



Institute for
Interlaboratory Studies

**Results of Proficiency Test
Preservatives in Skin Care products
(CMIT/MIT &
Parabens/Phenoxyethanol &
Benzoic Acid/Formaldehyde)
December 2024**

Organized by: Institute for Interlaboratory Studies
Spijkenisse, the Netherlands

Author: Mr. R.J. Starink, BSc
Correctors: Mrs. A. Ouwerkerk, BSc
Approved by: Mr. R.J. Starink, BSc

Report: iis24H72

February 2025

CONTENTS

1	INTRODUCTION	3
2	SET UP	3
2.1	QUALITY SYSTEM	4
2.2	PROTOCOL.....	4
2.3	CONFIDENTIALITY STATEMENT	4
2.4	SAMPLES	4
2.5	ANALYZES	6
3	RESULTS	7
3.1	STATISTICS	7
3.2	GRAPHICS	8
3.3	Z-SCORES	8
4	EVALUATION	9
4.1	EVALUATION PER SAMPLE AND PER COMPONENT	9
4.2	PERFORMANCE EVALUATION FOR THE GROUP OF LABORATORIES.....	11
4.3	COMPARISON OF THE PROFICIENCY TEST OF DECEMBER 2024 WITH PREVIOUS PTS.....	12
4.4	EVALUATION OF THE ANALYTICAL DETAILS	12
5	DISCUSSION	13
6	CONCLUSION.....	14

Appendices:

1.	Data, statistical and graphic results	15
2.	Analytical details	26
3.	Number of participants per country	27
4.	Abbreviations and literature.....	28

1 INTRODUCTION

Preservatives may be used in cosmetics to prevent the growth of harmful bacteria and mold. Chloromethylisothiazolinone (CMIT) and Methylisothiazolinone (MIT), Parabens, Phenoxyethanol, Formaldehyde and Benzoic Acid are widely used as preservatives in liquid cosmetic and personal care products. CMIT and MIT could be allergenic and cytotoxic, while Parabens and Phenoxyethanol are linked to hormonal disruption. These preservatives in skin care products are regulated through Annex V of Regulation (EC) No 1223/2009 ("Cosmetics Regulation").

Since 2018 the Institute for Interlaboratory Studies (iis) organizes a proficiency scheme for the determination of CMIT (5-Chloro-2-Methyl-4-Isouthiazolin-3-one) and MIT (2-Methyl-4-Isouthiazolin-3-one) in Skin Care Products and for Parabens and other preservatives since 2019. During the annual testing program of 2024, it was decided to continue the proficiency test for the determination of Preservatives in Skin Care Products. It was decided to conduct three different proficiency tests (PTs) of Preservatives in Skin Care Products: the PT CMIT/MIT, the PT Parabens/Phenoxyethanol and the PT Benzoic Acid/Formaldehyde.

In this interlaboratory study registered for participation:

- 9 laboratories in 5 countries for CMIT/MIT in Skin Care products iis24H72A
- 12 laboratories in 8 countries for Parabens/Phenoxyethanol in Skin Care products iis24H72B
- 12 laboratories in 8 countries for Benzoic Acid/Formaldehyde in Skin Care products iis24H72C

In total 21 laboratories in 12 countries registered for participation, see appendix 3 for the number of participants per country. In this report the results of the Preservatives in Skin Care products proficiency test are presented and discussed.

2 SET UP

The Institute for Interlaboratory Studies (iis) in Spijkenisse, the Netherlands, was the organizer of this proficiency test (PT). Sample analyzes for fit-for-use and homogeneity testing were subcontracted to a laboratory that has performed the tests in accordance with for ISO/IEC17043 relevant requirements of ISO/IEC17025.

In this proficiency test the participants received, depending on the registration, from one up to three different Skin Care products, see table below.

Sample ID	PT and product type	Quantity
#24775	CMIT/MIT in Body Cream iis24H72A	1x 10 mL
#24780	Parabens/Phenoxyethanol in Body Butter iis24H72B	1x 10 mL
#24785	Benzoic Acid/Formaldehyde in Uddercream iis24H72C	1x 10 mL

Table 1: samples used in PT iis24H72

The participants were requested to report rounded and unrounded test results. The unrounded test results were preferably used for statistical evaluation.

2.1 QUALITY SYSTEM

The Institute for Interlaboratory Studies in Spijkenisse, the Netherlands, has implemented a quality system based on ISO/IEC17043:2010. This ensures strict adherence to protocols for sample preparation and statistical evaluation and 100% confidentiality of participant's data. Feedback from the participants on the reported data is encouraged and customer's satisfaction is measured on regular basis by sending out questionnaires.

2.2 PROTOCOL

The protocol followed in the organization of this proficiency test was the one as described for proficiency testing in the report 'iis Interlaboratory Studies: Protocol for the Organization, Statistics and Evaluation' of October 2024 (iis-protocol, version 4.0). This protocol is electronically available through the iis website www.iisnl.com, from the FAQ page.

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

For the PT CMIT/MIT in Skin care products a batch of body cream was purchased from a local supermarket and was artificially fortified with CMIT/MIT. After homogenization 30 PE bottles of 10 mL were filled and labelled #24775.

The homogeneity of the subsamples was checked by determination of CMIT and MIT using an in-house test method on 4 stratified randomly selected subsamples.

	CMIT in mg/kg	MIT in mg/kg
sample #24775-1	41.0	18.5
sample #24775-2	40.9	18.4
sample #24775-3	41.5	19.0
sample #24775-4	41.4	18.5

Table 2: homogeneity test results of subsamples #24775

From the above test results the repeatabilities were calculated and compared with 0.3 times the estimated reproducibility calculated with the Horwitz equation in agreement with the procedure of ISO13528, Annex B2 in the next table.

	CMIT in mg/kg	MIT in mg/kg
r (observed)	0.8	0.7
reference method	Horwitz	Horwitz
0.3 x R (reference method)	3.2	1.6

Table 3: evaluation of the repeatabilities of subsamples #24775

The calculated repeatabilities are in agreement with 0.3 times the estimated reproducibility calculated with the Horwitz equation. Therefore, homogeneity of the subsamples was assumed.

For the PT Parabens/Phenoxyethanol in Skin Care products a batch of Body Butter was purchased from a local supermarket and was artificially fortified with the preservatives Methylparaben, Ethylparaben, Propylparaben, Isobutylparaben, Butylparaben and Phenoxyethanol. After homogenization 40 PE bottles of 10 mL were filled and labelled #24780.

The homogeneity of the subsamples was checked by determination of Methylparaben, Ethylparaben and Phenoxyethanol by using an in-house test method on 5 stratified randomly selected subsamples.

	Methylparaben in mg/kg	Ethylparaben in mg/kg	Phenoxyethanol in mg/kg
sample #24780-1	385.1	218.0	10545
sample #24780-2	385.0	218.5	10439
sample #24780-3	385.1	218.7	10526
sample #24780-4	384.0	218.9	10447
sample #24780-5	385.3	218.0	10514

Table 4: homogeneity test results of subsamples #24780

From the above test results the repeatabilities were calculated and compared with 0.3 times the estimated reproducibility calculated with the Horwitz equation in agreement with the procedure of ISO13528, Annex B2 in the next table.

	Methylparaben in mg/kg	Ethylparaben in mg/kg	Phenoxyethanol in mg/kg
r (observed)	1.5	1.2	134
reference method	Horwitz	Horwitz	Horwitz
0.3 x R (reference method)	21.1	13.0	350

Table 5: evaluation of the repeatabilities of subsamples #24780

The calculated repeatabilities are in agreement with 0.3 times the estimated reproducibility calculated with the Horwitz equation. Therefore, homogeneity of the subsamples was assumed.

For the PT Benzoic Acid/Formaldehyde in Skin Care products a batch of Uddercream was purchased from a local supermarket and was artificially fortified with Benzoic Acid and 4-Hydroxybenzoic. After homogenization 40 PE bottles of 10 mL were filled and labelled #24785.

The homogeneity of the subsamples was checked by determination of Benzoic Acid and 4-Hydroxybenzoic acid by using an in-house test method on 4 stratified randomly selected subsamples.

	Benzoic Acid in mg/kg	4-Hydroxybenzoic acid in mg/kg
sample #24785-1	2337	3243
sample #24785-2	2296	3277
sample #24785-3	2324	3238
sample #24785-4	2347	3220

Table 6: homogeneity test results of subsamples #24785

From the above test results the repeatabilities were calculated and compared with 0.3 times the estimated reproducibility calculated with the Horwitz equation in agreement with the procedure of ISO13528, Annex B2 in the next table.

	Benzoic Acid in mg/kg	4-Hydroxybenzoic acid in mg/kg
r (observed)	61	67
reference method	Horwitz	Horwitz
0.3 x R (reference method)	97	129

Table 7: evaluation of the repeatabilities of subsamples #24785

The calculated repeatabilities are in agreement with 0.3 times the estimated reproducibility calculated with the Horwitz equation. Therefore, homogeneity of the subsamples was assumed.

Depending on the registration of the participant the appropriate set of PT samples was sent on November 6, 2024.

2.5 ANALYZES

The participants were requested to determine on sample #24775 the concentrations of CMIT (5-Chloro-2-Methyl-4-Isothiazolin-3-one) and MIT (2-Methyl-4-Isothiazolin-3-one).

On sample #24780 it was requested to determine the concentrations of Methylparaben as ester, Ethylparaben as ester, Propylparaben as ester, Isobutylparaben as ester, Butylparaben as ester and Phenoxyethanol.

On sample #24785 it was requested to determine the concentrations of 4-Hydroxybenzoic acid, Benzoic acid and Formaldehyde.

It was also requested to report if the laboratory was accredited for the determined components and to report the amount of sample intake.

It was explicitly requested to treat the samples as if they were routine samples and to report the test results using the indicated units on the report form and not to round the test results but report as much significant figures as possible. It was also requested not to report 'less than' test results, which are above the detection limit, because such test results cannot be used for meaningful statistical evaluations.

To get comparable test results a detailed report form and a letter of instructions are prepared. On the report form the reporting units are given as well as the reference test methods (when applicable) that will be used during the evaluation. The detailed report form and the letter of instructions are both made available on the data entry portal www.kpmd.co.uk/sgs-iis-cts/. The participating laboratories are also requested to confirm the sample receipt on this data entry portal. The letter of instructions can also be downloaded from the iis website www.iisnl.com.

3 RESULTS

During five weeks after sample dispatch, the test results of the individual laboratories were gathered via the data entry portal www.kpmd.co.uk/sgs-iis-cts/. The reported test results are tabulated per determination in appendix 1 of this report. The laboratories are presented by their code numbers.

Directly after the deadline, a reminder was sent to those laboratories that had not reported test results at that moment. Shortly after the deadline, the available test results were screened for suspect data. A test result was called suspect in case the Huber Elimination Rule (a robust outlier test) found it to be an outlier. The laboratories that produced these suspect data were asked to check the reported test results (no reanalyzes). Additional or corrected test results are used for data analysis and the original test results are placed under 'Remarks' in the result tables in appendix 1. Test results that came in after the deadline were not taken into account in this screening for suspect data and thus these participants were not requested for checks.

3.1 STATISTICS

The protocol followed in the organization of this proficiency test was the one as described for proficiency testing in the report 'iis Interlaboratory Studies: Protocol for the Organization, Statistics and Evaluation' of October 2024 (iis-protocol, version 4.0).

For the statistical evaluation the *unrounded* (when available) figures were used instead of the rounded test results. Test 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. If a data set does not have a normal distribution, the (results of the) statistical evaluation should be used with due care.

The assigned value is determined by consensus based on the test results of the group of participants after rejection of the statistical outliers and/or suspect data.

According to ISO13528 all (original received or corrected) results per determination were submitted to outlier tests. In the iis procedure for proficiency tests, outliers are detected prior to calculation of the mean, standard deviation and reproducibility. For small data sets, Dixon (up to 20 test results) or Grubbs (up to 40 test results) outlier tests can be used. For larger data sets (above 20 test results) Rosner's outlier test can be used. 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. 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. In this PT the criterion of ISO13528, paragraph 9.2.1, was met for all evaluated tests. Therefore, the uncertainty of all assigned values may be negligible and need not be included in the PT report.

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

3.2 GRAPHICS

In order to visualize 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 test results are plotted. The corresponding laboratory numbers are on 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 reference test method. 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. This 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 (dotted line) was projected over the Kernel Density Graph (smooth line) for reference. The Gauss curve is calculated from the consensus value and the corresponding standard deviation.

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 (derived from e.g. ISO or ASTM test methods), the z-scores were calculated using a target standard deviation. This results in an evaluation independent of the variation in this interlaboratory study.

The target standard deviation was calculated from the literature reproducibility by division with 2.8. In case no literature reproducibility was available, other target values were used, like Horwitz or an estimated reproducibility based on former iis proficiency tests.

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:

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

The $z_{(\text{target})}$ scores are listed in the test result tables in appendix 1.

Absolute values for $z < 2$ are very common and absolute values for $z > 3$ are very rare. Therefore, the usual interpretation of z-scores is as follows:

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

End of preview of this report, requests to obtain the full report can be sent to nl.iis@sgs.com

Preview

Address: Malledijk 18, P.O. Box 200, 3200 AE Spijkenisse, The Netherlands
Telephone number: +31 (0)88 214 45 41
Email address: nl.iis@sgs.com
Website: www.iisnl.com

Institute for Interlaboratory Studies is a full member of SGS Nederland B.V. and registered at the Chamber of Commerce under number: 24226722. Unless otherwise agreed, all orders are executed in accordance with the SGS general conditions.