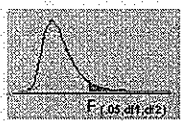
Degree of freedom：$f_1 = m - 1$ and $f_2 = N - m$. $f_1 = 9$, $f_2 = 10$

m(sample number)：10

N(test number)：20

Find the critical value from the F table for alpha=0.05. F 0.05(9,10)＝3.0204.

F Table for alpha=.05



df2/df1	1	2	3	4	5	6	7	8	9	10	12
1	161.4476	199.5000	215.7073	224.5832	230.1619	233.9860	236.7684	238.8827	240.5433	241.8817	243.9060
2	18.5128	19.0000	19.1643	19.2468	19.2964	19.3295	19.3532	19.3710	19.3848	19.3959	19.4125
3	10.1280	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130

Table 2.1 Result of homogeneity test of reference material (DEHP)

DEHP			
Analysis method：CPSC-CH-C1001-09.2			
Result (%)			
Test number / Sample number	NO.1	NO.2	Average
1	0.849	0.972	0.9105
2	0.980	0.983	0.9815
3	0.959	0.978	0.9685
4	1.030	1.020	1.0250
5	0.981	0.988	0.9845
6	0.955	0.969	0.9620
7	0.969	0.972	0.9705
8	0.977	0.995	0.9860
9	0.956	0.923	0.9395
10	0.997	0.995	0.9960
Average			0.9724
Standard deviation, SD			0.0312
The general average	0.9724		
Note	1.Analysis instrument：GC-MS 2. Sample Forms：Granular material		

Table 2.2 Result of ANOVA of reference material (DEHP)

The sources of variation	Degree of freedom	Sum of square	Mean square	square root of (mean square between - mean square within)/2)	F value
Between sample	9	1755780	195087	233.2	2.26
Within sample	10	863500	86350		

The data collected were analyzed by using one-way ANOVA of single variance. Find the critical value from the F table for alpha=0.05. F 0.05(9,10)=3.02. The F value 2.26 is calculated based on ANOVA. This value is less than the critical value. The level of significant is 95%. Proficiency testing sample is uniform.

Table 2.3 Result of homogeneity test of reference material (DNOP)

DNOP			
Analysis method：CPSC-CH-C1001-09.2			
Result (%)			
Test number / Sample number	NO.1	NO.2	Average
1	1.000	1.090	1.0450
2	0.942	0.993	0.9675
3	1.030	1.020	1.0250
4	1.010	1.040	1.0250
5	1.030	1.020	1.0250
6	0.961	0.986	0.9735
7	0.991	0.955	0.9730
8	0.999	0.996	0.9975
9	0.985	0.946	0.9655
10	0.999	0.996	0.9975
Average			0.9995
Standard deviation, SD			0.0290
The general average	0.9995		
Note	1.Analysis instrument：GC-MS 2. Sample Forms：Granular material		

Table 2.4 Result of ANOVA of reference material (DNOP)

The sources of variation	Degree of freedom	Sum of square	Mean square	square root of(mean square between - mean square within)/2)	F value
Between sample	9	1517445	168605	214.8	2.21
Within sample	10	763050	76305		

The data collected were analyzed by using one-way ANOVA of single variance. Find the critical value from the F table for alpha=0.05. F 0.05(9,10)=3.02. The F value 2.21 is calculated based on ANOVA. This value is less than the critical value. The level of significant is 95%. Proficiency testing sample is uniform.

Appendix C

Instructions to Accreditation Body

Instructions to Participants

Results Sheet

APLAC T079 Phthalates in Plastic materials

Proficiency Testing Program

Instructions to Accreditation Bodies

1. ORGNAIZATION

The Proficiency Testing Program on "Determination of Phthalates in Plastic materials" is organized by Taiwan Accreditation Foundation (TAF) under the auspices of Asia Pacific Laboratory Accreditation Cooperation (APLAC). The main purpose of this program is to evaluate the competence of participating laboratories on the concerned tests.

SAMPLES

Accreditation bodies (AB) will receive PT samples and instruction information (hard copy and soft copy) from the organizers. Each laboratory is supplied with two amber bottles that each contains about 7 gram of PVC pellets. Upon receipt, AB should carefully check the samples (number of bottles, physical conditions, etc.) and promptly acknowledge the TAF coordinator by returning the Receipt Form (for Accreditation Bodies) through E-mail and facsimile to

Taiwan Accreditation Foundation (TAF)
Ms. Cherry Cheng
FAX: +886-3-5726308
E-mail:cherry@taftw.org.tw

2. SAMPLES DISTRIBUTION

AB is recommended to promptly distribute the samples, together with the Instructions to Participating Laboratories and Results Sheet (hard and soft copy) to their respective participating laboratories. The samples shall be stored at room temperature before dispatch.

AB shall remind their participating laboratories to carefully check the samples upon receipt and promptly acknowledge the TAF coordinator by returning the Receipt Form (for Participating Laboratories) through E-mail and facsimile before **September 3rd 2010**. New samples will be re-sent for any damaged samples.

3. TESTS TO BE CARRIED OUT

Please refer to the Instructions to Participating Laboratories enclosed, which outline the tests to be performed and the instructions for test method and reporting results.

The Instructions to Participating Laboratories documents are master copies. The AB should make copies for distribution to the participating laboratories. Instructions could be e-mailed on request. If necessary, these documents should be translated into the appropriate language by the accreditation body.

4. SAMPLE DISTRIBUTION

Each accreditation body shall be responsible for the samples dispatching to the respective participating laboratories within one week of receipt date and the delivery costs in their own economy.

5. DOCUMENTS TO BE SUBMITTED

Within one week of the completion of the testing, participating laboratories are required to both facsimile and e-mail to their accreditation body the Results Sheet requested in the Instructions to Participating Laboratories. No other documentation is required. The accreditation body should ensure the information has been supplied by the participating laboratories and is complete. No later than **September 17 2010**, AB is required to forward these complete results forms to the TAF coordinator cherry@taftw.org.tw by **E-mail**/Facsimile.

6. GENERAL INFORMATION

For general queries, please contact:

Taiwan Accreditation Foundation (TAF)
Ms. Cherry Cheng
FAX: +886-3-5726308
E-mail:cherry@taftw.org.tw

APLAC T079 Phthalates in Plastic materials

Proficiency Testing Program

Receipt Form (for Accreditation Body)

In order to monitor the progress of the Proficiency Testing Program, we kindly ask each accreditation body, on receipt of the samples, to fill in this **Receipt Form** and E-mail/facsimile it to:

Taiwan Accreditation Foundation (TAF)
Ms. Cherry Cheng
FAX: +886-3-5726308
E-mail:cherry@taftw.org.tw

Thank you in advance for your cooperation.

1. _____* APLAC T079samples were received on: _____ (date).
2. After inspection, are the contents of the package damaged:
   ☐ Yes ☐ No, if yes, is it still suitable for use? ☐ Yes ☐ No
3. Is there a Declaration to Customs Officials and Shipping Agents enclosed in the plastic envelope and attached to the outside of the case? ☐ Yes ☐ No

* number of samples

Remarks: _____

_____

_____

Name of Accreditation Body: _____

Contact Person: _____

Fax: _____

E-mail: _____

APLAC T079 Phthalates in Plastic materials

Proficiency Testing Program

Instructions to Participating Laboratories

1. ORGNAIZATION

The Proficiency Testing Program on "Determination of Phthalates in Plastic materials" is organized by Taiwan Accreditation Foundation (TAF) under the auspices of Asia Pacific Laboratory Accreditation Cooperation (APLAC). The main purpose of this proficiency testing program is to evaluate the competence of participating laboratories on the concerned tests.

2. SAMPLES

Participating laboratories will be given two amber sample bottles labeled APLAC A and APLAC B from their respective accreditation bodies. Each bottle contains about 7 gram of PVC pellet. Upon receipt of the samples, participating laboratories should carefully inspect the sample for any physical damages and defects. Participating laboratories shall promptly acknowledge receipt of the sample by returning the Receipt Form (for Participating Laboratories) through E-mail and facsimile to the coordinator of your AB. New samples will be replaced for any damaged claims. Intact sample is recommended to be stored in a secure environment at room temperature (around 25°C) before use.

3. TESTS TO BE PERFORMED

Participating laboratories are required to determine the concentration (in %) of Phthalates in the sample. Participating laboratories should use the test method CPSC-CH-C1001-09.3 (~~CPSC-CH-C1001-09.2~~), *Standard Operating Procedure for Determination of Phthalates*. Where possible, uncertainties should be calculated using the method in the ISO/IEC Guide 98-3:2008, *Uncertainty of measurement -- Part 3: Guide to the expression of uncertainty in measurement.*

4. RESULTS AND DOCUMENTS TO BE SUBMITTED

Report all the data in Results Sheet, where possible, estimate and report the expanded uncertainty for individual analyte. **Submit the results to your AB before September 15 2010**. Unless with special reasons, results submitted after the deadline might not be accepted by the organizers.

Laboratories should make a copy of all documents and worksheets for their own records and

keep these on file for an adequate period of time (at least until a final report has been issued).

5. CONFIDENTIALITY

Each laboratory will be assigned with a unique identification code. This unique code will be used throughout the programme. All information on the identities and results of participating laboratories will be kept confidential and not be disclosed.

6. GENERAL INFORMATION

For general queries, please contact with your accreditation body, additional information may be obtained from:

Taiwan Accreditation Foundation (TAF)

Ms. Cherry Cheng

TEL：+886-3-5714848 Ext 239

FAX：+886-3-5726308

E-mail:cherry@taftw.org.tw

APLAC T079 Phthalates in Plastic materials

Results Sheet

Lab Code:

Name of Accreditation Body: _____

Name of Participating Laboratory: _____

Contact person/E-mail: _____

Test Date: _____ (day/month/year)

Part I: Test Results

	Test	Result(%)	Measurement Uncertainty
Sample A	**DEHP**		
No.:	**DNOP**		
Sample B	**DEHP**		
No.:	**DNOP**		

Note: Report all tests to 3 significant figures or 3 decimal places, whichever is appropriate.

※ Recommend typing to fill in the form.

Part II: Method Information

1. Please click the method used

Sample Preparation	Extraction Method		Analysis method	
CPSC-CH-C1001-09.3 CPSC-CH-C1001-09.2		CPSC-CH-C1001-09.3 (CPSC-CH-C1001-09.2)		CPSC-CH-C1001-09.3 CPSC-CH-C1001-09.2
		Health Canada Method C-34		
		EN 14372:2004		Health Canada Method C-34
		EPA 3540C, Soxhlet Extraction		
		EPA 3541, Automated Soxhlet Extraction		EN 14372:2004
		EPA 3545A, Pressurized Fluid Extraction		
		EPA 3546, Microwave Extraction		EPA 8270D
		EPA 3550C, Ultrasonic Extraction		California Dept. of Toxic Substances control Method
		ASTM D2124-99(2004)		
		California Dept. of Toxic Substances control Method		

2. Method accreditation: ☐ YES; ☐ NO, _____

3. Other additional information:

_____

_____

_____

* Please circle as appropriate

- **<u>It is the responsibility of the participants to avoid collusion and falsification of result so as to ensure a reliable assessment to be made in this proficiency testing programme</u>**
- Please submit this "Results sheet" E-mail and facsimile to the coordinator of your AB.

Signature: _____ Date: _____

INSTRUCTIONS – MEASUREMENT UNCERTAINTY

Part (1) Background information & justification for this change

ISO/IEC 17025 requires that, except under specified conditions, the uncertainty of measurement associated with the results of tests and measurements must be estimated.

What is uncertainty of measurement?

Uncertainty of measurement is defined as a "parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand" (the measurand is the particular quantity subject to measurement).

The result of a test or measurement is our best estimate of the true value of the measurand. The result is imperfect. The true value of the measurand is contained within a range of values about the measurement result and the "uncertainty of measurement" is an estimate of the magnitude of that range expressed at a given level of confidence (confidence interval). Uncertainty of measurement is usually given as a 95% confidence interval and would normally be expressed in the appropriate SI units (i.e. mm, °C, g/l, MPa etc).

For example, the result of a measurement might be 5.1 mg/l with an uncertainty of ±0.2 mg/l at a 95% level of confidence. This means that there is an estimated 95% probability that the true value is in the range 4.9 mg/l to 5.3 mg/l. The 95% probability means that there is an estimated one in twenty chance that the true value is outside that range.

Uncertainty of measurement may also be expressed as a percentage where appropriate.

To assist laboratories to comply with the requirements of ISO/IEC 17025 for estimating uncertainty and to promote a uniform methodology in its estimation, information packages for APLAC PT testing program participants now include general guidance relating to estimating uncertainties for the specific testing involved, and final program reports will now include relevant worked examples. All program participants are required to report their estimates of uncertainty to their accreditation bodies along with their results unless the Technical Adviser to the program specifically waives any requirement to estimate uncertainties. The estimates of uncertainty provided by participants will be incorporated into the final program reports enabling direct comparison of uncertainty estimates across the program participants. The uncertainty estimates will not be used in the evaluation of the results on the primary samples.

How is uncertainty of measurement to be estimated?

APLAC expects that program participants' uncertainties of measurement would be estimated in accordance with the requirements of the respective member accreditation bodies. There are different approaches and methodologies available. Worked examples provided in APLAC PT program reports will generally be based on ISO GUM but will recognize other methodologies in accordance with 5.4.6.3 NOTE 3 in ISO 17025.

Estimates of uncertainty of measurement provided by program participants are required to be given at the 95% level of confidence.

ISO GUM methodology

An estimate of uncertainty of measurement would usually be based on the combination of a number of influencing parameters (components of uncertainty) such as errors in reference values, instrument errors, repeatability, thermal effects, weighing errors, inhomogeneity etc. ISO GUM methodology requires that the influence of each component of uncertainty on the measurement result be quantified and expressed numerically as a standard deviation. These values are then combined according to the rules of the propagation of uncertainty to produce a combined standard deviation (combined standard uncertainty) and the combined standard uncertainty is multiplied by a coverage factor to produce an expanded uncertainty at the required level of confidence. Detailed descriptions and information on the implementation of this methodology have been published by ISO2, UKAS3 and Eurachem/CITAC4 and made available over the internet[7].

Uncertainty of measurement is best estimated within the individual laboratory environment. All factors

which will have a significant influence on the test or measurement result must be included in the estimation process. There must be suitable programs utilizing reference standards, instruments and materials to ensure on-going and adequate quality control and repeatability and reproducibility of methods and equipment over time. In many instances, it will be possible to utilize quality control data in assessing uncertainty components such as precision. Where these data are not available, it may be necessary to carry out precision studies or to rely on published information about the method or portions of it until the laboratory can obtain its own estimates.

APLAC is aware of the general need for better estimates of uncertainty, and estimates that are obtained under similar conditions in all laboratories. PT programs are useful mechanisms for spreading awareness of uncertainty of measurement and the effects of different ways of estimating it. We anticipate that the information made available through PT programs will help focus discussions on uncertainty of measurement.

APLAC Technical Committees will interpret the information and report on current practices. They will also make recommendations for improving the collection of uncertainty data, the estimation of uncertainties and incorporating data and information on uncertainty of measurement into PT program reports. Therefore we anticipate an evolution in the mechanisms for collecting and reporting uncertainty data and associated information over the next few years.

Participation in APLAC PT programs should assist laboratories to develop appropriate estimates of uncertainty, help to guide accreditation bodies to adopting common and consistent approaches leading to enhanced understanding and international comparability of measurements among the member nations.

APLAC will consider the use of estimates of uncertainty of measurement in the evaluation of its PT testing program results after it is satisfied that participating laboratories are estimating uncertainties of measurement in an appropriate and consistent manner.

Here are a few important terms:

Standard uncertainty (*u(xi)*) is an input component of uncertainty *xi* expressed as a standard deviation. It should be expressed in the units of the influencing parameter, but may be expressed as a percentage where convenient.

Type A evaluation estimates of standard uncertainty are evaluated by applying statistical techniques to a series of repeatability or curve fitting data. For example, a standard uncertainty estimated from the repeatability of measurements on replicate samples is a Type A evaluation.

Type B evaluation estimates of standard uncertainty are based on assumed probability distributions, experience, laboratory records, or other information. For example, a standard uncertainty estimated using data provided on a calibration certificate is a Type B evaluation.

Sensitivity coefficient (*ci*) is the mathematical relationship between an influencing parameter and its effect on the result of a measurement. In many instances it is unity. That is, there is a one to one relationship between the value of the influence and its effect on the measurement result. For example, when weighing a sample of material, any uncertainty due to errors in the balance reading will have a one to one effect on the measurement result. On the other hand, if we are considering the influence of temperature on the length of a metal bar then the sensitivity coefficient is equal to the coefficient of linear thermal expansion for the metal bar multiplied by the length of the bar. It is important to note that a sensitivity coefficient has units. It is also important to note that the calculation methodology used by Eurachem/CITAC4 incorporates sensitivity coefficients in a manner that does not require their specific evaluation.

Combined standard uncertainty (*uc(y)*) is the final estimate of uncertainty for the test or measurement result *y* expressed as a standard deviation. It is calculated by multiplying the standard uncertainty *u(xi)* for each input component *(xi)* with its respective sensitivity coefficient *ci* to produce *ciu(xi)* and then combining those values by taking the square root of the sum of their squares. Note that the products *ciu(xi)* must each be expressed in the same units as those required for expressing the combined estimate *uc(y)*.

Expanded uncertainty (*U*) is the final result of our estimate of uncertainty expressed as a confidence interval. It is calculated by multiplying the combined standard uncertainty by a coverage factor to produce the desired level of confidence (usually 95%).

Coverage factor (*k*) is a multiplier used to expand the combined standard uncertainty *uc(y)* to an interval that is estimated to contain the true value of the measurand at a given level of confidence ($U = k.uc(y)$). The coverage factor then represents the number of standard deviations in the expanded uncertainty and is determined according to the Student-t distribution. A coverage factor of 2 is commonly used to approximate the expanded uncertainty to the 95% confidence level.

References

1. ISO/IEC 17025:2005, General requirements for the competence of testing and calibration laboratories.
2. ISO/IEC Guide 98-3:2008, Uncertainty of measurement -- Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)
3. UKAS LAB 12: The Expression of Uncertainty in Testing. UKAS, London (2000)
4. Eurachem / CITAC Guide QUAM: 2000.P1. Quantifying Uncertainty in Analytical Measurement, 2nd Edition (2000)
5. ISO/DTS 21748:2002, Guide for the use of repeatability, reproducibility, and trueness estimates in measurement uncertainty estimation.
6. APLAC TC 005, Interpretation and Guidance on the Estimation of Uncertainty of Measurement in Testing
7. www.A2LA.org / (for A2LA policies, links to guidance documents, including the UKAS Guide, and the Eurachem/CITAC Guide, at no cost), www.measurementuncertainty.org / (Eurachem/CITAC Guide), www.fasor.com/iso25 / (general information, links, and discussion of ISO-IEC 17025)

Appendix D

Statistical Procedures

Calculations and Formulae

STATISTICAL PROCEDURES

The reported data from participating laboratories use the statistical method in ISO 13528 annex C robust analysis: algorithm A. The data sorted into increasing order and calculated their average and stand deviation. The new values of robust average and standard deviation may be derived by an iterative calculation until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust range. This algorithm yields robust values of the average and standard deviation of the data to which it is applied.

The statistical procedures are based on robust statistics and use z-score to assess performance of participant. Robust statistics are statistics that are not highly influenced by the presence of extreme results. In a mathematical environment, robustness is the ability of a statistical method center of a database and the mean (average) is not a robust parameter. A robust alternative to the mean is the median (the middle value). The z-score are normalized value that gives a score to each result, relative to the other numbers in the group, so a z-score value close to zero means that result agrees well with those from the other laboratories.

The z-score is a statistical procedure to identify outlier results. The calculation of the z-score depends on the statistical design of the program. Related samples can be identical (uniform pair) or similar (split pair). Occasionally it may be the case that a program can only be designed to have a single result on a single sample. An outlier will be any result(s), which has an absolute z-score value greater than three. Outliers are identified in the table by a marker (※) next to the relevant z-scores.

A very high between-laboratories z-score indicates that one or both of a laboratory's results is significantly higher than the consensus value (median). Similarly, a very low (i.e. negative) between-laboratories z-score shows that a laboratory's result(s) is lower than expected.

Each participant in this program receives two test samples and the two related samples are uniform pair. The pairs of return data are calculated to evaluate the error source in participant's laboratory as the between-laboratories z-score for laboratory i (denoted ZBi) and the within-laboratories z-score (ZWi).

Statistical Calculations and Formulae

Robust analysis of test results:

The data are sorted into increasing order, and then calculate the initial average (x*) and standard deviation (s*) of these data as below.

(x*)=median of the data

(s*)=1.483 multiply with median of the difference between the data and median

The δ value is the standard deviation (s*) multiply with 1.5.

$$\delta = 1.5s^*$$

For each data, the new value is calculated as the following formula:

$$x_i^* = \begin{Bmatrix} x^* - \delta & if\ x_i < x^* - \delta \\ x^* + \delta & if\ x_i > x^* + \delta \\ x_i & otherwise \end{Bmatrix}$$

The new average (x*) and standard deviation (s*) are calculated as the following formula:

$$x^* = \frac{\sum x_i^*}{p}$$

$$s^* = 1.134\sqrt{\sum (x_i^* - x^*)^2 / (p-1)}$$

The new values of robust average and standard deviation may be derived by an iterative calculation until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust range. This algorithm yields robust values of the average and standard deviation of the data to which it is applied.

Robust z-scores are calculated by replacing the mean and standard deviation in the "classical" z-score by robust values of the average and standard deviation, respectively - i.e. the z-score for a result is the result minus the robust average (all) divided by the robust standard deviation.

Suppose we have pairs of results (for related samples) from a number of laboratories. Let Ai be the result for the first sample (A) from laboratory i. Similarly, Bi is the result from laboratory i for the second sample (B). The number of (pairs of) results is denoted by N.

As indicated above, the between-laboratories and within-laboratories z-scores are based on the (standardized) sum and difference of the pair of results.

Let Si be the standardized sum of laboratory's results. Also, Di is the standardized difference between the two results from laboratory i. Now

$$Si = \frac{(Ai + Bi)}{\sqrt{2}}$$

$$Di = \frac{(Ai - Bi)}{\sqrt{2}} \qquad \text{if} \quad (\text{Ai}) > (\text{Bi}),$$

or

$$Di = \frac{(Bi - Ai)}{\sqrt{2}} \qquad \text{if} \quad (\text{Ai}) < (\text{Bi}).$$

So, the between-laboratories z-score for laboratory i (denoted ZBi) is the robust z-score of its Si and their within-laboratories z-score (ZWi) is the robust z-score of its Di - i.e.

$$ZBi = \frac{Si - median(Si)}{IQR(Si) \times 0.7413}$$

and

$$ZWi = \frac{Di - median(Di)}{IQR(Di) \times 0.7413}$$

The median is the middle value of the between-laboratories and within-laboratories groups, i.e. half of the results are higher than it and half are lower. It is calculated from the sorted values (from lowest to highest). If N is an odd number, the median is the singular central value. If N is even, it is the average of the two central values.

The interquartile range (IQR) is the difference between the lower and upper quartiles. The lower quartile (Q1) is the value below which a quarter of the results lie. Similarly, the upper quartile (Q3) is the value above which a quarter of the results lie. The quartiles are calculated analogously to the median and IQR=Q3-Q1. The "Normalized IQR" equals IQR×0.7413. The factor 0.7413 comes from the standard normal distribution, which has a mean of zero and a standard deviation equal to one. The width of the interquartile range of such a distribution is 1.34898 and 1/1.34898 = 0.7413.

Multiplying the IQR by this factor makes it comparable to a standard deviation. The "Robust CV" is a coefficient of variation and is equal to the normalized IQR divided by the median, expressed as a percentage (i.e. multiplied by 100). The minimum, maximum and range are the lowest value, the highest value and the

difference between them (respectively).

The list of summary statistics in the between-laboratories and within-laboratories group appears at the neighbor table of the Z-score of results and consists of:

The number of results for that test/sample (No. of Results);

The median of laboratory's results - i.e. the middle value (Median);

The normalized interquartile range of the results (Normalized IQR) – the interquartile range times 0.7413;

The robust coefficient of variation, expressed as a percentage (Robust CV) - i.e.

100× Normalized IQR ÷ Median;

The minimum and maximum laboratory results; and

The range is calculated from (Maximum - Minimum).

The Histograms and Bar Charts of Z-Score

The histograms of z-score for two kinds of phthalates (DEHP and DNOP) with two test samples were shown in appendix A. The Bar charts of z-score were ordered by the laboratories z-score value. Ordered between laboratories z-score bar charts and ordered within laboratory z-score bar charts were also used to illustrate the laboratories data. Each laboratory z-score was drawn on a chart and identified by their code number. The z-score limits of ± 3 were also drawn which makes it easy to identify any bars extending past this point which would be outliers.

These charts contained dashed lines at +3 and –3, so the test outliers were clearly identifiable as the laboratories whose "bar" extends beyond these cut-off lines. The y-axis was also limited to range from –6 to +6, so in some cases very large (positive) or small (negative) Z-scores appeared as extending beyond the limit of the chart.

These between-laboratories z-score (ZB) and the within-laboratories z-score (ZW) were indicated to the error source in laboratory. When the absolute ZB score was greater than 3, it means there was a serious systematic error in this laboratory. When the absolute ZW score was greater than 3, it means there was a serious random error in this laboratory.