A quick and automated method for profiling heroin samples for tactical intelligence purposes

Valérie Dufey, Laurence Dujourdy, Fabrice Besacier*, Huguette Chaudron

Laboratoire de Police Scientifique de Lyon, 31 Avenue Franklin Roosevelt, 69134 Ecully, France

Received 25 July 2005; received in revised form 1 August 2006; accepted 3 August 2006
Available online 14 September 2006

Abstract
A computerized procedure is presented for the profiling and comparison of illicit heroin seizures. The system involves the derivation and gas chromatographic separation of five major constituents followed by a fully automatic data analysis and transfer to a PHP/MySQL database. Comparisons via the square cosine function between a single chromatographic profile and the whole database (several hundred of samples) are performed in a few seconds.

Advantages of this profiling method compared to the classical minor constituents one are discussed.
© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Heroin; Major alkaloid profiling; Neutral and acidic impurities; Drug intelligence; Diamorphine; Inter-sample correlation; PHP/MySQL database

1. Introduction
Comparison of heroin samples is not a new matter. Since the beginning of the 1980’s, many studies have been conducted to determine chemical links between heroin exhibits [1–4].

Most of the profiling methods for characterization of heroin derived from narcotine and norlaudanosine-related compounds have been developed by Allen et al. [5]. However, this technique presents several disadvantages, first it requires a rather high amount of material (15 mg of pure diamorphine, i.e. for heroin 10% one needs 150 mg of powder) critical for highly diluted heroin samples. Furthermore, it needs good professional skills to correctly perform the extraction step. Finally, it is a rather long preparation method due to centrifugation and evaporation.

Recently, Gueniat followed by Esseiva introduced an alternative profiling method dealing with the characterization of major constituents [6]. The main difference between the two methods is that almost no preparation is required for the one described here.

The Police Laboratory of Lyon has utilized the minor constituents method for more than 10 years on a routine basis. Comparisons between heroin exhibits were achieved via calculations of a Euclidian distance.

Since 1994, the average purity of street heroin samples has been decreasing down to 5–10% in 2004 (in 2004, 65% of the heroin samples analyzed by the five French police laboratories were below 10% purity [7], see Fig. 1).

Thus, it was more and more difficult to use the minor constituents method, especially for street samples, because this method needs 15 mg of pure diamorphine. Thus, if we have a small quantity of heroin at 4%, for instance, we have to weigh 375 mg of powder and sometimes it is impossible to obtain it.

To avoid this annoyance, in 2003, it was decided to modify the standard operating procedure for heroin comparisons with the following objectives:

- to provide a fast profiling method,
- to analyze nearly all small heroin exhibits whatever the purity,
- to automate the data analysis and transfer as well as the correlation calculations,
- to store the data in a dedicated database.

The major constituents method seemed to be more suitable to reach these goals. Therefore, it was decided to implement the analytical method used in Switzerland and to build a fully automated tool.
2. Material and methods

2.1. Minor constituents method

2.1.1. Sample preparation

About 1 g of heroin sample is homogenized in an agate mortar. An aliquot equivalent to 15 mg of pure diamorphine base is weighed and transferred into a 30 mL flask. The sample is dissolved in 5 mL of 0.5 M H₂SO₄ and then 5 mL of toluene are added. Vortex thoroughly and then centrifuge to separate the phases. Remove 4 mL of the organic phase, transfer into a vial and evaporate to dryness. Add 200 μL of chloroform and inject into the GC.

2.2. Major constituents method

2.2.1. Sample preparation

About 1 g of heroin sample is homogenized in an agate mortar. Fifteen milligrams of sample are weighed and transferred into a vial. The sample is dissolved in a chloroform/pyridine mixture (5:1) and then 100 μL of MSTFA is added. The sample is warmed for 1 h at 80 °C and injected into the GC.

2.3. Gas chromatography-flame ionization detection

Analyses of minor and major constituents were performed on an Agilent 6850 Gas Chromatograph equipped with a Flame Ionization Detector and an auto sampler. Sample volumes of 2 μL were injected with split mode.

The chromatographic separation was achieved on a J&W DB-1 capillary column (length 30 m, i.d. 0.25 mm, film thickness 0.25 μm). The helium carrier gas flow rate was 1.4 mL/min.

The oven temperature was programmed as follows:

- Minor constituents, 200 °C for 1 min, 4 °C/min to 280 °C, then 5 °C/min to 300 °C for 13 min.
- Major constituents, 150 °C for 1 min, 10 °C/min to 250 °C, then 4 °C/min to 300 °C for 4 min.

The injector and detector temperatures were set, respectively, to 280 and 300 °C for minor constituents, 290 and 325 °C for major ones.  

Typical chromatograms corresponding to both methods are shown in Figs. 2 and 3.

2.4. Data evaluation and storage

For minor constituents, similarity and dissimilarity among samples were evaluated by multivariate analysis using Euclidian distance, and square cosine function for major constituents. At that time, the aim for the minor constituents method was to discriminate samples between them. Distance calculations are more adapted to highlight dissimilarities, so Euclidian distance was chosen.

The aim for the major constituents method, in our perspective, was to show similarities between a new sample and candidates stored in a database. Angle measures as cosine function are more adapted to do this work [8].

Normalization was performed in both cases using the sum of peak areas for the minor constituents method, and the diamorphine peak area for the major constituents method.

A database for the major constituents method was built on an Apache web server, using PHP software and MySQL open source database.

For each heroin sample, a chromatographic profile is obtained thanks to the GC-FID and imported automatically by the operator from the gas chromatograph by selecting the appropriate data directory (see Appendix 1). Then, correlation scores are calculated between the sample and the whole database and a list of hits is determined (see Fig. 4).
3. Results and discussion

3.1. Method optimization

As the initial method was developed for Swiss samples, it was necessary first to check if French samples were the same as Swiss and if not to adapt the method to our samples.

Six target constituents were selected in the former method: meconin, papaverine, noscapine, acetylcodeine, 6-acetylmorphine and acetyltethebaol.

Unfortunately, with our rather impure samples, meconin was not present in most of the chromatograms (first difference between Swiss samples [9]). Increasing the amount of heroin for analysis did not solve the problem as the acetaminophen and caffeine peaks were too large and too close to the meconin peak.

A study of meconin influence on the discrimination power of the method was then performed. Nine samples, named A to I, which contained meconin at a sufficient level were utilized. They belonged to three different batches: first one comprised

![Typical chromatogram of major heroin constituents.](image)

![List of the correlation scores calculated between the sample and the whole database.](image)
samples A to F, second batch consisted of G and H, and last batch was I alone. All nine samples were analyzed and the impact of meconin on comparisons was measured by calculating the correlation value, via the square cosine function, with and without the meconin area. The correlation matrices are shown, respectively, in Tables 1 and 2.

Distribution of the correlation value calculated with five or six constituents is shown in Fig. 5.

The range of values vary from 0 to 100, 0 indicating a complete dissimilarity, 100 a complete similarity.

It was noticeable that meconin had no significant influence on the correlation values obtained from within batch samples and between batches.

Therefore, meconin was no longer considered in the profiling method and the oven program was changed in order to achieve a good chromatographic separation for the five remaining constituents only. The total run time was then lowered from 30 to 23 min.
Another difference with Swiss samples is that noscapine was often present in large amounts in our heroin exhibits (up to 48% [7]). Possibly it was added as a “cutting” agent [10], so the frequency distribution of noscapine/HMM ratio (HMM being the sum of morphine, monoacetylmorphine and diacetylmorphine concentrations) was plotted (see Fig. 6). The distribution is based on 670 samples issued from seizures realized between 2000 and 2004. The upper limit of this ratio is 3.2.

The figure shows that in most samples (85%) the noscapine/HMM ratio does not exceed the value of 1. This is in close agreement with the fact that heroin seized in France comes mainly from South West Asia [11], which contains morphine and noscapine in ratio from 2:1 to 1:2 [12]. Only few samples (around 2%) show a high noscapine/HMM ratio so we can conclude that noscapine is an original component of illicit heroin samples. Nevertheless, to avoid the influence of the largest constituents, especially noscapine, on the correlation value, it was decided to apply standardization (after normalization to diamorphine) to the peak areas.

3.2. Reproducibility and repeatability

Performance of the method was checked via reproducibility and repeatability studies.

First, several measurements were made by one operator, using the same equipment and over a relatively short time span. Two heroin samples, respectively, at 3 and 49% were injected 10 times each. Relative standard deviation of each constituent was calculated and profiles were compared by the square cosine function (Tables 3a and 3b).

The same samples were then analyzed by one operator, using the same equipment, over two weeks (one analysis per day). Relative standard deviation of each constituent was calculated and the profiles compared by the square cosine function (Tables 4a and 4b).

Furthermore one sample was analyzed using the same equipment, over an eight-month period. Box plots representation for each constituent and distribution of the correlation values for these samples are plotted in Figs. 7 and 8.

According to the results, the method seemed to be repeatable and reproducible: all the R.S.D. for each major component are under 8% and all the correlation values were higher than 99.90%.

3.2.1. Discriminatory power of the analytical method

We focused then on the study of linked and unlinked samples in order to determine a threshold to assert that two samples came from the same batch.

We first performed the profiling of heroin samples with similar profiles. We chose three heroin batches of linked samples, containing, respectively, 29, 16 and 10 exhibits (total number of pairs reached 571) followed by comparing unlinked samples. We supposed that a batch life did not exceed more than one year and to be sure to select different batches, we chose heroin seizures stored for more than two years and selected 30 samples seized in

<table>
<thead>
<tr>
<th>Heroin percentage</th>
<th>3%</th>
<th>49%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average correlation value</td>
<td>99.998</td>
<td>99.991</td>
</tr>
<tr>
<td>Minimum correlation value</td>
<td>99.991</td>
<td>99.937</td>
</tr>
</tbody>
</table>

Table 3a
Correlation values calculated from the repeatability study

<table>
<thead>
<tr>
<th>Heroin percentage</th>
<th>3%</th>
<th>49%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average correlation value</td>
<td>99.999</td>
<td>99.996</td>
</tr>
<tr>
<td>Minimum correlation value</td>
<td>99.998</td>
<td>99.974</td>
</tr>
</tbody>
</table>

Table 4a
Correlation values calculated from the reproducibility study

<table>
<thead>
<tr>
<th>Reproducibility study</th>
<th>Acetylcodeine (%)</th>
<th>Acetylthebaol (%)</th>
<th>6-MAM (%)</th>
<th>Papaverine (%)</th>
<th>Noscapine (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.S.D. for the 3% heroin sample</td>
<td>0.6</td>
<td>2.4</td>
<td>1.6</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td>R.S.D. for the 49% heroin sample</td>
<td>0.6</td>
<td>1.1</td>
<td>4.0</td>
<td>2.2</td>
<td>1.5</td>
</tr>
</tbody>
</table>
2001, 4 in 2002, 70 in 2003, 72 in 2004 and compared them to each other (total number of pairs reached 10,500).

The distribution of the correlation values for linked samples and unlinked ones are plotted, respectively, in Figs. 9 and 10. For unlinked samples, the range of correlation values was large varying from 6.86 to 99.88%.

Conversely, 97.7% of the values corresponding to the linked samples were higher than 99.92%. The comparison of these two distributions allowed us to ascertain the capacity of the method to determine chemical links between seizures.

The number of false positives and false negatives have also been determined during this study. We plotted the recovering of curves representing the respective frequency of appearance of the correlation values within and between batch samples (see Fig. 11). Bar diagram was smoothed using a curve modeling software (Statistica®).

The intersection of these two curves (99.91) and the area covered gave the percentage of false positives and negatives, respectively, 0.2 and 6%.

We then chose arbitrarily to set two thresholds which minimize the risk of false positive and negative and correspond to two decision levels:

- similarity (99.92) which gave a percentage of false positives of 0.003.
- dissimilarity (99.85) which gave a percentage of false negatives of 0.2.

3.2.2. Adulteration impact

Influence of adulteration in heroin comparison was conducted. In France, 90% of our base heroin seizures are “cut” with the mixture caffeine/acetaminophen with the following proportions: 45/55 [7]. Therefore, we investigated the influence of these products on our comparison tool.

We cut a 49% pure base heroin sample in order to give a range of purity between 5 and 49%. The different sets of samples were then compared to each other and correlations were obtained. Results are summarized in Table 5.
This study shows that the adulteration did not have any significant influence on the correlation values.

A second study was then conducted to assess up to what percentage this method remained reliable. A 13% pure base heroin sample was cut several times with the same mixture of caffeine/acetaminophen until a purity of 1.5%. In the same way, correlations between the different sets of samples were calculated and minimum, maximum and mean are listed in Table 5.

First, we noticed that below 3%, the acetylthebaol peak did not appear any more in the chromatogram. Moreover, the correlation values calculated between the 1.5% heroin chromatogram and the other ones were below the dissimilarity threshold (C mean is 99.78 compared to 99.92). Finally, we noticed that replicates were not reproducible with samples whose purity was below 2%.

Therefore, we decided to use that profiling method only for heroin samples with a purity level above 3% to be sure to integrate acetylthebaol and to have reproducibility between replicates.

Table 5

<table>
<thead>
<tr>
<th>Heroin percentage</th>
<th>Between 49 and 25%</th>
<th>10%</th>
<th>5%</th>
<th>Between 13 and 6%</th>
<th>3%</th>
<th>1.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>99.98</td>
<td>99.96</td>
<td>99.94</td>
<td>99.98</td>
<td>99.94</td>
<td>99.78</td>
</tr>
</tbody>
</table>

3.2.3. Comparison between the characterizations of the major and minor constituents

The last study consisted of comparing results from the former and the new method by analyzing 30 heroin samples. Results showed that 95% of the links performed by the minor constituents method were confirmed by the major constituents one. This is in agreement with the conclusion of Esseiva et al. [6]. Fig. 12 explains the results in details.

The Y left scale represents Euclidian distances obtained with the minor constituents method (graph in black).

The Y right scale represents cosine correlations obtained with the major constituents method (graph in grey).

The X scale represents the various samples analyzed.

The linked samples would give a score below 0.05 for the former method and above 99.92 for the new method (zone in white).

The unlinked samples would give a score above 0.1 for the former method and below 99.85 for the new method (zone in dark grey).

The zone in light grey corresponds to an uncertainty area. In this case, the analyst takes the decision by examining the different factors concerning the seizure.

In this figure, we can see that the links obtained with the former method are confirmed by the new method. Moreover
with the major constituents method, the samples in the uncertainty area move, in most cases, to a decision zone.

However, the study was only done on 30 samples and it would be better to confirm these results on a larger set of samples.

4. Conclusion

In conclusion, this method, based on the characterization of major constituents, proved to be efficient and reliable. Profiling of a heroin exhibit can now be achieved in one hour, whereas the former method, based on the characterization of minor constituents, needed at least 4 h.

Moreover, the major constituents method showed good discriminatory capabilities with a low percentage of false negatives and false positives. On the other hand, as the chemical link is determined using five parameters only, it is recommended that other data such as heroin percentage or information given by the investigation unit is also taken into account.

Last but not least, the major constituents method is not only an analytical tool but, because it also comprises an innovatory computerized process, allowing the laboratory to easily transfer the profiling data from the instrument to the database and to calculate correlation values with the already stored profiles.

Acknowledgements

We would like to thank Didier Loyen for his technical assistance in the development of the database and Ania Parenty for English corrections.

Appendix 1. Heroin profiling database

3 choices are available for heroin:

- automatic import,
- manual import,
- consultation.
Appendix 1. (Continued)

With automatic import, as the full pathway is already entered in the software, the data directory in the GC hard disk is directly showed. The user can therefore select the week corresponding to his analyses and then the data file (see below):

First the user clicks on the sample file, then he selects the file « report01.CSV »:

Peak area values are transferred automatically and the user just enters the case number and sample / seizure references (see below):
Appendix 1. (Continued)

With manual import, the user has to enter the data himself. It is possible to differentiate between a Quality Control sample and a case sample (by clicking yes or no in the corresponding area).

Impurity area values (acetylcodeine, acetylthebaol, mono-acetylmorphine, diacetylmorphine, papaverine, noscapine) can be checked visually.

Finally, the user clicks on the button: “insertion” to insert this sample in the database.

Correlation calculation:
The correlation value between two chromatograms is given by:

\[
C = 100 \cos^2 \theta
\]

\[
= 100 \left[ \frac{(A_1B_1 + A_2B_2 + \cdots + A_nB_n)^2}{A_1^2 + A_2^2 + \cdots + A_n^2} \right.
\]

\[
\left. + (B_1^2 + B_2^2 + \cdots + B_n^2) \right]
\]

where \( A_1, A_2, \ldots, A_n \) represent the respective values of the variables 1–\( n \) for chromatogram a, and \( B_1, B_2, \ldots, B_n \) represent the respective values of the variables 1–\( n \) for the chromatogram b.

Two thresholds have been determined:

- \( C < 99.85 \): the samples are not linked,
- \( C > 99.92 \): the samples are linked (belong to the same batch).

References