Laboratory PT #1 September 2022

1st Round of Mercury Laboratory Proficiency Testing

– Asia and the Pacific –

Final Report

1. Objective

This proficiency testing (PT) programme was to evaluate the performance of mercury analyses conducted by the laboratories. It was expected to provide the individual proficiency levels of participating laboratories and collective mercury monitoring capacity in the region.

2. Proficiency testing provider

This PT was organized by United Nations Environmental Programme, Regional Office for Asia-and the Pacific (UNEP-ROAP) with the overall PT design by National Institute for Minamata Disease (NIMD). Asian Institute of Technology Regional Resource Centre for Asia and the Pacific (AIT RRC.AP) distributed the test item and collected the analytical results. IDEA Consultants, Inc. prepared the test item for this PT.

3. Implementation period

Call for participation:	August – December 2021
Test item distribution:	February – March 2022
Duration of test (analysis):	Until 8 April 2022

4. Participation fee

Free.

5. Test item (sample) and parameter

One (1) human hair sample was used for analysing total mercury concentration.

5.1. Test item preparation

The hair was collected from multiple people who was not occupationally exposed of mercury. Collected hair was cut (<1 mm) without crushing and mixed to ensure sufficient homogeneity. The test item was also sterilized by gamma-ray.

Approximately 3 g each of the prepared hair was packed in brown glass bottle. The bottles of test items were sealed in aluminium-lined laminate packs for distribution to the participants.

5.2. Homogeneity testing

To ensure no significant differences of the mercury concentrations in the test items between the bottles, which might affect the result of the PT, following homogeneity testing of the test item was conducted.

After the preparation of test item (packed in bottles), 10 bottles were selected and the total mercury analysis (acid digestion – aeration CVAAS measurement) was performed twice per each test item in a bottle.

Homogeneity of the test item was then analysed from the results of the total mercury concentrations. Because the analysis results contain the uncertainty by the (chemical) analysis procedure, homogeneity was judged by following criterion:

Criterion: $S_s \leqq \sqrt{F_1 \times (0.3 \times \sigma_{ep})^2 + F_2 \times {S_w}^2}$

S_s: relative standard deviation of homogeneity testing

σ_{ep}: (expected) relative standard deviation of the reported results from participants

$w_i^2 = \Sigma (x_{gm}^2 - \bar{x}_g^2) / (m-1)$ x_{gm}: result of m times analysis of the bottle

$S_w^2 = \Sigma w_i^2 / g$ $\bar{x}_g$: average of the result of each bottle

F_1 and F_2 are values which are calculated from the probability distribution. In this homogeneity testing (10 bottles testing), F_1 and F_2 were applied following numbers:

$F_1 = 1.88$ $F_2 = 1.01$

(Even though these values are referred from the Annex B of ISO13528:2015, they are introduced from the random variables of χ^2 distribution and F distribution.)

Analysis results of this homogeneity testing are as follows:

$S_s = 0.0256$

$S_w^2 = 0.000897$

Also, relative standard deviation of the results from the participants was as follows:

$\sigma_{ep} = 0.0837$

This standard deviation should be used the value which was used for the performance evaluation for participants. As described in 9.2, the performance of participants was evaluated from the median and normalized interquartile range (NIQR) of the results, thus relative NIQR was used for the confirmation of the criterion.

Therefore, above criterion was judged as follows:

$$\sqrt{F_1 \times (0.3 \times \sigma_{ep})^2 + F_2 \times {S_w}^2}$$

$$= \sqrt{1.88 \times (0.3 \times 0.0837)^2 + 1.01 \times 0.000897^2}$$

$= 0.0344 > 0.0256\ (S_s)$

It was confirmed that the test item was sufficiently homogeneous to evaluate the performance of participants' results.

5.3. Stability testing

To ensure the concentration of the target parameter (total mercury) being maintained without significant changes during the PT, following stability testing was conducted after the duration of the analysis.

Stored (not distributed to participants) 10 bottles of test items were selected and the total mercury analysis (acid digestion – aeration CVAAS measurement) was performed twice per each test items in a bottle.

Stability of the test item was then analysed by comparing the results before and after the distribution of the test item. The stability of test item was judged by following criterion:

Criterion: $| \bar{x} - \bar{y} | \leqq 0.3 \times \sigma_{pt} + 2 \times \sqrt{u_{(x)}{}^2 + u_{(y)}{}^2}$

$\bar{x}$: average of the item before distribution

$\bar{y}$: average of the item after proficiency testing

$u_{(x)}$: uncertainty of the item before distribution

$u_{(y)}$: uncertainty of the item after proficiency testing

σ_{pt}: standard deviation for the proficiency evaluation. In this program, NIQR was applied to evaluation of performance of the participant.

Analysis results of test items before and after the PT are as follows:

$\bar{x} = 1.289$ $\quad u_{(x)} = 0.0104$

$\bar{y} = 1.294$ $\quad u_{(y)} = 0.0157$

Standard deviation of the result of all participants was as follows:

$\sigma_{pt} = 0.122$

This standard deviation should be used the value which was used for the performance evaluation for participants. As described in 9.2, the performance of participants was evaluated from the median and normalized interquartile range (NIQR) of the results, thus NIQR was used for the confirmation of the criterion.

Therefore, above criterion was judged as follows:

$0.3 \times \sigma_{pt} + 2 \times \sqrt{(u_{(x)}{}^2 + u_{(y)}{}^2}$

$$= 0.3 \times 0.122 + 2 \times \sqrt{0.0104^2 + 0.0157^2}$$
$$= 0.0743 > 0.005 \ (|\bar{x} - \bar{y}|)$$

It was confirmed that the concentration of total mercury in test item was not changed during the PT.

6. Target parameter

Target parameter of the PT was total mercury. Participants conducted 3 times of total mercury analysis and reported all results of them. Participants also conducted analysis of the moisture in the test item. The result of moisture was used for the analysis of reported data; however, it was not the target of the PT, and the result of total mercury was not calculated by the moisture.

The analysis procedure of moisture was instructed to participants as follows:

1. Take a hair sample of 100 mg or more and weigh it precisely.
2. Dry the taken hair sample (85 °C, 4 hours)
3. Weigh the dried hair sample again and calculate the moisture of sample from the reduced mass.

Also, it was instructed that the sample used for moisture analysis should not be used for total mercury analysis.

7. Participating institutions

The public laboratories or laboratories in universities that undertakes mercury analysis were intended for this PT. Performance of the analysis that lower of detection limit than 0.1 mg/kg on 10 mg test item was requested, however there was an institution who was unable to perform that analysis and report the result.

34 institutions registered in the PT and 26 institutions reported analysis results.

Table 1 Changes in number of laboratories per step

Category	Number of laboratories
Registered	34
Sample received	31
Result delivered	26

Most participating laboratories were from Asia and the Pacific region but a few laboratories from other regions were also participated in the PT. Number of laboratories participated from each region is shown in Table 2.

Table 2 Number of participating laboratories per region

UN Region	Number of laboratories
Africa	2
Asia and the Pacific	21
Central and Eastern Europe	0
Latin America and the Caribbean	2
Western Europe and Other Group	1

8. Methods and procedures

The method of analysis was not specified for the PT and the participants performed analysis by the method that they usually used in their routine analysis, or they were planning to use in the future.

The participants performed analysis of total mercury by the following methods:

- Thermal decomposition Cold Vapour Atomic Absorption Spectrometry (TDAAS)
- Acid digestion, reduction or hydride generation aeration Cold Vapour Atomic Absorption Spectrometry (CVAAS)
- Acid digestion, Cold Vapour Atomic Fluorescence Spectrometry (CVAFS)
- Acid digestion, Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

In CVAAS method, there were participants who used chemical reduction (using by such as stannous chloride) and who used hydride generation (using by such as borohydride).

9. Result

9.1. Basic statistic data of the PT result (total mercury)

The basic statistics of the result of PT are shown in Table 3.

Table 3 Summary of the results of the PT

Statistic data of the results (unit: mg/kg)	
Average:	1.405
Median:	1.458
Standard deviation:	0.238
Minimum	0.720
Maximum	1.824
25 percentiles	1.339
75 percentiles	1.504
Interquartile range (IQR)	0.165
Normalized IQR (NIQR)	0.122
Parameter related to distribution	
Skewness of distribution	-0.865
Kurtosis of distribution	1.880

The distribution of the results from the participants is shown in Fig. 1.

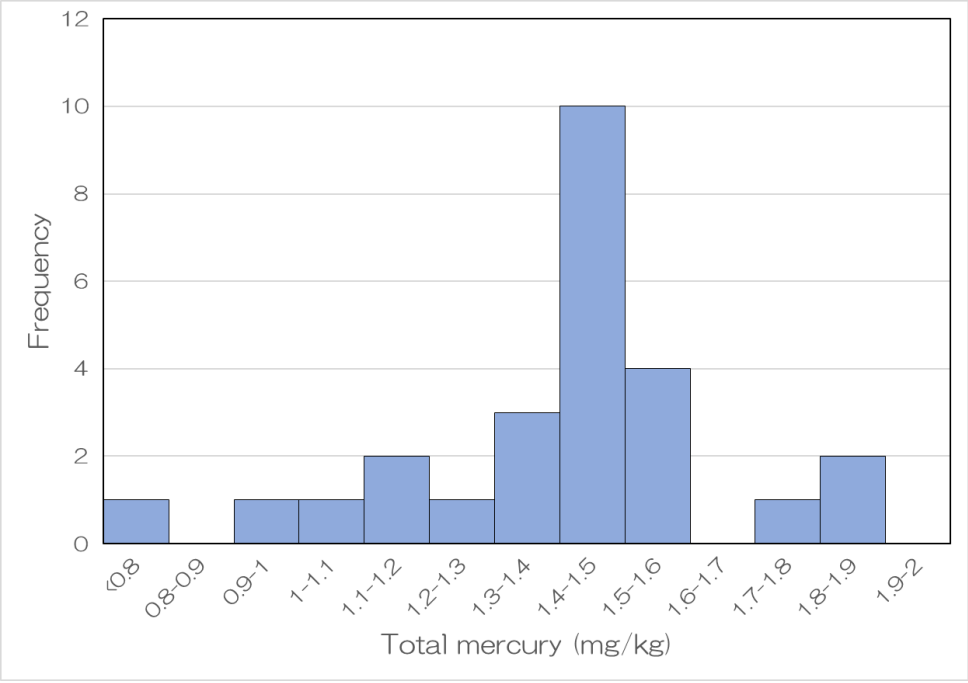

Fig. 1 Histogram of the report data (total mercury)

Some registered participants were unable to receive the test item because of several issues such as custom clearance, etc. Also, there were participants who were not able to report the analysis results because of trouble or insufficient performance of measuring instrument.

These statistical data were calculated from the average of each participant. There were participants who did not report 3 results (1, 2 or 4 results were reported), however all data were used for the statistical analysis.

It was also shown in large kurtosis of distribution, a relatively many reported values were concentrated in a narrow range around the median, while some reported values deviate from the median. Thus, the results in a relatively small interquartile range than the standard deviation.

There were more reported values lower than the median than larger, as evidenced by negative skewness values.

There were few reported results that were far from the median, and it was considered that the reported data did not have a singular distribution. Since no results were reported which were extremely different from other reported values and may affect the evaluation, the performance was evaluated based on the data obtained from all reported values without processing outliers.

9.2. Performance evaluation for participants

Median data of all laboratories was applied as agreement value. Performance of the results was evaluated by the robust z score which was calculated from the median and normalized interquartile range (NIQR).

z score of each participant was calculated from the following equation.

$$z = [(\text{average of reported result}) - (\text{median of all participants})] \,/\, \text{NIQR}$$

Performance of the result is classified by z score as follows:

$|z| \leqq 2$: Performance is satisfactory (satisfactory)

$2 < |z| < 3$: Performance is questionable (caution)

$|z| \geqq 3$: Performance is unsatisfactory (action)

The results and performances of laboratories are shown in Fig. 2.

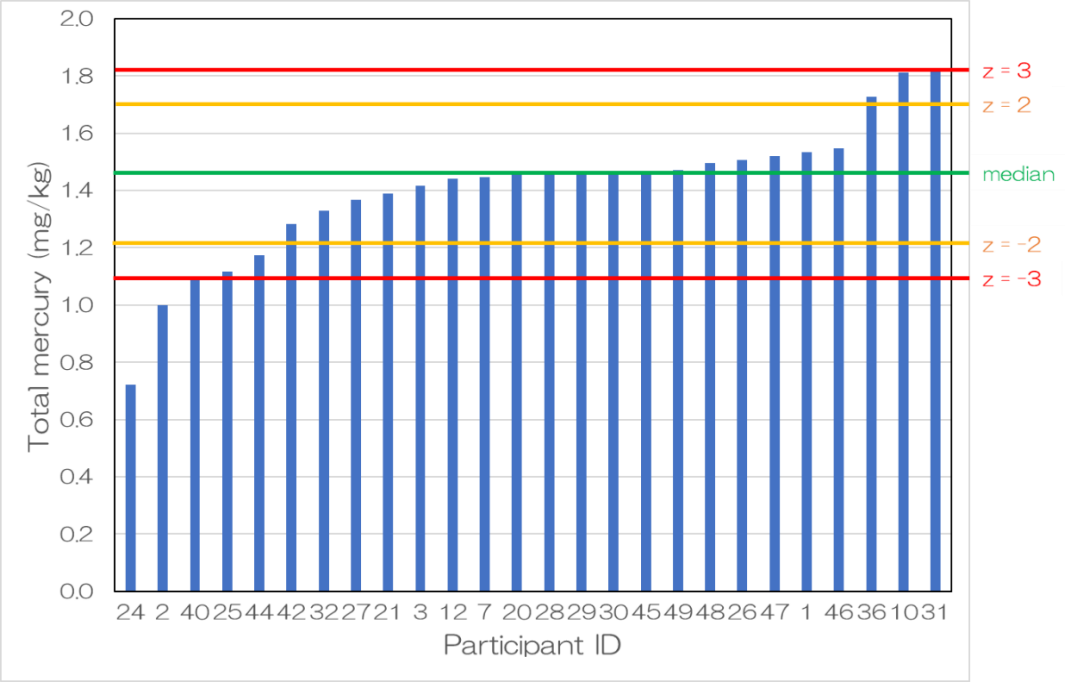

Fig. 2 Result of each participant

The numbers of laboratories disaggregated for each z score range are shown in Table 4.

Table 4 Number of the laboratories in the range of z score

z score	z ≦ -3	-3 <z< -2	-2 ≦z≦ 2	2 < z < 3	z ≧3
n	2	3	18	3	0

As described in 9.1, IQR of reported results were relatively small. Therefore, satisfactory range of the result (mercury concentration) was relatively close and results around 25 % difference from the median was the classified range of unsatisfactory of the result.

9.3. Regions of participating laboratories and the performances

Number of laboratories per region where the laboratories are located and the performance of the PT are shown in Fig. 3.

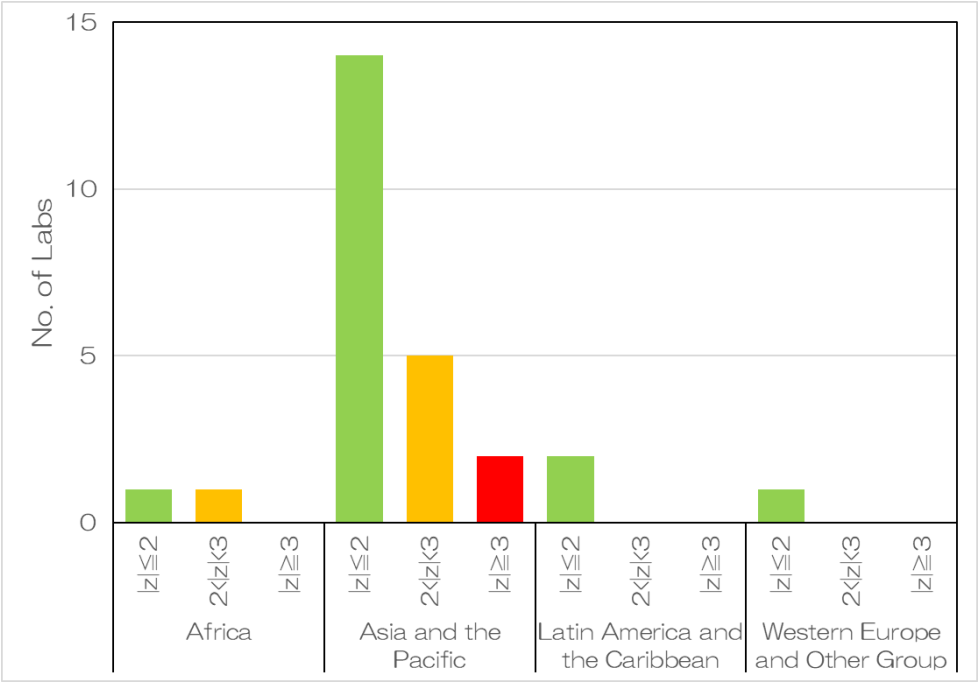

Fig. 3 Number of laboratories per regions and performances

Regional difference was not examined due to the limited numbers of laboratories participating from regions other than Asia and the Pacific being insufficient for comparison.

9.4. Types of laboratories and the performances

Number of laboratories per type (academic, government, or non-government) and the performance of the PT are shown in Fig. 4.

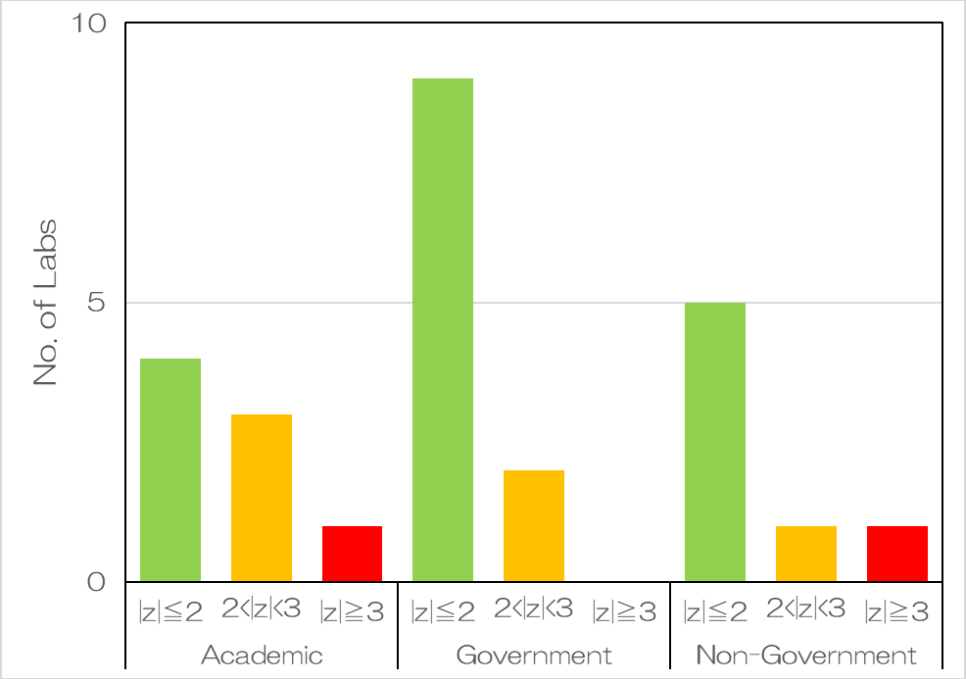

Fig. 4 Number of laboratories per types and performance

Statistical differences among the types of laboratories were not detected from the reported results (Kruskal-Wallis one-way ANOVA on ranks, p=0.44).

9.5. Analysis methods and results

The distribution of the results from participants by analysis method is shown in Fig. 5.

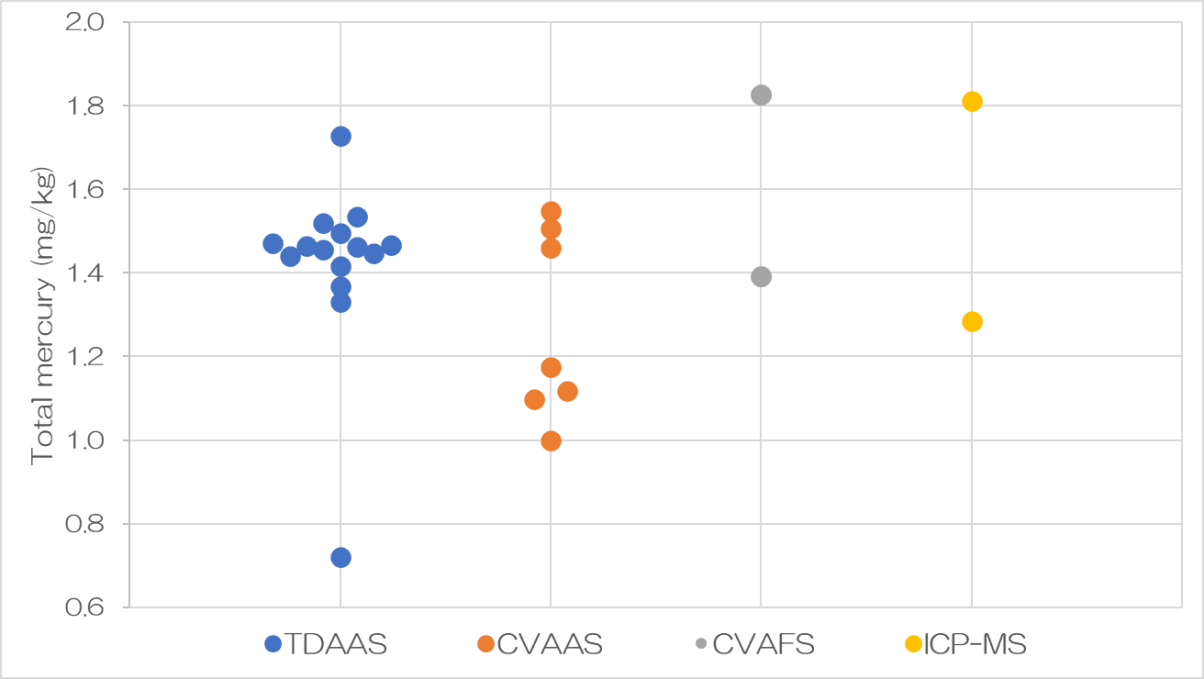

Fig. 5 Distribution of the report data by analysis method

Most TDAAS results agreed with median, but noticeable differences were not detected among the analysis methods employed (Kruskal-Wallis one-way ANOVA on ranks, p=0.12).

9.6. Moisture

The basic statistics of the reported moisture of test item is shown in Table 5.

Table 5 Summary of the result of moisture (Unit: %)

Average	11.6
Median	9.75
Standard deviation*1	16.3 (3.09)
Minimum	0.51
Maximum*2	88.4 (13.5)

*1: Value in parentheses is standard deviations excluding the maximum result.

*2: Value in parentheses is second largest result.

The distribution of the results of moisture is shown in Fig. 6.

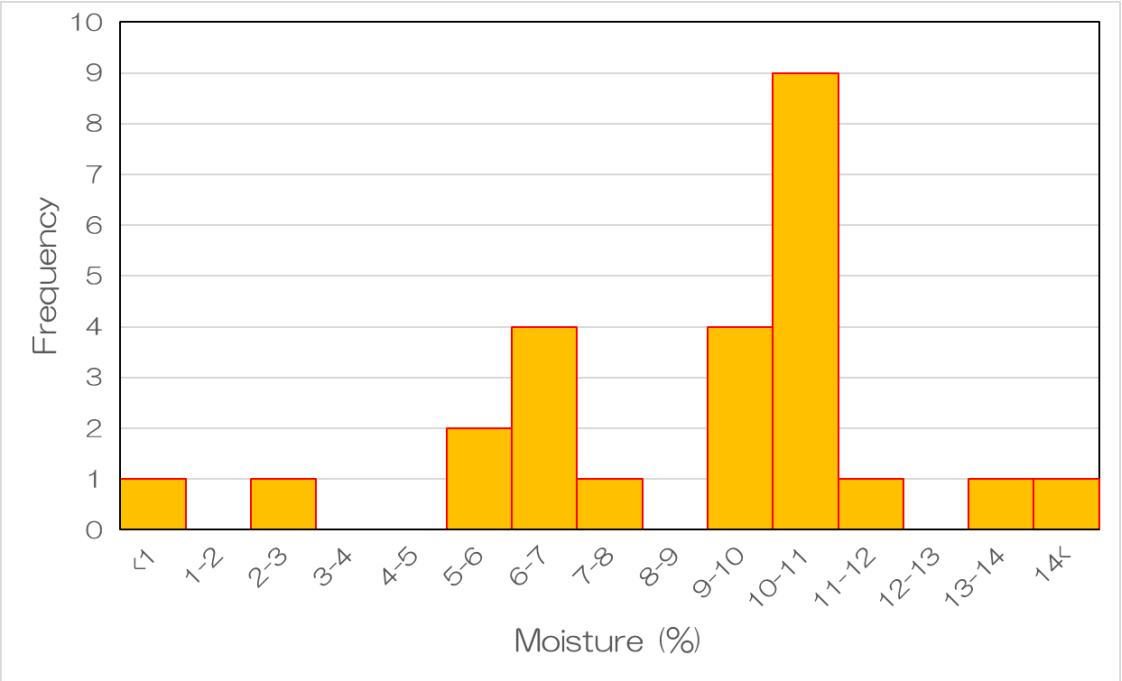

Fig. 6 Histogram of the result of moisture

The plots of the relations of moisture and mercury concentration are shown in Fig. 7. (The largest value of moisture (88.1 %) is not shown in this figure.)

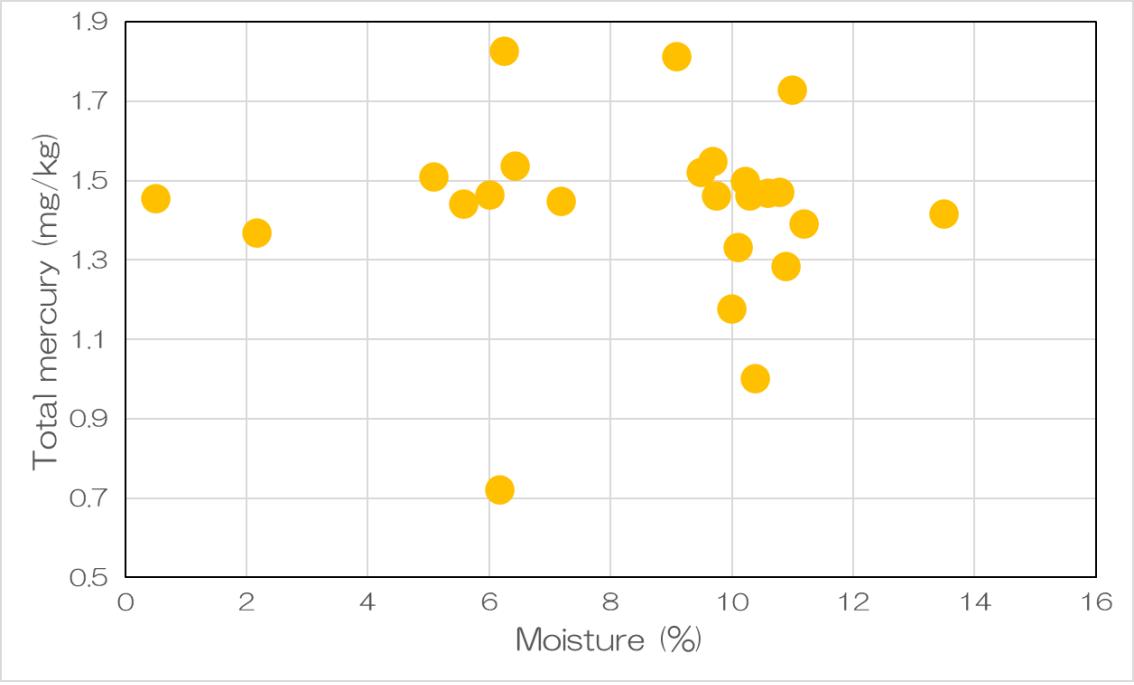

Fig. 7 Total mercury and moisture

The deviation of the moisture was larger than the mercury concentration, but noticeable relation between moisture and mercury concentration was not detected. It was considered that the uncertainty of the analysis of moisture was larger than the change of the moisture during the PT analysis duration.

10. Conclusion

Even the participants of this time PT were not so many, many of the reported results from the participants were concentrated around the median value and IQR of the results was relatively close. Thus, it was considered that the difference among with many mercury analysis laboratories were relatively small.

The range of moisture reported from the participants were larger than the deviation of the mercury results even it was not the target of the PT. Also, the correlation between the mercury and moisture was not detected. This result suggests that precise measurement of moisture for the small amount of sample such as the hair is not easy and there may be a challenge to correct the result of such sample by the moisture.