



DETERMINATION OF THE STATISTICAL ROBUSTNESS OF A GAS CHROMATOGRAPHIC METHOD DESIGNED FOR HEROIN PROFILING BY PATTERN ANALYSIS USING CASE SAMPLES OF UNKNOWN ORIGINS

Kar-Weng CHAN^{1,2}, Guan-Huat TAN¹, Richard C. S. WONG¹

¹ *Department of Chemistry, University of Malaya, Kuala Lumpur, Malaysia*

² *Department of Chemistry Malaysia, Ministry of Science, Technology and Innovation, Petaling Jaya, Malaysia*

Abstract

This paper proposes pattern analysis for assessing the robustness of an analytical method *via* a statistical means. Two gas chromatographic methods optimized for heroin profiling were used to compare the method robustness. An established method and a relatively poor method were used respectively as a standard positive control method (PCM) and a negative control method (NCM) to demonstrate how a statistically robust method should perform in data interpretation. A total of 43 illicit heroin samples of unknown origins were analysed using both methods and the data distributions representing the relationships between the samples were obtained by principal component analysis. In the PCM, the data maintained its general pattern irrespective of how the data were manipulated. Pattern distortion was shown in the NCM and this suggested that the method was not statistically robust for heroin profiling.

Key words

Statistical robustness; GC; Heroin; Profiling.

Received 17 December 2011; accepted 28 March 2012

1. Introduction

Forensic drug profiling is a powerful tool for tactical and strategic intelligence. In the realm of narcotics, many analytical methods such as gas chromatography (GC), high performance liquid chromatography (HPLC), and electrophoresis have been widely used for profiling since they have superior separation power over other techniques. GC in particular has been vital in illicit drug profiling since most compounds of interest are relatively volatile. With suitable statistical techniques, the chemical data are useful to establish the relationships between samples, which in turn can be used to infer the trafficking route, manufacturing plant and geographical origin of the samples. Specifically,

this often involves the analysis of compounds that are linked to the synthesis or processing method. In the case of illicit heroin, acidic and neutral manufacturing impurities have been profiled, and they proved useful for clustering hundreds of samples according to the geographical regions [5, 7]. Natural opium alkaloids such as morphine and noscapine can also be profiled to trace the agricultural origins [3, 10]. For profiling, the steps involved in GC method validation can be found in the literature [1, 9, 12, 13]. Some of these studies were devoted to the investigation of method robustness through inter-laboratory precision.

As the ultimate use of drug profiles is to determine the relationships between samples, thus many statistical classification techniques have been established for

this aim [6, 8, 11]. It is also known that drug profiling does not employ chemical standards to calibrate the instrument, so errors may be present when normalization of peaks is utilized for sample comparison. To ensure the validity of the target analytical and statistical methods, statistical validation of the analytical data must be performed [4]. Conventionally, a batch of known samples can be used to evaluate whether they can be statistically clustered according to their groups. The correctness of grouping will reveal the robustness of the overall analytical and statistical method. In other words, if all the samples are successfully clustered into their real groups, the overall method is very robust. Unfortunately, this is not possible if the laboratory lacks samples of known histories. As a result, this study proposes to employ pattern recognition to evaluate the robustness of the overall method by comparing the data acquired from a questioned method against that of established method. In this paper, the positive control method (PCM) is the established method while the negative control method (NCM) is the questioned method to be assessed. A robust method should show consistent patterns irrespective of whether calibration is performed. A poor method on the other hand is expected to display pattern distortion. The NCM was in fact a planned method that was deliberately developed for this study to show such distortion. It was chosen because it initially appeared to be good and was believed to be able to give reliable results for illicit heroin profiling. However, this paper eventually revealed its poor robustness in terms of sample classification although some of its validation criteria were marginally fulfilled.

In this study, a statistical approach was proposed to evaluate the robustness of a GC method for heroin profiling. Since the relative relationships are more important than the absolute relationships between the samples, it was thus decided to treat all the samples used in this study as if they were of unknown origins, in order to supervise the pattern distribution. A total of 43 locally seized case samples were analysed using both methods by GC. The robustness of both methods were assessed by decomposing the GC data into two components by principal component analysis (PCA). The statistical procedure eventually demonstrated how method robustness is achieved based on the data distribution pattern.

2. Materials and method

2.1. Standards and solvents

Acetylcodeine hydrochloride was prepared by the Department of Chemistry Malaysia. Codeine phosphate, morphine hydrochloride and heroin hydrochloride were purchased from Johnson Matthey Macfarlan Smith. 6-monoacetylmorphine (6-MAM) hydrochloride was commercially obtained from Lipomed and 2,2,2 triphenyl acetophenone (internal standard, IS) from Aldrich Chemical Company. HPLC grade methanol and analytical reagent grade chloroform, both manufactured by Fisher Scientific, were used.

2.2. Gas chromatography-flame ionization detector (GC-FID)

The GC methods used in this study were as follows:

1. The PCM [2] employed an HP6890N GC-FID pre-installed with a J&W HP Ultra 2 (length 25 m, i.d. 0.20 mm, film thickness 0.33 μm) capillary column. Chromatographic separation was accomplished by holding the oven temperature at 240°C for 1 min and heating up to 270°C at the rate of 12°C/min. The oven was then held for 8 min at this temperature. Injector and detector temperatures were set at 290°C with a split ratio of 40:1 and injection volume of 1 μl .
2. The NCM employed an HP6890N GC-FID pre-installed with a J&W HP 5 (length 30 m, i.d. 0.25 mm, film thickness 0.25 μm) capillary column. Chromatographic separation was accomplished by holding the oven temperature at 250°C for 4 min and heating up to 260°C at the rate of 7°C/min and to 280°C at the rate of 6°C/min. Injector and detector temperatures were set at 280°C with a split ratio of 40.5:1 and injection volume of 1 μl .

2.3. Partial method validation

The injector and detector temperatures, split ratio, flow rate and temperature programming were optimized using the J&W HP Ultra 2 and J&W HP 5 columns. They were optimized to achieve separation that offers good peak resolution and sharp peak shapes within a short analysis time. The optimization details of these two methods have been reported [2]. Subsequently, the performance of each optimized method was verified based on five critical aspects: repeatability, reproducibility, linear working range, linearity and recovery. A standard solution containing approximate-

ly 0.05 mg/ml codeine (CD), 0.08 mg/ml morphine (MP), 0.10 mg/ml acetylcodeine (AC), 0.26 mg/ml 6-monoacetylmorphine (MM) and 0.31 mg/ml heroin (HR) respectively, which were the preferred concentrations, was prepared in methanol:chloroform (1:9) containing 2,2,2-triphenyl acetophenone (internal standard, IS). After optimization, the PCM and NCM were determined based on their performance and the optimum concentrations of the IS were set at 0.18 mg/ml and 0.20 mg/ml respectively for the two methods. A single mixture was used to study within-day and between-day precisions with six injections respectively. Linearity was studied using a series of eight concentration levels prepared from a single mixture of the five analytes through dilution, while the IS was held constant. Each concentration level was injected at least twice. The analytes were spiked into the sample matrix for recovery studies.

2.4. Sample preparation

A total of 43 illicit heroin samples were analysed. For each sample, approximately 80–85 mg homogenized substance was weighed in a dissolution vessel to which 10 ml of IS solution was added. Subsequently, each aliquot was analysed in duplicate.

2.5. Data analysis

Quantitative data of the five major opium alkaloids (codeine, morphine, acetylcodeine, 6-monoacetylmorphine and heroin) were obtained in two forms: peak area relative to IS and concentration (mg/ml). The dataset was subjected to Minitab 15 (USA) for multivariate analysis after the data were pretreated using Microsoft Excel. Pretreatments included forming suitable quotients and normalization. Standardization was done in the correlation mode during principal component analysis. Principal components 1 and 2 were used for comparison.

3. Results and discussion

3.1. Performance of the methods

Peak resolution was first examined by comparing two chromatograms of a single sample analysed with the two methods. The NCM (Figure 1b) demonstrates unsatisfactory peak shapes for the target analytes, whereby peak tailing with morphine and heroin is notable. Hence, errors are inherently associated with the NCM. This is largely attributed to peak integration

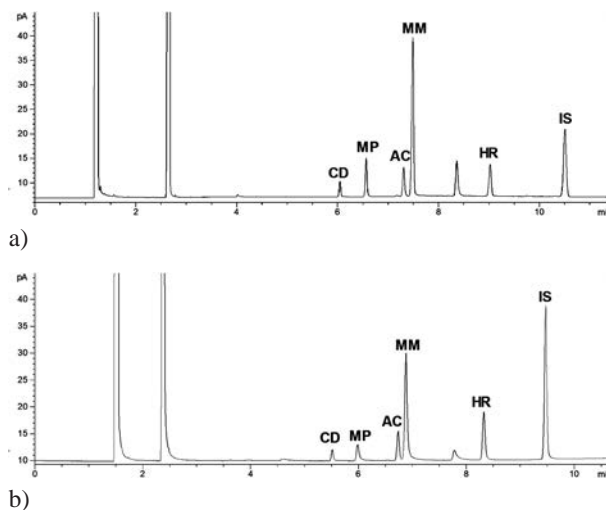


Fig. 1. Chromatograms of a case sample analysed using: a) PCM and b) NCM.

when base-to-base definition of a peak becomes uncertain. For quantitative assessment using GC, a sharper peak with insignificant peak tailing (Figure 1a) is desirable for each target compound. Therefore, the NCM did not fulfil this criterion.

The performance of the system was evaluated quantitatively based on the five critical aspects of validation. Repeatability is a measure that evaluates the stability of the instrument within the same day, whereas reproducibility is a more reliable measure for assessing instrumental variation on a between-day basis. According to Table I, the repeatability and reproducibility calculated from the six injections, which are expressed as relative standard deviation (*RSD*) are satisfactory for both methods. This is because the errors are better than $\pm 15\%$, the acceptance level set by the United Nations Office of Drugs and Crime [10].

A wider linear working range within which the detector responds to the known concentration in a linear manner is preferable for accurate quantification. In this regard, the PCM showed very good linear responses with $r^2 > 0.999$ within 50–150% working concentration ranges. However, the linear working ranges of the NCM were significantly narrower than those of the PCM. A non-linear response with r^2 between 0.96–0.98 was the major issue of the NCM. The r^2 values could hardly be accepted as a good method as the analytes cannot be quantified accurately even though they fall within the working ranges. At certain concentrations, the NCM tended to present high or low readings that give rise to the poor r^2 .

The accuracy of the method was measured by the recovery test. Apparently, the PCM displayed better

TABLE I. IMPORTANT ANALYTICAL PARAMETERS IN PCM AND NCM^a

Compound	PCM					NCM				
	Repeatability ¹ (n = 6)	Reproducibility ² (n = 6)	Range ³ [mg/ml]	Linearity r^2	Recovery [%]	Repeatability ¹ (n = 6)	Reproducibility ² (n = 6)	Range ⁴ [mg/ml]	Linearity r^2	Recovery [%]
CD	1.19	1.33	0.01–0.20	0.9998	99.01	1.16	1.56	0.02–0.05	0.9805	101.03
MP	1.03	4.65	0.01–0.20	0.9997	101.70	3.02	5.29	0.05–0.11	0.9601	105.26
AC	1.19	0.32	0.05–0.10	0.9998	99.25	0.39	0.85	0.05–0.11	0.9750	101.67
MM	1.07	1.47	0.05–1.00	0.9997	100.66	0.77	1.98	0.15–0.34	0.9742	106.51
HR	1.09	1.19	0.05–1.00	0.9999	101.37	0.56	1.11	0.20–0.47	0.9802	98.06

^a All GC data are expressed as peak area relative to IS (area ratio).

¹ Repeatability was carried out (determined) using a single mixture containing the five target compounds at the working concentrations. The mixture was injected six times consecutively.

² Reproducibility was carried out as for repeatability, but the six injections were performed on different days.

³ The desired working range covered 50%–150% of the target concentration. The range showed a linear response.

⁴ From the series of dilutions, only the consecutive points that give the best r^2 value for the regression were used to define the working range. The range is narrow and not linear.

recoveries than the NCM, which indicated excessive recoveries for morphine and 6-monoacetylmorphine.

Other validation aspects for the PCM have been reported in detail [2]. The overall validation study shows that the PCM is a valid method for the profiling of heroin and hence suitable to serve as the standard positive control for robustness assessment. In contrast, the performance of the NCM is unacceptable (with the above-mentioned undesirable linear ranges, linearity and recoveries) and therefore was chosen as the negative control in this study.

3.2. Evaluation of method robustness using PCA

The 43 case samples were analysed using both methods in duplicate. They were selected because these samples contained all the target peaks. Hence the presence of all the target peaks would help minimize statistical errors arising from zero-values (absence of peaks) during statistical treatment. The GC data were obtained in two forms. The first method is by using the “peak area relative to IS” through which errors due to inconsistent split ratios and evaporative losses can be greatly minimized. The second method is to take the concentration in mg/ml based on the one-point calibration performed daily using area ratio (peak area relative to IS). In addition to the errors addressed by the first method, the second method could also compensate for any unknown variables that would affect the GC system. In other words, the system was corrected once it was calibrated using known standards.

Data normalization was performed on the GC data, targeting five major opium alkaloids found in the il-

licit heroin in order to minimize environmental and statistical errors as well as the cutting effects due to the presence of adulterants. As heroin (diacetylmorphine) is not stable, the decomposition of heroin to 6-monoacetylmorphine and sometimes further deacetylation to morphine must be taken into consideration. Hence, the data were tentatively normalized according to CD / (MM + HR), MP / (MM + HR), AC / (MM + HR), CD / (MP + MM + HR), AC / (MP + MM + HR) and (CD + AC) / (MP + MM + HR) to compensate for the effects of decomposition. The sum of the morphine contents as the denominator will help minimize the change in the contents due to decomposition. Thus, the data from these two methods can be compared without prejudice. These pre-treated data were decomposed by PCA to show the relationships between the samples. The variability of the data retained by the principal component is expressed as % variability (%V) on the score plot.

According to Figure 2a and 2b, the score plots show that the general patterns of data distribution seem to be vertically rotated when the GC data are interpreted in the two different forms of readings. However, the relationships (in terms of distance) between the data points are relatively in close agreement irrespective of whether the system was calibrated or not. For example, “65” is close to “50” and “111”, and it maintains a relatively long distance from “277” in both plots. Insignificant differences in distance were also observed on the score plots. In this case, the negligible deviations of the patterns in Figure 2a and 2b are believed to be some small random errors that should be corrected using standards on a daily basis. In this regard, the dis-

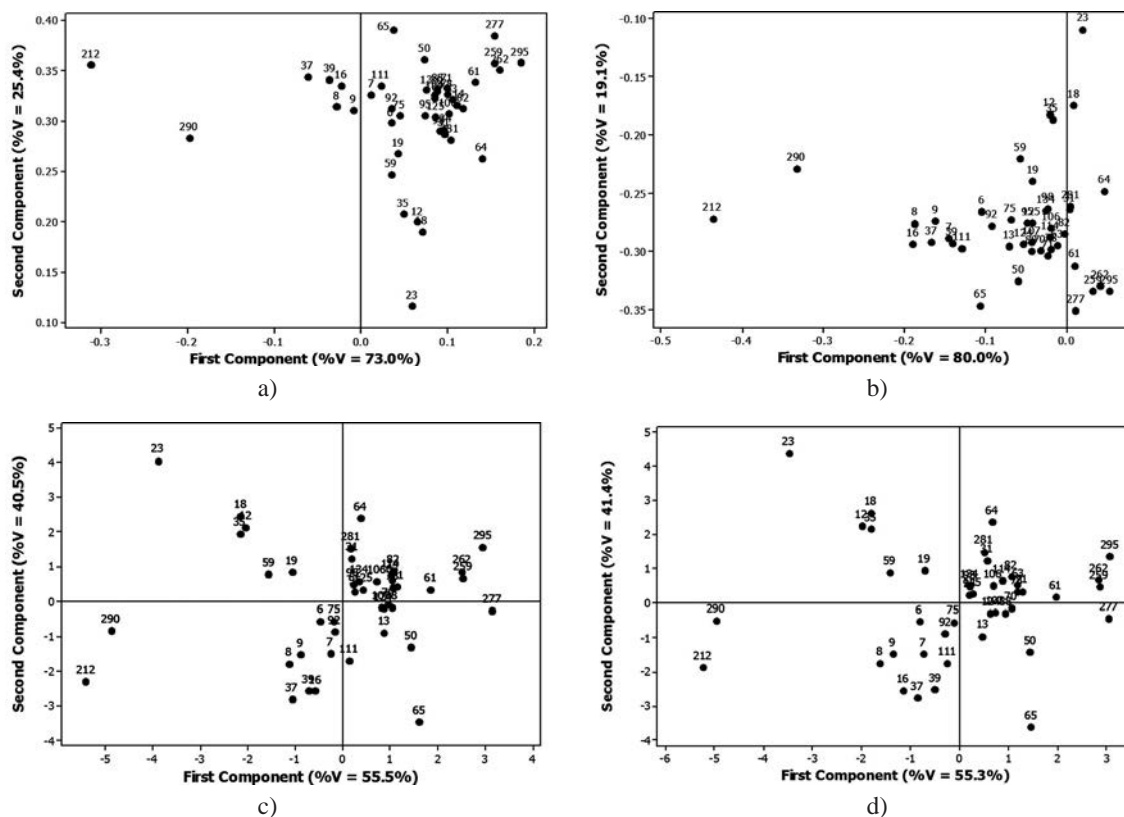


Fig. 2. Score plots of 43 data points obtained from the PCM and decomposed by PCA according to: a) peak to IS, b) concentration in mg/ml, c) peak to IS with standardization and d) concentration in mg/ml with standardization. The labelling is the original case assignment. The patterns in a) and b) are due to the difference in the types of data readings. Using standardization, the pattern in c) and d) becomes very similar and the relationships between data points are not affected.

tribution of the data in Figure 2b is presumably more reliable than that in Figure 2a because the daily errors have been corrected by the standards.

In order to view both plots on the same scale, standardization is an ideal method to transform the data in different types of readings to a common scale. To do this, each normalized variable was treated as the variable-*i*, e.g. $CD / (MM + HR)$ and all individual variables-*i* were then divided by the standard deviation of that variable-*i*, e.g. standard deviation computed from all cases for $CD / (MM + HR)$. When the standardized data were decomposed by PCA in Figure 2c and 2d, the distribution pattern and the relationships between the samples revealed in the new score plots are generally similar. Hence, the data distribution in the normalized form or in the standardized form of the normalized data is consistent with one general relational pattern irrespective of whether the GC data were corrected by calibration or not. Therefore, the established method/PCM is very robust and free of noticeable analytical errors. The distribution and relationships depicted

in the plots were used as a standard map against which the NCM method was compared.

Measurable pattern distortions are observed in the data obtained from the NCM. According to Figure 3a and 3b, data in the normalized form show different sample relationships between uncalibrated and calibrated readings. When the data were standardized and decomposed by PCA, the patterns in Figure 3c and 3d are still not in close agreement. The calibrated data distribution pattern is believed to be distorted as compared to the uncalibrated data distribution pattern when daily calibration is considered for analysis. This means that the analytical errors occurring in the NCM are significant even though they were corrected by daily calibration. As such, different interpretations (in terms of data relationships) could be generated from the poor method. Sample relationships become very dependent on the types of readings used in PCA. Hence, methods with poor robustness will reveal pattern distortion such as that demonstrated in the poor method/NCM when different types of readings are used.

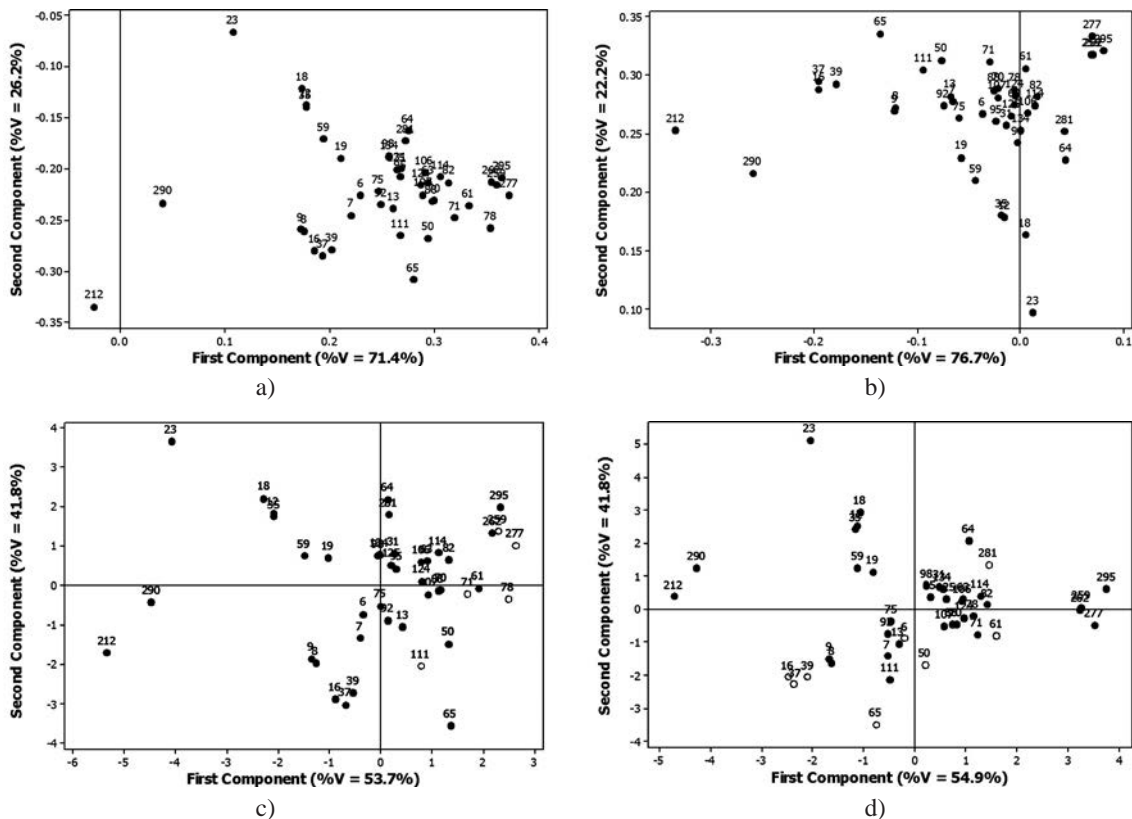


Fig. 3. Score plots of 43 data points obtained from the NCM and decomposed by PCA according to: a) peak to IS, b) concentration in mg/ml, c) peak to IS with standardization and d) concentration in mg/ml with standardization. Distortion is observed when the pattern is compared with that of the valid method. Measurable impacts were experienced by data points in bright circles.

A comparison between Figure 2 and 3 shows the extent to which the data differ in the two systems. For the standardized data, measurable distortion impacts were experienced by data points in bright circles when the score plots in Figure 3c and 3d are compared against the corresponding score plots in Figure 2c and 2d. Figure 3c demonstrates some discrepancy in the upper right portion as compared to Figure 2c. Also, the NCM has relatively loosely packed data. A severe shift was observed in “78” which was displaced from the cluster in the middle area. Although the possible errors in the NCM were corrected by daily calibration, the data points in Figure 3d are somewhat inconsistent with those in Figure 2d. Obvious shifts were observed in “50”, “61” and “65” and the general pattern in the lower left portion of the score plot. For example, the cluster comprising “16”, “37” and “39” is supposed to be closely related to “7” and “111” but they have been incorrectly separated from the latter in Figure 3d. Taking Figure 2d as the most reliable pattern, the data of the NCM generally show distorted parts in their pattern and altered relationships. Misinterpretation can

arise if this unreliable method is used for establishing sample relationships.

4. Conclusion

In summary, an analytically valid method should also perform statistically well in data interpretation. The PCM is a good example that shows its analytical and statistical robustness irrespective of how the data are manipulated (e.g. calibrated *versus* non-calibrated data). The NCM, on the other hand, is indeed a poor method. It also shows that analytical errors associated with the method will lead to misinterpretation in pattern analysis. Hence, this paper encourages forensic chemists to evaluate the robustness of a questioned method based on a well established method using pattern analysis.

Acknowledgements

The authors would like to thank the University of Malaya for the IPPP (PV004/2011B) grant, the Department of Chemistry Malaysia (MOSTI) for the facilities and Y.S.P. Industries (M) Sdn. Bhd. for the standards to carry out this project.

References

1. Andersson K., Jalava K., Lock E. [et al.], Development of a harmonised method for the profiling of amphetamines III. Development of the gas chromatographic method, *Forensic Science International* 2007, 169, 50–63.
2. Chan K. W., Tan G. H., Wong R. C. S., Gas chromatographic method validation for the analysis of major components in illicit heroin seized in Malaysia, *Science & Justice* 2012, 52, 9–16.
3. Dufey V., Dujourdy L., Besacier F. [et al.], A quick and automated method for profiling heroin samples for tactical intelligence purposes, *Forensic Science International* 2007, 168, 108–117.
4. Esseiva P., Dujourdy L., Anglada F. [et al.], A methodology for illicit heroin seizures comparison in a drug intelligence perspective using large databases, *Forensic Science International* 2003, 132, 139–152.
5. Morello D. R., Cooper S. D., Panicker S. [et al.], Signature profiling and classification of illicit heroin by GC-MS analysis of acidic and neutral manufacturing impurities, *Journal of Forensic Sciences* 2010, 55, 42–49.
6. Myors R., Well R. J., Skopec S. V. [et al.], Preliminary investigation of heroin fingerprinting using traces element concentrations, *Analytical Communication* 1998, 35, 403–410.
7. Neumann H., Gloger M., Profiling of illicit heroin samples by high-resolution capillary gas chromatography for forensic application, *Chromatographia* 1982, 16, 261–264.
8. Ratle F., Terrettaz A. L., Kanevski M. [et al.], Pattern analysis in illicit heroin seizures: A novel application of machine learning algorithms [paper presented at European Symposium on Artificial Neural Networks, Bruges, Belgium, 2006].
9. Strömberg L., Lundberg L., Neumann H. [et al.], Heroin impurity profiling: A harmonization study for retrospective comparisons, *Forensic Science International* 2000, 114, 67–88.
10. Terrettaz-Zufferey A. L., Ratle F., Ribaux O. [et al.], Pattern detection in forensic case data using graph theory: Application to heroin cutting agents, *Forensic Science International* 2007, 167, 242–246.
11. United Nations Office of Drugs and Crime, Guidance for validation of analytical methodology and calibration of equipment used for testing of illicit drugs in seized materials and biological specimens, 2009, Vienna: United Nations.
12. van Deursen M. M., Lock E. R. A., Poortman-van der Meer A. J., Organic impurity profiling of 3,4-methylenedioxyamphetamine (MDMA) tablets seized in the Netherlands, *Science & Justice* 2006, 46, 135–152.
13. Visser H. A. A. H., Visser-van Leeuwen M., Huizer H., Residual solvents in methylenedioxyamphetamine tablets as a source of strategic information and as a tool for comparative analysis: the development and application of a static headspace gas chromatography/mass spectrometry method, *Bulletin on Narcotics* 2005, 67, 167–182.
14. Zhang D., Shi X., Yuan, Z. [et al.], Component analysis of illicit heroin samples with GC/MS and its application in source identification, *Journal of Forensic Sciences* 2004, 49, 81–86.

Corresponding author

Kar-Weng Chan
Department of Chemistry Malaysia
46661 Petaling Jaya, Malaysia
e-mail:chankarweng@yahoo.com

ZASTOSOWANIE STATYSTYKI DO OCENY RZETELNOŚCI METOD CHROMATOGRAFICZNYCH STOSOWANYCH W PROFILOWANIU HEROINY W OPARCIU O WYNIKI UZYSKANE DLA PRÓBEK O NIEZNANYM POCHODZENIU

1. Wstęp

Profilowanie narkotyków jest jednym z ważniejszych narzędzi w taktycznej i strategicznej pracy organów ścigania. Powszechnie stosowane są w tym celu różne metody rozdzielania, a wśród nich chromatografia gazowa (GC), wysokosprawna chromatografia cieczowa (HPLC) oraz elektroforeza. Charakteryzują się one zdecydowanie lepszą zdolnością rozdzielczą niż inne techniki. Szczególnie istotna jest metoda GC, ponieważ większość związków zawartych w narkotykach to związki lotne. Dane uzyskane w trakcie analiz chemicznych, wsparte analizą z wykorzystaniem metod statystycznych, mogą być stosowane do określenia podobieństw między próbkami, a to z kolei jest pomocne przy ustalaniu dróg przetrzutu, miejsca produkcji narkotyków lub pochodzenia geograficznego analizowanej próbki.

W profilowaniu narkotyków bardzo często dokonuje się analizy danych odnoszących się do związków charakterystycznych dla metody produkcji ksenobiotyków. W przypadku profilowania heroiny, kwasowe i obojętne zanieczyszczenia powstające podczas procesu jej produkcji były stosowane w celu zaklasyfikowania setek próbek pochodzących z różnych regionów świata [5, 7]. Naturalne alkaloidy opium, takie jak morfina i noskapina, mogą być również wykorzystywane w celu profilowania, którego celem jest określenie regionu, z którego pochodzi opium [3, 10]. Opis poszczególnych etapów walidacji metody GC w przypadku profilowania narkotyków można znaleźć w literaturze przedmiotu [1, 9, 12, 13]. Część z tych prac poświęcona była badaniom powtarzalności międzylaboratoryjnej.

Jednym z celów profilowania narkotyków jest określenie podobieństw pomiędzy próbkami. Opisane już zostały różne procedury klasyfikacji stosowanych w tym celu [6, 8, 11]. Należy jednak zdawać sobie sprawę, że mogą wystąpić błędy w procesie profilowania, np. gdy porównując próbki, nie wykonuje się normalizacji pików lub gdy nie przeprowadza się kalibracji aparatury przy wykorzystaniu serii wzorców. Ponadto musi być wykonana walidacja danych analitycznych w celu potwierdzenia użyteczności stosowanej metody analitycznej i statystycznej [4]. Zazwyczaj w tym celu wykorzystywana jest grupa próbek o znanym pochodzeniu, ponieważ wówczas można sprawdzić, czy są one poprawnie zaklasyfikowane. Z kolei poprawność procesu klasyfikacji jest miernikiem użyteczności stosowanej procedury

profilowania narkotyków, tj. stosowanych łącznie metod: chromatograficznej i statystycznej. W przypadku, gdy w laboratorium znajdują się próbki o nieznanym pochodzeniu, to wykonanie takiej oceny nie jest możliwe. Dlatego też autorzy niniejszego artykułu proponują zastosowanie metod rozpoznawania obrazów w celu oceny stosowanej procedury profilowania poprzez porównywanie jej z rezultatami uzyskanymi dla metody profilowania o uznanej rzetelności. W artykule metodę o uznanej rzetelności, wykazującą niezmienną uzyskiwanych wyników pomimo tego, czy kalibracja była czy też nie była wykonywana, oznaczono jako PCM (ang. positive control method). Z kolei metodę, której użyteczność dopiero oceniano, oznaczono jako NCM (ang. negative control method). Metoda chromatograficzna oznaczona jako NCM została wybrana spośród różnych metod chromatograficznych, ponieważ wstępnie oceniono ją jako rzetelną i dającą użyteczne rezultaty podczas profilowania heroiny. Następnie dla potrzeb prowadzonych badań zmodyfikowano ją tak, aby wykazywała pewne błędy, tj. charakteryzowała się złą rzetelnością przejawiającą się w niepoprawnej klasyfikacji próbek. Należy dodać, że metoda NCM spełniała większość kryteriów walidacyjnych, chociaż na granicy przyjętych kryteriów.

W artykule zaproponowano zastosowanie metod statystycznych do oceny rzetelności metod GC użytych do profilowania heroiny. Uznano też, że względne zależności pomiędzy próbkami są bardziej istotne niż zależności bezwzględne. Ponadto w opisanych badaniach zdecydowano się traktować wszystkie analizowane próbki jako te, których pochodzenie nie jest znane i dlatego też nie mogą być one użyte w procesie uczenia używanego klasyfikatora.

43 próbki zarekwirowanej heroiny analizowano dwiema metodami GC. Rzetelność obu technik była oceniana przez zastosowanie metody analizy głównych składowych (PCA), a konkretnie przez analizę uzyskanych w jej wyniku dwóch pierwszych głównych składowych. W artykule wskazano również, jak dokonać oceny rzetelności metody na podstawie uzyskiwanych wyników analizy statystycznej.

2. Materiały i metody

2.1. Wzorce i rozpuszczalniki

Chlorowodorek acetylokodeiny został zsyntezowany na Wydziale Chemii Uniwersytetu Malezyjskiego. Fosforan kodeiny, chlorowodorek morfiny i chlorowodorek heroiny zostały zakupione w firmie Johnson Matthey Macfarlan Smith. Chlorowodorek 6-monoacetylmorfiny (6-MAM) zakupiono w firmie Lipomed, a 2,2,2-trifenyloacetofenon (standard wewnętrzny, IS) w firmie Aldrich Chemical Company. Metanol o czystości HPLC oraz chloroform cz.d.a. wyprodukowała firma Fisher Scientific.

2.2. Chromatograf gazowy z detektorem płomieniowo jonizacyjnym (GC-FID)

W badaniach zastosowano następujące metody chromatograficzne:

1. W metodzie PCM [2] użyto chromatograf HP6890N GC-FID z kolumną kapilarną J&W HP Ultra 2 (długość 25 m, średnica wewnętrzna 0,20 mm, grubość filmu 0,33 μm). Rozdział chromatograficzny przeprowadzono w oparciu o następujący program temperaturowy: 240°C przez 1 min, ogrzewanie do 270°C z przyrostem temperatury 12°C/min. Tę temperaturę pieca utrzymywano przez 8 min. Temperatura dozownika i detektora wynosiła 290°C. Zastosowano podział strumienia 40:1, a objętość nastrzyku wynosiła 1 μl .
2. W metodzie NCM zastosowano chromatograf HP6890N GC-FID z kolumną kapilarną J&W HP 5 (długość 30 m, średnica wewnętrzna 0,25 mm, grubość filmu 0,25 μm). Rozdział chromatograficzny przeprowadzono w oparciu o następujący program temperaturowy: 250°C przez 4 min i ogrzewanie do temperatury 260°C z przyrostem 7°C/min, a następnie z przyrostem 6°C/min do 280°C. Temperatura dozownika i detektora wynosiła 280°C. Zastosowano podział strumienia 40,5:1, a objętość nastrzyku wynosiła 1 μl .

2.3. Walidacja metody

Temperatura dozownika i detektora, współczynnik podziału, szybkość przepływu gazu nośnego oraz program temperaturowy były optymalizowane przy zastosowaniu kolumn J&W HP Ultra 2 i J&W HP 5. Celem optymalizacji było uzyskanie dobrego rozdziału pików, których kształty w stosunkowo krótkim czasie były odpowiednio wyraźne. Szczegóły dotyczące wykonanego procesu optymalizacji opisano wcześniej [2]. Ponadto dla każdej z optymalizowanych metod chromatograficznych przeprowadzono walidację, sprawdzając pięć parametrów: powtarzalność, odtwarzalność, liniowość,

zakres liniowości i odzysk. Roztwór wzorcowy zawierał 0,05 mg/ml kodeiny (CD), 0,08 mg/ml morfiny (MP), 0,10 mg/ml acetylokodeiny (AC), 0,26 mg/ml 6-monoacetylmorfiny (MM) oraz 0,31 mg/ml heroiny (HR). Jako rozpuszczalnik zastosowano mieszaninę metanolu i chloroformu w stosunku 1:0. 2,2,2-trifenyloacetofenon użyto jako standard wewnętrzny (IS). Po optymalizacji wykonano oznaczenia dla PCM oraz NCM i zdecydowano się zastosować odpowiednio 0,18 mg/ml i 0,20 mg/ml IS. Roztwór o tej samej zawartości oznaczanych składników był stosowany w badaniach precyzji wyników uzyskiwanych w danym dniu pomiarowym, jak też w różnych dniach pomiarowych. Wykonywano po sześć nastrzyków. Liniowość określono na podstawie wyników uzyskanych dla serii wzorców o ośmiu poziomach stężeń, które sporządzono przez rozcieńczenie mieszaniny wyjściowej zawierającej pięć analitów. Zawartość standardu wewnętrznego była stała. Każdy z roztworów z serii wzorców był analizowany co najmniej dwa razy. W badaniach odzysku mieszaninę analitów nastrzykiwano na matryce.

2.4. Przygotowanie próbek

Analizowano 43 próbki heroiny. Odważano około 80–85 mg uśrednionej próbki i rozpuszczano w kolbie, a następnie dodawano do niej 10 ml standardu wewnętrznego. Każdy roztwór był analizowany dwukrotnie.

2.5. Analiza danych

Informacja o pięciu głównych alkaloidach opium oznaczanych w każdej z analizowanych próbek (kodeina, morfina, acetylokodeina, 6-monoacetylmorfina i heroina) była wyrażana na dwa sposoby, tj. jako stosunek pola powierzchni pików do pola powierzchni IS oraz jako ich zawartość (mg/ml). Stosowano program Minitab 15 (Stany Zjednoczone) do wielowymiarowej analizy danych oraz Microsoft Excel do wstępnej obróbki danych, tj. określenia zawartości poszczególnych związków w analizowanych próbkach oraz do procesu normalizacji uzyskanych danych. Przed wykonaniem analizy głównych składowych dane poddano standaryzacji. Pierwsza i druga główna składowa były używane w ocenie rzetelności stosowanych metod.

3. Rezultaty i dyskusja

3.1. Opis stosowanych metod

Zdolność rozdzielcza była oceniana przez porównanie chromatogramów danej próbki, które uzyskano za pomocą dwóch metod chromatograficznych stosowanych podczas badań. Stosując metodę NCM (rycina

1b), uzyskano niezbyt satysfakcjonujące kształty pików oznaczanych związków, m.in. wystąpiło istotne „ogonowanie” pików morfiny i heroiny. Miało to z kolei wpływ na ustalenie linii bazowej w procesie integracji pików, a tym samym przyczyniło się do błędów oznaczania poszczególnych składników próbki metodą NCM. Wskazane jest bowiem, aby piki widoczne na chromatogramach uzyskanych techniką GC były ostre (rycina 1a). Metoda NCM nie spełniała tych kryteriów.

Sprawność systemu oceniono, określając w procesie walidacji pięć parametrów. Powtarzalność zdefiniowano jako stabilność instrumentu w danym dniu pomiarowym, a odtwarzalność jako stabilność pracy aparatu w różnych dniach. Powtarzalność i odtwarzalność określano za pomocą względnego odchylenia standardowego (*RSD*) i obliczano na podstawie wyników pomiarów uzyskanych z sześciu powtórzeń. Jak wynika z tabeli I, odtwarzalność i powtarzalność obu metod można uznać za zadowalającą, ponieważ względne odchylenie standardowe było mniejsze niż 15%, co jest poziomem akceptowalnym przez Biuro Organizacji Narodów Zjednoczonych ds. Narkotyków i Przestępstw (United Nations Office of Drugs and Crime) [10].

W analizie ilościowej wskazane jest, aby stosowana procedura analityczna charakteryzowała się stosunkowo szerokim liniowym zakresem odpowiedzi detektora na zawartości analitu w próbce. Metoda PCM wykazała bardzo dobrą liniowość ($r^2 > 0,999$) w zakresie 50–150% analizowanych stężeń roztworów wzorcowych. Niemniej jednak zakres liniowości pracy w metodzie NCM był nieznacznie mniejszy niż w metodzie PCM. Nieliniowy zakres odpowiedzi z r^2 pomiędzy 0,96–0,98 był głównie widoczny w metodzie NCM. Tym samym trudno było ją uznać za dobrą metodę analityczną, ponieważ nie dało się precyzyjnie i dokładnie oznaczyć analitów, mimo iż znajdowały się w zakresie liniowym metody. Dla pewnych stężeń wzorców metoda NCM wykazuje bowiem zaniżone lub zawyżone wartości, prowadząc do względnie małej wartości r^2 .

Dokładność metody była mierzona jako jej odzysk. Metoda PCM wykazała lepsze wartości odzysku niż metoda NCM, która dawała wyższe wartości odzysku dla morfiny i 6-monoacetylmorfiny.

Inne aspekty procesu walidacyjnego dla metody PCM opisano szczegółowo w literaturze [2]. Na podstawie uzyskanych wyników oznaczeń można stwierdzić, że jest ona odpowiednia do profilowania heroiny, dlatego też może być zastosowana jako metoda odniesienia przy ocenie innych metod profilowania narkotyków. Wyniki oznaczeń uzyskane za pomocą metody NCM nie mogą być brane pod uwagę ze względu na wspomniany powyżej nieakceptowalny zakres liniowości oraz uzyskane wartości odzysku. Dlatego też metoda NCM była zastosowana w badaniach jako negatywna metoda odniesienia.

3.2. Ocena rzetelności metody z zastosowaniem PCA

Każda z 43 próbek była dwukrotnie analizowana za pomocą obu metod chromatograficznych. Wybrane próbki zawierały wszystkie z oznaczanych związków, co pozwoliło zminimalizować błąd statystyczny związany z sytuacją, w której w danych występuje zero, tj. brak pików. Dane ilościowe o analizowanych związkach uzyskano na podstawie wyników analiz GC dwoma sposobami. Pierwszy polegał na wyznaczeniu stosunku pola powierzchni pików danego związku do pola powierzchni standardu wewnętrznego. Procedura ta pozwalała minimalizować błędy związane z niestabilnością stosowanej metody. Drugi sposób opierał się na wyznaczeniu stężenia w mg/ml na podstawie wykonanej jednopunktowej kalibracji, w której określano w danym dniu pomiarowym stosunek pola powierzchni pików oznaczanego związku do pola powierzchni standardu wewnętrznego. Metoda ta, oprócz kompensacji błędów, które występują również w pierwszej z opisanych metod, umożliwia kompensację innych, nieraz trudnych do zidentyfikowania czynników, które mogą mieć wpływ na uzyskiwane dane. Innymi słowy, system był korygowany w momencie jego skalibrowania poprzez uwzględnienie informacji o standardzie.

W celu minimalizacji wpływu aparatury i błędów związanych z zastosowaniem metod statystycznych, wykonano normalizację danych uzyskanych dla pięciu związków będących głównymi alkaloidami heroiny. Ponadto heroina (diacetylmorfina) jest związkiem niestabilnym, tj. rozkłada się ona do 6-monoacetylmorfiny, a czasami dalej poprzez deacetylację do morfiny i to musi być również brane pod uwagę. Dlatego też, w celu kompensacji efektu rozkładu morfiny, dane były wstępnie normalizowane według zasad: $CD / (MM + HR)$, $MP / (MM + HR)$, $AC / (MM + HR)$, $CD / (MP + MM + HR)$, $AC / (MP + MM + HR)$ oraz $(CD + AC) / (MP + MM + HR)$. Suma zawartości morfiny jako licznik w powyższych wyrażeniach minimalizuje zmiany związane z procesem jej rozkładu. Oznacza to, że dane uzyskane za pomocą obu metod analitycznych nie mogą być porównywane bez wstępnej obróbki danych. W celu określenia relacji między próbkami wykonano analizę za pomocą metody PCA. Część zmienności tłumaczona przez główne składowe wyrażona została w procentach (%V). Informację o tym zamieszczono na wykresach rozrzutu danych.

Wykresy rozrzutu danych (rycina 2a i 2b) otrzymane dla danych, które zebrano obiema opisanymi wcześniej metodami, wyglądają jak pionowo obrócone względem siebie. Niemniej jednak podobieństwa pomiędzy próbkami określane miarą odległości między nimi są takie same, tj. niezależne od tego, czy dane uzyskano z wykorzystaniem procedury kalibracji, czy też bez niej. Na przykład

próbka „65” na obu wykresach jest położona blisko próbek „50” oraz „111” i znajduje się w dużej odległości od próbki „277”. Niemniej jednak na rycinie 2a i 2b widoczne są nieznaczne różnice w odległościach pomiędzy tymi samymi próbkami. Mogą być one wytłumaczone błędem przypadkowym, który jest niwelowany tylko wówczas, gdy dane pomiarowe koryguje się, wykorzystując informację o wynikach pomiaru roztworów standardowych w danym dniu pomiarowym. W takim przypadku rozkład danych na rycinie 2b jest przypuszczalnie bardziej wiarygodny niż ten na rycinie 2a.

Aby otrzymać wykresy rozrzutu danych w tej samej skali, należy wykonać standaryzację danych uzyskanych różnymi metodami. W tym celu wartość danej zmiennej, np. $CD / (MM + HR)$, należy podzielić przez jej odchylenie standardowe obliczone na podstawie wszystkich wyników uzyskanych dla stosunku $CD / (MM + HR)$. Kiedy dane po standaryzacji były poddane analizie PCA, to uzyskane nowe wykresy (rycina 2c i 2d) wykazują duże podobieństwo. W związku z tym rozrzut danych pomiarowych poddanych tylko normalizacji, jak również znormalizowanych i poddanych standaryzacji, pozwala uzyskać stały rozrzut danych pomiarowych niezależnie od tego, czy dane z analizy GC były, czy też nie, skorygowane przez zastosowanie kalibracji. Dlatego też metoda PCM jest nie tylko wolna od widocznych błędów analitycznych, ale jednocześnie ujawnione zależności podczas analizy wykresów rozrzutu danych można wykorzystywać w celu oceny metody NCM.

Zaobserwowano różnice na wykresach rozrzutu danych (rycina 3a i 3b) uzyskanych w przypadku, gdy analizę statystyczną wykonano, posługując się danymi znormalizowanymi otrzymanymi metodą NCM z wykorzystaniem, bądź też nie, procesu kalibracji. Również gdy dane poddane zostały dodatkowo procesowi standaryzacji, to wykresy rozrzutu danych uzyskane metodą PCA (rycina 3c i 3d) nie były zbyt podobne do siebie. Wykres rozrzutu danych uzyskany dla danych skalibrowanych, gdy kalibracje wykonywano codziennie, jest różny od uzyskanego dla danych nieskalibrowanych. Oznacza to, że błędy analityczne występujące w metodzie NCM są znaczne nawet wówczas, gdy uzyskane dane analityczne są korygowane przez wykonywanie codziennej kalibracji. Dowodzi to, że zależność między próbkami uwidoczniona w formie wykresów rozrzutu zależy w dużym stopniu od tego, jakie dane poddaje się analizie metodą PCA. Dlatego też metody analityczne charakteryzujące się słabą odtwarzalnością (np. metoda NCM) dostarczają wykresy rozrzutu danych różniące się znacznie w zależności od tego, jakie dane są stosowane w analizie PCA.

Porównując ryciny 2 i 3 uzyskuje się informację, w jakim stopniu różnią się dane otrzymane za pomocą porównywanych metod analitycznych. Dla danych standaryzowanych próbki, dla których obserwowano różnice, są oznaczone na rycinie 3c i 3d w formie białych

kropek. Lokalizacja tych próbek była następnie porównywana z ich pozycją na rycinie 2c i 2d. Ponadto ryciny 2c i 3c wykazują rozbieżność w rozrzucie danych w górnym prawym rogu. Punkty występujące na wykresach rozrzutu danych uzyskanych metodą NCM są względnie luźno rozłożone. Istotna różnica występuje w przypadku próbki „78”, która zmieniła swoje położenie ze skupienia w środkowej części wykresu na rycinie 2c do miejsca po prawej stronie na rycinie 3c. Położenie punktów na rycinie 3d nie jest zgodne z położeniem punktów na rycinie 2d, pomimo że błędy związane z metodą NCM były skorygowane przez wykonywanie codziennej kalibracji. Wyraźne różnice w położeniu można również zaobserwować dla próbek o numerach „50”, „61” i „65” oraz obszaru w dolnej lewej części wykresu rozrzutu danych. Na przykład skupienie zawierające próbki „16”, „37” i „39” znajduje się stosunkowo blisko skupienia „7” i „111”. Skupienia te są z kolei od siebie oddzielone na rycinie 3d. Biorąc pod uwagę, że wykres rozrzutu danych na rycinie 2d jest stosowany jako wzorzec odniesienia, to wykresy rozrzutu danych uzyskane dla wyników zebranych metodą NCM są generalnie zniekształcone. Tym samym jej zastosowanie do profilowania narkotyków może prowadzić do błędnej interpretacji zależności występujących pomiędzy porównywanymi próbkami.

4. Podsumowanie

Podsumowując należy stwierdzić, że użyteczna i zwalidowana metoda analityczna powinna dostarczać danych, które z kolei mogą być wykorzystane po zastosowaniu metod analizy statystycznej. Metoda PCM jest przykładem pozwalającym wykazać statystyczną rzetelność danych bez względu na to, jak są one przygotowywane do analizy statystycznej, tj. czy dane są lub też nie są skalibrowane. Wykazano, że metoda NCM jest mało rzetelną metodą analityczną oraz fakt, że błąd analityczny związany ze stosowaną metodą analityczną prowadzi do błędnej interpretacji dokonywanej na podstawie analizy wykresów rozrzutu danych. W związku z tym artykuł zachęca chemików sądowych do oceny rzetelności stosowanej w profilowaniu narkotyków procedury analitycznej oraz metody statystycznej w analizie uzyskanych danych.

Podziękowania

Autorzy pragną podziękować władzom Uniwersytetu Malezyjskiego za grant IPPP (PV004/2011B), Wydziałowi Chemii Uniwersytetu Malezyjskiego (MOSTI) za możliwość wykonywania badań z wykorzystaniem aparatury wydziałowej i w pomieszczeniach tego wydziału oraz firmie Y.S.P. Industries (M) Sdn. Bhd. za dostarczenie standardów wykorzystywanych w tym projekcie.