Journal of the Korean Society of
Tobacco Science Vol. 20, No. 1, 108~114 (1998)
Printed in Republic of Korea.

Statistical Evaluation of Smoke Analysis Technique through Asia Collaborative Study V.

Do-Young Ra, Moon-Soo Rhee, Yoon-Dong Kim and Keon-Joong Hwang*
Korea Ginseng and Tobacco Research Institute, Taejon 305-345, Korea
(Received May 13, 1998)

ABSTRACT : **This study was conducted to evaluate the techniques of analyzing tobacco smoke by statistical treatment method for the analytical data through Asia Collaborative Study V. In addition to five smoke components analysis, consisting of TPM, water, nicotine, NFDPM, and puff count of four cigarettes samples, statistical parameters such as mean, standard deviation, box-and-whisker plots, h plots, k plots, regression coefficients, reproducibility (R), and repeatability (r) were also calculated. Analysis of water content of cigarette smoke was the most difficult task, whereas puff count analysis was the easiest as well recognized by all laboratories. Analysis of nicotine and puff count accounted for both the lowest and the highest variation among four parameters. The water coefficients indicated more randomness or variation in the slops. The NFDPM data exhibited both types of deviations from linearity. Water content of sample D indicated the highest difference between two single results and between two inter-laboratory test results. As a whole, KGTRI ranked higher in the analytical techniques for statistical evaluation of results when compared with the practices of 28 other laboratories.**

Key words : **statistic, cigarette smoke, nicotine, TPM, NFDPM**

Because of the recent increase in attention for human health caused by cigarette smoke, accurate information is indispensable on the amount of smoke components associated with cigarettes smoking harmful to human being. Nevertheless no standard procedures are available for cigarette testing. Cigarette smoking testing belongs to destructive test and even though appropriate samples from single cigarette makings were used, product variability is inevitable and has great impact. Variations in individual cigarette, i.e. weights, paper porosity, filter efficiencies and others will cause some effect on the data to be obtained.

Measurement of tar and nicotine in smoke is widely used to evaluate the quality of cigarette. Up to the present, all laboratories of tobacco industry has made continued measurements on the amount of tar and nicotine in smoke. In 1991, the International Organization for Standardization (ISO) accepted the CORESTA harmonized smoking methodology for tar and nicotine testing. It was apparent that many national standards bodies would soon adopt these standardized and internationally recognized method as their own regulatory procedures. In view of the impending changes, great number of personnel in the government laboratories as well as the industry felt the need for performing inter-laboratory testing prior to adoption of the new CORESTA procedures. The

*Corresponding author : *Korea Ginseng & Tobacco Research Institute, 302 Shinsung-Dong, Yusong-Ku, Taejon 305-345, Korea*

first Asian Collaborative Study was completed in 1993 for determining the degree of agreement of nicotine-free dry particulate matter(NFDPM) and nicotine among the participating laboratories based on ISO smoking procedures (Bright, 1994). As the game advances, this collaborative study was expanded to include additional laboratories and country. In 1996, the fifth Asia Collaborative Study conducted a test in determining of NFDPM and nicotine in the smoke condensate of cigarettes in accordance with the ISO standards(Rhee, 1997). These results provided an opportunity for an open exchange of technical information regarding the ISO smoking methodology and furthermore, promotes standardized testing practices. By analyzing a set of common samples throughout all laboratories under the same conditions, overall variability of the methodology can be evaluated. These results will provided benefits for needed precise data on enforcement application of regulations for government agencies and essential assurance that their products are within specific limits for manufacturers.

This study was undertaken to evaluate laboratory techniques of smoke analysis being practiced at the Korea Ginseng & Tobacco Research Institute (KGTRI) applying statistical methodology of analyzing results of total particulate matter(TPM), water, nicotine, NFDPM in smoke in cooperation whth the fifth asia collaborative study.

MATERIALS AND METHODS

Cigarette samples : All participants received 2 cartons each of the 4 different brands of cigarettes. Four different brands were used in this study. Details of the samples were as follows:

Code	Brand	Supplier	Intended ISO Tar (mg/cig.)
A	Dunhill Ultimate Light	Rothmans	1
B	Mild Seven Lights	Japan Tobacco	3
C	Philip Morris Lights	Philip Morris	5
D	Kent Milds	Brown & Williamson	10

Experimental system : Twenty-nine laboratories including KGTRI capable of performing the ISO Smoking Method(ISO 4387, 3308 and 1991) participated in this study. The majority of the laboratories(24) used a Linear 20-Port (Filtrona-Type) smoking machine. However, other five laboratories, including KGTRI, used Rotary 20-Port (Borgwaldt-Type) in this study. Each laboratory was requested to determine the TPM, water, nicotine and NFDPM in the smoke and puff count using a routine analytical smoking machine in accordance with the ISO standards. The ISO 3402, 1991 and 3308 specify the ambient conditions for the conditioning and testing of cigarette samples were as follows:

Conditioning atmosphere:
 22 ± 1℃, and RH 60 ± 2 %
Testing atmosphere:
 22 ± 2℃, and RH 60 ± 5 %
Average air velocity at the cigarette level:
 200 ± 30 mm/sec

The testing conditions and air velocity surrounding the cigarettes of KGTRI were 22℃, RH 60%, and 190 mm/sec, respectively.

Statistical analysis methodologies

- Exploratory data analysis: Box-and-Whisker Plots
- Comparative Graphs: Comparative Table
- Ranking: Ranking Table
- Data Consistency: *h* Plots, *k* Plots
- Regression : Regression Coefficients
- Reproducibility(R) and Repeatability(r): R & r table

RESULTS AND DISCUSSION

There was no single statistical tool that was as powerful as a well chosen graph or table for transforming data into information and communication patterns and relationships to technical and non technical audiences. Average, standard deviation (STD), and the ranking scores for KGTRI in the measurement of the five parameters were presented in Table 1. Also, the difference of the 29 laboratory for the five parameters from their respective overall average, maximum value, minimum minimum

Table 1. Analytical Results of TPM, Water, Nicotine, NFDPM, and Puff Count for Four Cigarette Samples by KGTRI

	Sample											
	A			B			C			D		
	Average (mg/cig.)	STD*	Rank**	Average (mg/cig.)	STD	Rank	Average (mg/cig.)	STD	Rank	Average (mg/cig.)	STD	Rank
TPM	0.95	0.138	11	3.50	0.181	22	5.52	0.173	23	12.64	0.258	21
Water	0.12	0.045	20	0.24	0.033	16	0.44	0.030	20	1.64	0.068	22
Nicotine	0.09	0.009	11	0.32	0.013	25	0.40	0.008	26	0.80	0.010	25
NFDPM	0.75	0.121	7	2.94	0.147	19	4.69	0.148	21	10.21	0.220	17
Puff count	8.15	0.214	28	8.08	0.366	29	6.47	0.137	28	7.53	0.082	27

* : Standard Deviation

** : Ranked low to high by average values of each parameters in 29 laboratories

value, and percent difference were shown in Table 2. All of the average value except NFDPM of sample A for the five parameters by KGTRI were higher than that of overall average. The analytical value of nicotine for four samples had the lowest STD among the five parameters. Puff count of sample A and B, TPM of sample C and D had the highest STD among the five parameters. Analysis of nicotine indicated the lowest variation, and analysis of puff count showed the highest variation in four parameters.

A rank is used to indicate the relationship between an individual and the group. For individuals that are tied, the arithmetic mean is used to indicate their rank. a rank of 1 is assigned to the smallest amount, rank 2 to the next smallest, etc. It was assumed that laboratories which were consistently high or consistently low, had a systematic error (Youden, 1975). The ranking data for KGTRI exhibited a significantly higher level of TPM, nicotine, and NFDPM of sample A, conversely, KGTRI's data indicated a significantly lower level of puff count of all samples, and nicotine of sample B, C, D than that of the other laboratories. Table 2 was used to ascertain the difference of all laboratory from the overall average of all labor-

Table 2. Summary of Overall Average, Maximum Value, Minimum Value, and Percent Difference for Four Cigarette Samples by 29 Laboratories

(unit : mg/cig.)

	Sample															
	A				B				C				D			
	OA[1]	Max[2]	Min[3]	PD[4]	OA	Max	Min	PD	OA	Max	Min	PD	OA	Max	Min	PD
TPM	0.99	1.28	0.33	±67%	3.34	3.82	2.44	±27%	5.30	5.69	4.51	±15%	12.26	13.27	10.66	±13%
Water	0.10	0.34	-0.04	±240%	0.23	0.52	0.10	±120%	0.39	0.86	0.11	±122%	1.38	2.36	0.69	±71%
Nicotine	0.10	0.13	0.07	±38%	0.30	0.35	0.23	±21%	0.38	0.44	0.33	±17%	0.76	0.90	0.71	±19%
NFDPM	0.79	0.99	0.28	±65%	2.80	3.12	2.00	±29%	4.54	4.87	3.93	±13%	10.12	11.31	8.90	±12%
Puff count	7.31	8.40	6.48	±15%	7.15	8.08	6.23	±13%	6.07	6.80	5.48	±12%	7.13	7.59	6.20	±13%

1) OA : Overall Average, 2) Max : Maximum value, 3) Min : Minimum value, 4) PD : Percent difference

atories for each parameter. The % difference was calculated as follows:

{(Lab. average - Overall average) / Overall average} x 100

The analytical data of water and puff count had the highest and the lowest percent difference, respectively. These results indicated that analysis of water from cigarette smoke was the most difficult task, whereas puff count analysis was the easiest as well recongized by all laboratories.

The Notched Box-and-Whisker Plots for the five measurement parameters of four samples by KGTRI were shown in Figure 1. The Box-and-Whisker Plot procedure is an exploratory data analysis tool that is useful in studying of symmetry, checking distributional assumptions, and detecting outliers. Notched Box-and Whisker Plot is a modification of the standard Box-and-Whisker Plot. A notch is added to each box at the median.

The length of the notch represents an approximate 95 percent confidence interval for the median. The width of the box is proportional to the square root of the number of observations on the group. We can perform pairwise comparison of the group medians at the 95 percent confidence level by examining whether two notches overlap. If tow notches overlap, the medians are not significantly different. If two notches do not overlap, the medians are significantly different. When the notch extends beyond a quartile the ends of the box fold in. The fold-in pattern is caused by the median extending past a quartile boundary. It happens when extreme values are found in the data or when your have an extremely small data set. Extreme values are those data points more than three interquartile ranges below the first quartile. KGTRI had no outside values, except puff count of sample B, in the five components of four samples.

As TPM, nicotine, and puff count of sample A had a wider distribution of analytical data, water of all four samples had the narrowest distribution.

The consistency statistics, *h* and *k* for the KGTRI were prepared in Table 3 according to laboratory by individual samples. The critical values of the consistency statistics *h* and *k* at the 0.5% significance level for the present were Puff count of all four samples showed higher value of

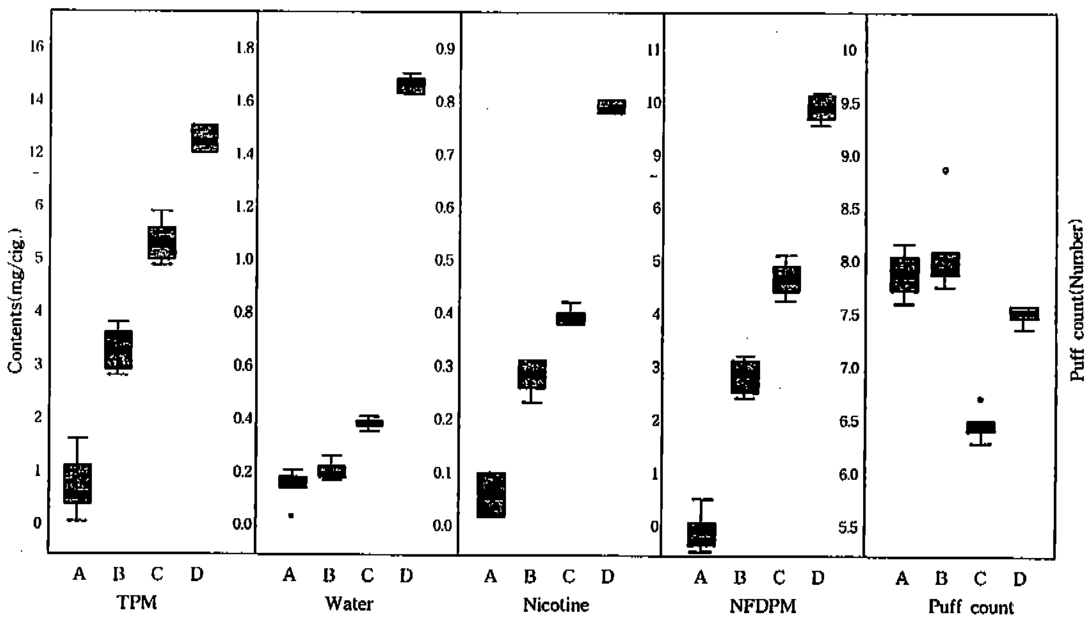

Fig. 1. Notched Box-and-Whisker Plot for five parameters of four samples by KGTRI.

Table 3. Data Consistency of TPM, Water, Nicotine, NFDPM, and Puff Count for Four Cigarette Samples by KGTRI

	Sample							
	A		B		C		D	
	h	k	h	k	h	k	h	k
TPM	0.20	0.77	-0.51	0.69	-0.72	0.51	-0.61	0.43
Water	-0.29	0.55	-0.08	0.30	-0.32	0.22	-0.71	0.29
Nicotine	0.39	0.21	-0.82	0.62	-0.90	0.34	-0.91	0.27
NFDPM	0.33	0.69	-0.44	0.59	-0.61	0.47	-0.18	0.45
Puff count	-2.16	0.83	-2.35	1.71	-1.70	0.69	-1.48	0.40

Table 4. Regression coefficients of analytical data of TPM, water, nicotine, and NFDPM by KGTRI

	Regression Coefficients			
	Intercept	Slop	R-squared	Standard Error of Estimate
TPM	-0.00	1.0335	99.99	0.071
Water	-0.02	1.2054	99.93	0.022
Nicotine	-0.01	1.0614	99.96	0.008
NFDPM	0.04	1.0091	99.95	0.106

obtained from statistical tables as ±2.64 and 1.80 respectively. This table gave a picture of the overall character of the variability of the test methods, as well as singling out particular results that require investigation.

The *h* and *k* consistency statistics were calculated in accordance with ASTM E-691(ASTM, 1992) from the following formula:

$h = d / s_x$

where *h* = The between laboratory consistency statistic,

d = The difference of an individual laboratory's average from the overall average of all laboratories

s_x = The standard deviation of the laboratories's averages

$k = s / s_r$

where *k* = the within laboratory consistency statistic,

s = Standard deviation of an individual laboratory

s_r = The repeatability standard deviation between laboratories

h and *k* comparing with other components. This results indicated that analytical data of puff count had a low consistency. Especially, *h* value was higher than *k* value in puff count means that the consistency between laboratory was lower than that of within laboratory.

A Linear regression model($y = a + bx$) was used to provide a general description of the bias and random error in the data from each laboratory plotted against the overall average for all the laboratories at several levels. This model was applied to TPM, water, nicotine and NFDPM. The puff count data were not regressed because these data were not level dependent. Using a linear x - y plot, if there was no constant error in the analytical process, the various points should tend to lie closely along a line passing through the origin. If there was a constant error or bias(a), all the points are displaced by a constant amount and will not pass through the origin. This was shown in the intercept values deviation from zero. Since the slop(b) is independent of the bias or intercept, the slop of a straight line is the change on y for a unit change in x. If a given laboratory had perfect agreement across the eight levels, then the slop would be equal to 1 and intercept would be equal to 0. My laboratory's table of regression coefficients were shown in Table 4. Using this regression coefficient tables, water and NFDPM data showed some bias or variation in the intercepts. However, the water coefficients indicated more randomness or variation in the slops. The NFDPM data exhibited both types of deviations from linearity

Reproducibility(R) and Repeatability(r) are two measures of precision and both are used to describe the variability of a test method. Reproducibility(R) is a closeness of agreement between individual results obtained with the same method on identical test material but under different condition. In our study, the difference between two single results on matched samples reported by two laboratories will not differ by more than R, on average, than once on twenty cases under normal and correct operation of the method. In other words, R, can be defined as an established value at a given tar level below which the absolute difference between two inter-laboratory test results may be expected to fall with a specified probability (0.05). Repeatability(r) is a closeness of agreement between successive results obtained with the same method on identical test material and under the same conditions. In our study, the difference between two single results found on matched samples by one operator using the same apparatus within the shortest feasible time interval will not exceed the r, on average, more than on in twenty cases under normal and correct operation of the method. In other words, r, can be defined as an established value at a given tar level below which the absolute difference between two intralaboratory test results may be expected to fall with a specified probability (0.05). Table 5 presented repeatability(r) and reproducibility(R) of analytical data by KGTRI. Our results indicated that water of sample D had the highest difference between two single results and between two interlaboratory test results. The concepts of reproducibility and repeatability by themselves were insufficient to serve as a representative of the interlaboratory situation underlying the data. Reproducibility and repeatability measures fail to provide information on the behavior of individual laboratories. It was difficult to compare R and r based on data generated on previous interlaboratory collaborative study has a varying number of participating laboratories and sample levels which would make comparisons invalid. When considering the interpretation of r and R, it is vitally important to remember that r and R do not take account of any long term variability of the product itself. In the context of long-term product variability, values of r and/or R will only be applicable to estimating confidence regions, etc. of such a product taken at a single point in time. To obtain an estimate of the overall variability of the product it must be sampled continuously for many months, and the components of variance due to the medium and long-term sources of variability incorporated into the statistical analysis of the data.

As a whole, KGTRI ranked higher in the analytical techniques for statistical evaluation of results when compared with the practices of 28

Table 5. Repeatability(r) and Reproducibility(R) of TPM, Water, Nicotine, NFDPM, and Puff Count for Four Cigarette Samples by KGTRI

	Sample							
	A		B		C		D	
	r	R	r	R	r	R	r	R
TPM	0.112	0.160	0.165	0.318	0.207	0.288	0.452	0.730
Water	0.053	0.095	0.072	0.103	0.087	0.185	0.587	1.385
Nicotine	0.023	0.046	0.014	0.027	0.021	0.031	0.065	0.138
NFDPM	0.111	0.157	0.159	0.252	0.200	0.202	0.283	0.548
Puff count	0.141	0.469	0.142	0.453	0.122	0.322	0.128	0.342

other laboratories. Further, statistical evaluation through collaborative study should be continued on a routine basis to be assured that our data remain comparable. Also, the routine exchange of information creates understanding and leads to the development of great precision.

CONCLUSION

This study was undertaken to evaluate the techniques of smoke analysis laboratory in KGTRI by statistical methodology of the analytical results of TPM, water, nicotine, NFDPM in smoke and puff count through the fifth Asia collaborative study. Analysis of water content of cigarette smoke was the most difficult task, whereas puff count analysis was the easiest as well recognized by all laboratories. Analysis of nicotine and puff count accounted for both the lowest and the highest variation among four parameters. The water coefficients indicated more randomness or variation in the slops. The NFDPM data exhibited both types of deviations from linearity. Water content of sample D indicated the highest difference between two single results and between two inter-laboratory test results. As a whole, KGTRI ranked higher in the analytical techniques for statistical evaluation of results when compared with the practices of 28 other laboratories. Statistical evaluation through collaborative study should be continued on a routine basis to be assured that our data remain comparable. Also, the routine exchange of information creates understanding and leads to the development of great precision.

REFERENCES

ASTM E691 (1992) Standard on precision and bias for various applications, Fourth Edition.

Bright, Cynthia C. (1994) Asia Collaborative Study #2, Philip Morris Products Inc.

CORESTA (1991-3) Special Smoking Methods, Information Bulletin Cooperation Center for Scientific Research Relative to Tobacco.

Mandel, J. (1991) Evaluation and Control of Measurement. Marcel Dekker, Inc., N.Y., U.S.A.

Rhee, M. S. (1997) Asia Collaborative Study #5, Korea Ginseng & Tobacco Research Institute.

Youden, W. J. and Steiner, E. H. (1975) Statistical Manual of the Association of Official Analytical Chemists, *Association of Official Analytical Chemists*.