

Chemometrics and chromatographic fingerprints to discriminate and classify counterfeit medicines containing PDE-5 inhibitors

E. Deconinck^{1*}, P.Y. Sacré^{1,2}, P. Courseille¹, J. O. De Beer¹

¹ Division of food, medicines and consumer safety, Section Medicinal Products, Scientific Institute of Public Health (WIV-ISP), J. Wytmansstraat 14, B-1050 Brussels, Belgium

² Analytical Chemistry Laboratory, CIRM, University of Liège, Avenue de l'Hôpital 1, B36, B-4000 Liège, Belgium

Abstract:

Chromatographic fingerprints recorded for a set of genuine and counterfeit samples of Viagra® and Cialis® were evaluated for their use in the detection and classification of counterfeit samples of these groups of medicines. Therefore several exploratory chemometric techniques were applied to reveal structures in the data sets as well as differences among the samples. The focus was on the differentiation between genuine and counterfeit samples and on the differences between the samples of the different classes of counterfeits as defined by the Dutch National Institute for Public Health and the Environment (RIVM). In a second part

the revealed differences between the samples were modelled to obtain a predictive model for both the differentiation between genuine and counterfeit samples as well as the classification of the counterfeit samples.

The exploratory analysis clearly revealed differences in the data for the genuine and the counterfeit samples and with projection pursuit and hierarchical clustering differences among

the different groups of counterfeits could be revealed, especially for the Viagra® data set.

or both data sets predictive models were obtained with 100% correct classification rates for the differentiation between genuine and counterfeit medicines and high correct classification rates for the classification in the different classes of counterfeit medicines. For both data sets the best performing models were obtained with Least Square-Support Vector Machines (LS- SVM) and Soft Independent Modelling by Class Analogy (SIMCA).

*corresponding author: Eric.Deconinck@wiv-ispl.be

Tel. +32 2 642 51 36

Fax. +32 2 642 53 27

1. Introduction

Counterfeit medicines are a growing problem both in developing and industrialized countries. The problem of counterfeit medicines in the industrialized countries, like the European Union, is mostly due to the extension of the internet. Research revealed that about one of five Belgians bought already at least once medicines through internet, from which is known that about 50% of the products, sold by sites disclosing their identity, are counterfeit. It is estimated that about 1% of the European market in medicines is covered by counterfeits [1]. In most cases the concerned products are produced without respecting the GMP quality norms and imported illegally. Health risks, associated with these products, can be due to the absence of the active ingredient, the presence of a wrong active substance, a wrong dosage, possible high concentrations of toxic impurities and a combination of (non-approved) active substances.

The World Health Organisation (WHO) first cited the counterfeiting of medicines in 1985 at the conference of experts on the rational use of drugs in Nairobi. In 1988, a World Health Assembly Resolution (41.16) recommended to “initiate programs for the prevention and detection of export, import and smuggling of falsely labeled, spurious, counterfeited or substandard pharmaceutical preparations” [2]. This led to the launch of many international initiatives among which the International Medicinal Products Anti-Counterfeiting Taskforce (IMPACT) started by the WHO in 2006. In parallel, the major pharmaceutical companies established the Pharmaceutical Security Institute (PSI) in 2002. On the European level the European Alliance for Access to Safe Medicines (EAASM) was created [1]. This is a pan- European patient safety initiative committed to promoting the exclusion of counterfeit and substandard medicines from the supply chain. The European Parliament and the Council of Europe created recently an amendment to the European directive 2001/83/EC [3] on the community code relating to medicinal products for human use, describing the policy of the European Union towards counterfeit and substandard medicines. On the 26th of October of 2011 Europe launched the Medicrime convention which was held in Moscow, Russia. Medicrime is the first international instrument for the criminalization of counterfeiting of medical products and similar crimes in order to protect public health [4].

The World Health Organization (WHO) [5] defines a counterfeit drug as: “one which is deliberately and fraudulently mislabeled with respect to identity and/or source. Counterfeiting can apply to both branded and generic products and counterfeit products may include products with the correct ingredients or with the wrong ingredients, without the active ingredients, with insufficient active ingredient or with fake packaging.”

Even if this is the internationally accepted definition of a counterfeit medicine, it does not apply to the majority of the illegal products encountered on the European market, since they do not copy the packaging and brand names of the genuine products. Therefore the classification proposed by the Dutch National Institute for Public Health and

the Environment (RIVM) was followed [6]. This classification (table 1) distinguishes counterfeits, which appearance corresponds to the one of the genuine products, and imitations, which do not. Most of these imitations originate from Asia, where European and American patents are not recognized.

Table 1: Definition of the RIVM classes [5]

| Main category | Subcategory | Inclusion and exclusion criteria |
|---------------|------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| Counterfeit | Professional | Appearance in conformity with genuine medicine; Content of correct API within 90 - 110 % of declared value; No other APIs; not genuine medicine. |
| | Non-professional | Appearance in conformity with genuine medicine; Content of correct API outside 90 - 110 % of declared value; No other APIs. |
| | Mixed | Appearance in conformity with genuine medicine; Contains correct API and another, known API |
| | Fraudulent | Appearance in conformity with genuine medicine; Contains a different, known API. |
| | Analog | Appearance in conformity with genuine medicine, Contains other, unapproved API |
| | Placebo | Appearance in conformity with genuine medicine; Does not contain APIs. |
| Imitation | Professional | Appearance not in conformity with genuine medicine; Content of correct API within 90 - 110 % of declared value; No other APIs. |
| | Non-professional | Appearance not in conformity with genuine medicine; Content of declared API outside 90 - 110 % of declared value; No other APIs. |
| | Mixed | Appearance not in conformity with genuine medicine; Contains declared API and another API. |
| | Fraudulent | Appearance not in conformity with genuine medicine; Contains an undeclared API. |
| | Analog | Appearance not in conformity with genuine medicine; Contains other, unapproved API |
| | Placebo | Appearance not in conformity with genuine medicine; Does not contain APIs. |

The fight against the counterfeiting of medicines resulted in numerous articles where several analytical techniques have already been used for the detection of counterfeit medicines. These techniques are separated in two main groups: chromatographic and spectroscopic techniques. The chromatographic techniques are used for the separation, identification and quantification of the active substances. They may also contribute to the elucidation of the structure of new analogues. Commonly used chromatographic techniques contain cheap and easy ones such as thin layer chromatography (TLC) [7,8] but also more sophisticated and expensive ones: such as High Pressure Liquid Chromatography with Ultraviolet detection (HPLC-UV) [9-11], liquid chromatography with mass spectrometric detection (LC-MS) [9-22] and LC-diode array detection (DAD)-circular dichroism [20].

The spectroscopic techniques are often preferred to chromatography for the identification of counterfeit drugs because of the fact that they are fast, need less (or no) sample preparation and some of them are non destructive. Fourier-transformed Infrared spectroscopy (FT-IR) [18,19,23], near infrared (NIR) [7,23], Raman spectroscopy [23-25], X-ray diffraction (XRD) [26], colorimetry [27,28] and Nuclear Magnetic Resonance (NMR) [25,29] have demonstrated their usefulness to detect counterfeit or adulterated drugs.

All the papers and research mentioned above are focused on the identification and/or the quantification of the active substances in the samples. The disadvantage of this is that a counterfeit sample can be evaluated as quite save based on the active substance present and the dosage, but in fact can contain some toxic secondary components like impurities, plant based materials, residual solvents, etc. Therefore the approach of using chromatographic fingerprints, a technique widely used in the field of pharmacognosy for quality control of plants [30-35], is interesting.

In the frame of counterfeit drug detection, chromatographic impurity fingerprints is interesting, since a major hazard of counterfeit products is the presence of toxic impurities in unknown amount. Thus, as it is the case for herbal medicines, the fingerprint approach allows the discrimination of tablets according to their chromatographic profiles without knowing *a priori* the identity nor the quantity of the constituents.

In previous work at our section [36], the use of chromatographic fingerprints was evaluated for its ability to discriminate between counterfeit and genuine medicines. It was clearly shown that chromatographic fingerprints allow the distinction between genuine and counterfeit medicines and can be useful in the fight against counterfeit medicines. Apart from the fact that almost every laboratory is equipped with a HPLC instrument, an important advantage of this approach is that it allows to discriminate, to quantify the active ingredients and some known impurities and that it gives an idea about the impurity profile of the sample which can be compared with the impurity content of the genuine products.

In this paper several chemometric techniques, both explorative as classification methods, were explored and compared for their performance in the discrimination between genuine and counterfeit medicines and the classification of counterfeit medicines following the RIVM classification, based on the chromatographic fingerprints obtained by P.-Y. Sacré et al. [36].

2. Theory

2.1. Exploratory methods

2.1.1. Principal Component Analysis (PCA)

PCA is a projection method that allows projecting high dimensional data into a low dimensional space of new latent variables, called principal components.

The general principle of PCA is to reduce the number of variables by calculating linear combinations of these variables, in which the weights or loadings of the different variables are chosen in such way that the first principal component represents the highest variance in the data set and the higher principal components the highest remaining variance. By definition the different principal components are orthogonal. The loadings of the variables show their respective contribution to a given PC and the correlation between the different explanatory variables. The projections of the objects on the PCs are called the scores and give information about the similarities among the objects [37].

2.1.2. Discriminant Partial Least Squares (D-PLS)

Partial Least Squares is a supervised technique, based on exactly the same principles as PCA. The difference is situated in the definition of the latent variables, called PLS-factors. The PLS-factors, also linear combinations of the original explanatory variables in the data set are defined in such way that they maximize the co-variance with the response variables (categorical for D-PLS). In this way latent variables are obtained that are more directly related to the response variables [37].

2.1.3. Projection Pursuit (PP)

PP is also a projection technique that is able to project high dimensional data into a low dimensional space, defined by a few latent variables, called projection pursuit features (PPFs). Contrary to PCA, PPFs are obtained by maximizing a projection index describing the homogeneity of the data. In such way PP can be complementary to PCA in revealing the data structure and groups of similar samples that could not be observed with PCA. In this study the algorithm as described by Croux and Ruiz-Gazen was used [38]. Three projection indices were tested: entropy, kurtosis and yenyukov.

2.1.4. Hierarchical Clustering

In order to cluster the samples according to their similarity different types of hierarchical clustering were performed on both data sets. In fact the objects are clustered in a hierarchical and agglomerative way. Hierarchical means that smaller clusters are included in larger ones or vice versa, and agglomerative means that objects are sequentially merged [39]. The goal of the method is to classify m objects in $m-1$ steps. In each consecutive step, the two most similar objects (clusters) are merged. The objects (clusters) to merge are derived from the dissimilarity matrix, representing the dissimilarity between each pair of objects (clusters). Dissimilarities or dissimilarity coefficients are positive numbers that are small if two objects are closely related and large if they differ [37, 39]. In this study different coefficients based on the distances between objects in the data space are applied and compared for their performance in clustering the different samples in the data set (see results).

2.2. Modeling methods

2.2.1. K-Nearest Neighbors (k-NN)

The k-nearest neighbor method [37] (k-NN) is a classification technique where neighborhoods of training set objects are used for the construction of classification rules. During the classification procedure, the Euclidian distances between an unknown object and each of the objects of the training set are computed. For a dataset with n samples, n distances are calculated. Then, for a new object its k closest neighbors from the training set are examined. The unknown object is classified into the group to which the majority of the k neighboring objects of the training samples belongs.

2.2.2. Classification And Regression Trees (CART)

CART is a non-parametric statistical technique, developed by Breiman et al. [40] in 1984, which is able to solve classification (categorical dependent variables) as well as regression problems (continuous dependent variables). In both cases the method builds a decision tree, describing a response variable as a function of different explanatory variables.

A CART analysis generally consists of three steps. In a first step the maximum tree is build, using a binary split-procedure. To choose the most appropriate variable and split value, CART uses an algorithm in which all descriptors and all possible split values are considered. The split resulting in the highest decrease in impurity between the mother group and the daughter groups is selected. For classification trees the impurity can be defined by different split criteria [40]. The three commonly used split criteria are the Gini index, the Twoing index and the Information index.

The maximum tree is overgrown and closely describes the training set, usually resulting in overfitting. In a second step this overfitted model is pruned. This procedure results in a series of less complex trees, derived from the maximum tree. In the third and final step the optimal tree is selected using a cross validation procedure. For more details we refer to references [40-42].

2.2.3. Soft Independent Modeling by Class Analogy (SIMCA)

SIMCA is a classification technique that models each class of a dataset separately using PCA. The optimal number of principal components, required to describe a training class is evaluated using a cross-validation procedure. To construct classification rules, two critical values are considered obtained for the Euclidian distances towards the SIMCA model (the so-called orthogonal distances) and the Mahalanobis distances computed in the space of scores. The two critical values define a limited space around the samples of the training set in the model space with respect to their orthogonal distance and Mahalanobis distance. The position of a new object in the studied space is computed using the scores and loadings of the constructed PCA model. If the object is located within the defined limited space of orthogonal and Mahalanobis distances for the training class then it is said to belong to this class. Otherwise, the object is considered as an outlier, i.e. not belonging to this group. Confidence limits were set at 95%. Since SIMCA belongs to the so-called soft classification methods, it is possible that a new sample can be assigned to one or more existing groups or to any. This is a direct consequence of building disjoint classification models for each group of samples [37].

2.2.4. Support Vector Machines (SVM)

SVMs were originally designed to solve binary classification problems. Their basic concept is based on the calculation of hyperplanes defining decision boundaries to separate data points belonging to different classes. SVMs are able to deal with simple linear but also with more complex non linear classification problems. The idea behind SVMs is to map the original data points from the input data to a high dimensional feature space such that the classification problem becomes easier to solve in the feature space. The mapping is done by a suitable choice of a Kernel function [43]. The Kernel function transforms the data to a high dimensional space where the classification problem can be solved. Two types of Kernel functions were tested in this study: a Kernel function based on the standard inner product and a Radial Basis Function (Gaussian) Kernel function. More details about the mathematics behind SVM, the Kernel functions and the algorithms used in this study can be found in references [43, 44].

3. Methods and materials

3.1. Data

The data for the Viagra® like samples consists of the chromatographic fingerprints for 48 samples of counterfeit and imitations of Viagra® and 9 genuine samples. For the Cialis® like samples the data consists of the chromatographic fingerprints obtained for 44 samples of counterfeit and imitations of Cialis® and 6 genuine samples. The chromatographic fingerprints were recorded using the method described in Pharmeuropa for sildenafil (Viagra®) [45] and a slightly adapted method from Pharmeuropa for tadalafil (Cialis®) [46]. For more details about how this data was acquired we refer to Sacré et al. [36]. During the study of Sacré et al. it was seen that the variability in the chromatographic fingerprints for

the genuine samples is very low. The limited number of genuine samples should therefore be enough to represent the genuine class in the models.

All counterfeit and imitation samples were donated by the Federal Agency for Medicines and Health Products in Belgium (AFMPS/FAGG). One batch of each dosage of genuine Viagra® (25 mg, 50mg and 100mg) was kindly provided by Pfizer SA/NV (Belgium). Eli Lilly SA/NV (Benelux) kindly provided one batch of each commercial packaging (10 mg and 20 mg) of genuine Cialis®. Two other batches of each dosage of the genuine products were purchased from local pharmacies in Belgium.

The counterfeit samples were classified following the classification proposed by the RIVM [6]. The classification for both data sets is given in table 2.

Table 2: Composition of the data sets in function of the RIVM classes [6]

| Main category | subcategory | category number | Number of Viagra® like samples | Number of Cialis® like samples |
|---------------|------------------|-----------------|--------------------------------|--------------------------------|
| Counterfeit | Professional | 1 | 0 | 0 |
| | Non-professional | 2 | 2 | 0 |
| | Mixed | 3 | 1 | 5 |
| | Fraudulent | 4 | 0 | 0 |
| | Analog | 5 | 0 | 0 |
| | Placebo | 6 | 0 | 0 |
| Imitation | Professional | 7 | 34 | 28 |
| | Non-professional | 8 | 11 | 4 |
| | Mixed | 9 | 1 | 6 |
| | Fraudulent | 10 | 0 | 1 |
| | Analog | 11 | 0 | 0 |
| | Placebo | 12 | 0 | 0 |
| Genuine | | 0 | 9 | 6 |

3.2. Data pre-processing

In general the pre processing of chromatographic data consists of an improvement of the signal-to-noise ratio, normalization and peak alignment [47]. For the data used in this study a simple peak correction was sufficient to obtain a good overall alignment of all signals.

For more details about the data pre-processing and the major steps of the applied alignment method we refer to Sacré et al. [36] and Daszykowski et al. [48].

In order to evaluate the predictive properties of the models both data sets were divided in training and test set. It was chosen that the test sets would contain about 20 % of the samples. Two selection algorithms i.e. Duplexx [49] and Kennard and Stone [50] were applied to both data sets. For the Viagra® data set the test set (14 samples) selected by the Duplexx algorithm was considered as more representative for the complete data set than the one selected with Kennard and Stone. The main reason was that no genuine samples were selected for the test set with Kennard and Stone. For the Cialis® data set Kennard and Stone selected the best test set (12 samples), based on the coverage of the data space.

From the PCA plots shown in figure 1 it can be seen that the selected test sets cover quite well the data space of the data sets.

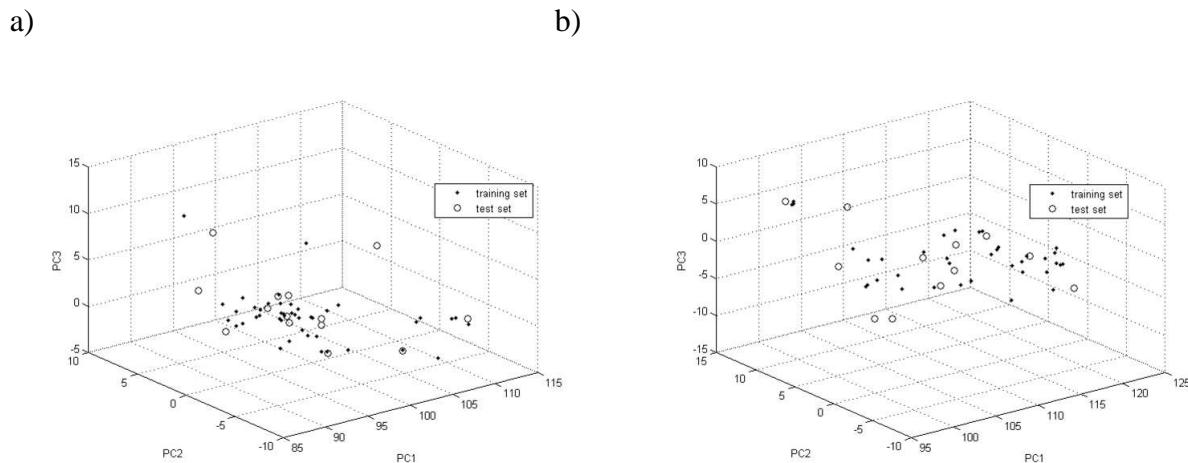


Figure 1: PCA plot representing the spread of the test set for the Viagra® data set (a) and the Cialis® data set (b) over the data space.

3.3. Software

Data processing and modeling was performed using Matlab version 7.9.0 (The Mathworks, Natick, MA, USA). The algorithms for PCA, D-PLS, PP, k-NN and hierarchical clustering were part of the ChemoAC toolbox (Freeware, ChemoAC Consortium, Brussels, Belgium, version 4.0). The programming of the CART algorithm was done according to the original CART algorithm proposed by Breiman [40]. The SIMCA toolbox was downloaded from the Matlab Central (www.mathworks.com/matlabcentral/fileexchange/30762-soft-independent-modeling-of-class-analogy-simca). SVM analysis was performed using the LS-SVM Toolbox version 1.8 (www.esat.kuleuven.be/sista/lssvmlab).

4. Results

Before data treatment the chromatograms were aligned using a simple linear interpolation technique [36,48], autoscaled and log transformed. The log transformation was necessary to reduce the influence of noise and small baseline shifts. The absorbances at the different time points were used as explanatory variables and the RIVM class numbers of the samples (table 2) were used as response variables (for the supervised techniques).

4.1. Exploratory data analysis

4.1.1. PCA

Viagra®-like samples

When performing a simple PCA analysis on the data set a clear differentiation could be observed between genuine and counterfeit samples. Figure 2 shows the score plot for PC1 and PC2. Investigation of the loadings on PC2 did not reveal a clear section of the chromatograms to be responsible for this discrimination. Even though the separation between counterfeit and genuine is clear, the counterfeit and imitation samples are all clustered together. No pattern could be observed.

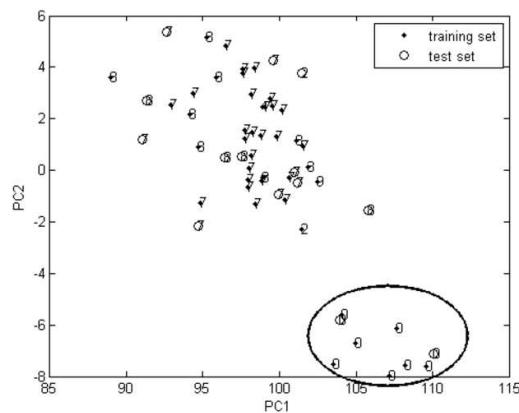


Figure 2: scoreplot for PC1-PC2 obtained for the Viagra® data set. The samples are identified by their class. The circle defines the genuine class.

Cialis®-like samples

For this data set no differentiation or clustering could be observed with PCA. The genuine samples are clustered together with some professional imitations (class 7). This is probably due to the fact that the counterfeit samples in the Cialis® data set are more similar to the genuines than is the case in the Viagra® data set. The samples in the Cialis® data set have more or less the same color and dimensions what is not the case for the samples in the Viagra® data set (see table 2). Also within the group of the counterfeits no clear clustering or differentiation in groups could be observed.

4.1.2. D-PLS

Since no clear clustering could be observed with PCA, it was decided to apply D-PLS as supervised technique in order to enhance discrimination of the samples in their respective classes.

Viagra®-like samples

The application of D-PLS did not result in better separation of the different groups of samples compared to PCA. The discrimination between genuine and counterfeits is more pronounced as can be expected with a supervised technique, but the counterfeit samples are still together and do not show any clustering or pattern (figure 3a). Also here the loadings do not reveal certain regions of the chromatograms responsible for the difference between genuine and counterfeit samples.

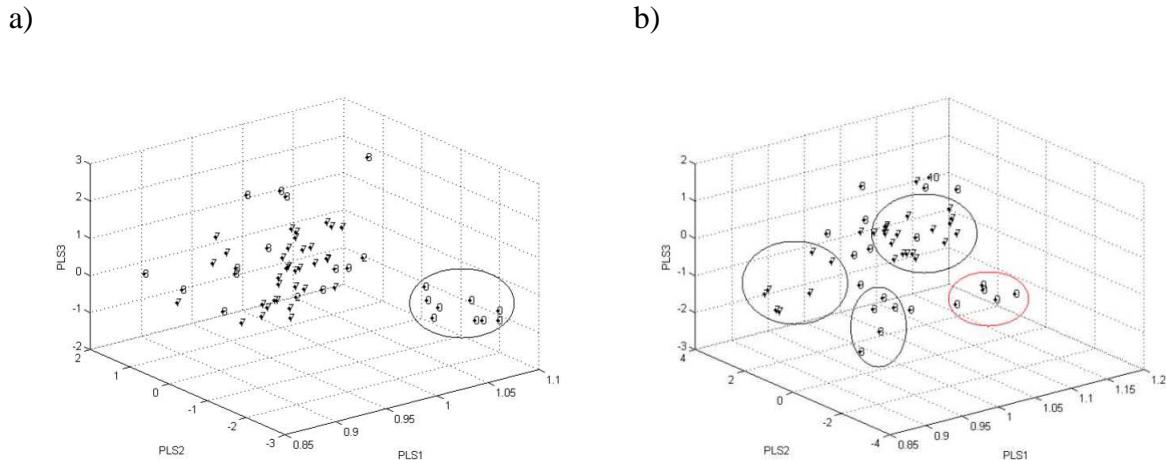


Figure 3: (a) scoreplot for PLS1-PLS2-PLS3 obtained for the Viagra® data set. The samples are identified by their class. The circle defines the genuine class. (b) scoreplot for PLS1-PLS2- PLS3 obtained for the Cialis® data set. The samples are identified by their class. The black circles define clusters of counterfeit samples. The red circle defines the genuine class.

Cialis®-like samples

On the score plot obtained for this data set (figure 3b) a clear differentiation between counterfeit and genuine samples can be observed, as well as a tendency of clustering within the group of the counterfeit samples. The separation between genuine and counterfeit is mainly defined along PLS2, but based on the loadings no determining regions of the chromatogram could be selected. Within the group of the counterfeit samples, three slightly separated groups could be observed. One big group containing professional imitations (class 7) and one non-professional imitation (class 8), a small group containing only professional imitations (class 7) and one containing only non professional imitations (class 8). These clusters, even if they are only slightly separated indicate that there is difference in data

between the samples of at least the classes 7 and 8 and that it should be possible to model these differences using classification algorithms.

4.1.3. PP

Projection Pursuit was applied using each of the three considered projection indices: entropy, kurtosis and Yenyukov.

Viagra®-like samples

For the Viagra® data set the application of PP with the entropy and the kurtosis index to be optimized, did not reveal any structure in the data. No differentiation between genuine and counterfeit samples could be observed and no significant clustering among the counterfeit samples could be observed.

The Yenyukov index is based on a nearest neighborhood approach, which says that the clustering tendency of data can be judged based on the ratio of the mean of all inter object distances and the average nearest neighbor distance. The use of this index resulted in a clear differentiation between counterfeit and genuine samples as well as in a tendency of clustering between the two main groups in the data set (group 7 and 8). Figure 4 shows the 3- dimensional plot of the projection of the different samples in the PP-space, defined by the three first projection pursuit factors. In comparison with PCA and PLS, PP revealed a certain structure in the data for the different groups of counterfeit samples. This indicates that these differences could possibly be modeled to obtain a predictive classification model.

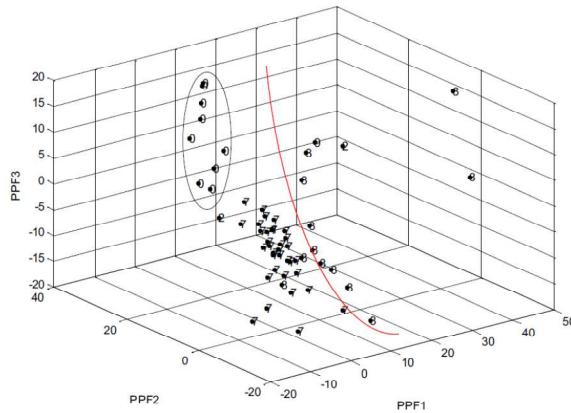


Figure 4: scoreplot for PPF1-PPF2-PPF3 obtained for the Viagra® data set. The samples are identified by their class. The circle defines the genuine class and the arc the separation between the samples of class 7 and 8.

Cialis®-like samples

The application of PP to the Cialis® data set, using the three selected indices, did not result in a clear structure of the data. No differentiation between genuine and counterfeit samples could be observed and within the counterfeit samples no clusters are present. Only when using the Yenyukov index a tendency of clustering of the genuine samples could be observed, but the structure is clearly worse than the one revealed by D-PLS.

4.1.4. Hierarchical Clustering

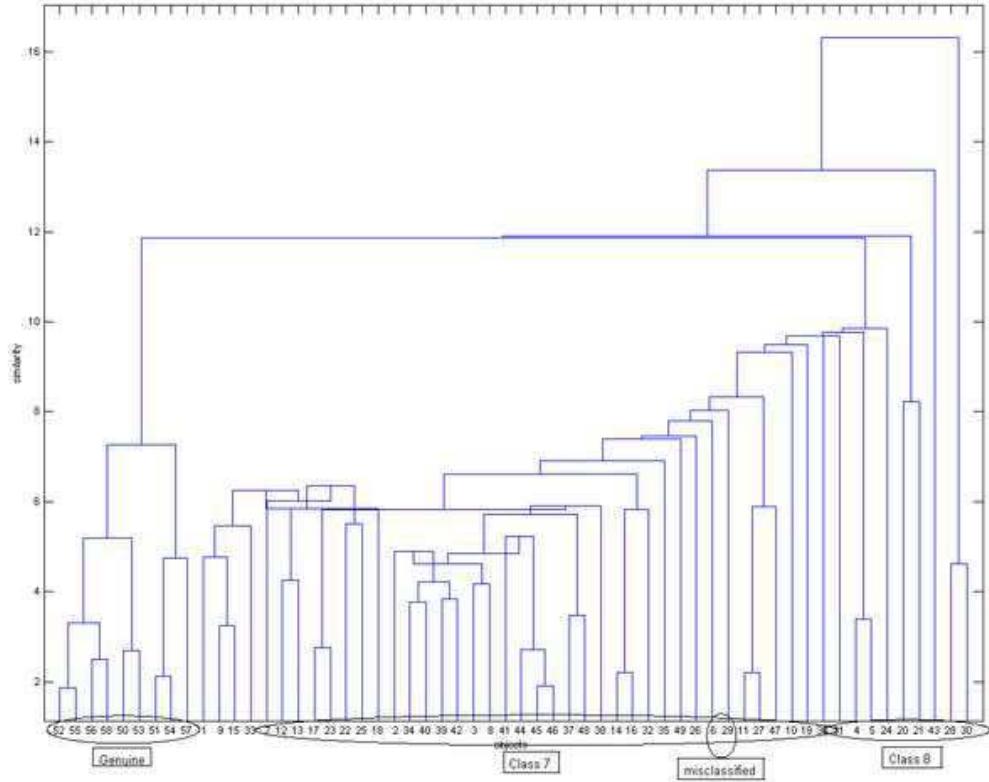
Five different hierarchical clustering techniques, i.e. single linkage, complete linkage, average linkage, centroid linkage and the Ward's method, were applied to both data sets. Each of the methods was tested using five similarity measures: the Euclidean distance, the standardized Euclidean distance, the Manhattan, the Mahalanobis and the Minkowski distance.

For both data sets the obtained clustering, with the different method-similarity measures combinations, was compared based on the clustering obtained for the differentiation genuine/counterfeit and for the clustering among the different classes of counterfeit samples.

Viagra®-like samples

The best results were obtained using the centroid linkage method with the Euclidean distance as similarity measure. Figure 5 shows the dendrogram (figure 5a) as well as the color map (figure 5b), obtained after calculation of the similarity matrix. When analyzing the dendrogram it can be seen that from left to right the genuine samples are clustered together, followed by the samples of the less represented classes (class 2 (sample 9 and 15), 3 (sample 33) and 9 (sample 1)). Next are the samples of class 7 followed by the samples of class 8, which are clustered completely to right. Two samples, sample 6 and 29, belonging to class 8 are clustered with the samples of class 7, but overall this clustering can be considered as a good result, considering the non supervised nature of the technique. The color map clearly shows the correlation between the genuine samples (sample 50-58) and the samples of class 7 in the middle of the map (blue color). The more the previous methods also showed that samples of class 8 were clustered with class 7 and vice-versa, even if a quite good overall differentiation could be observed. Again a structure in the data is revealed that indicate that there are differences between the data, not only for genuine and counterfeits, but also between the different groups of counterfeit samples. These differences indicate that a predictive classification model is possible.

a)



b)

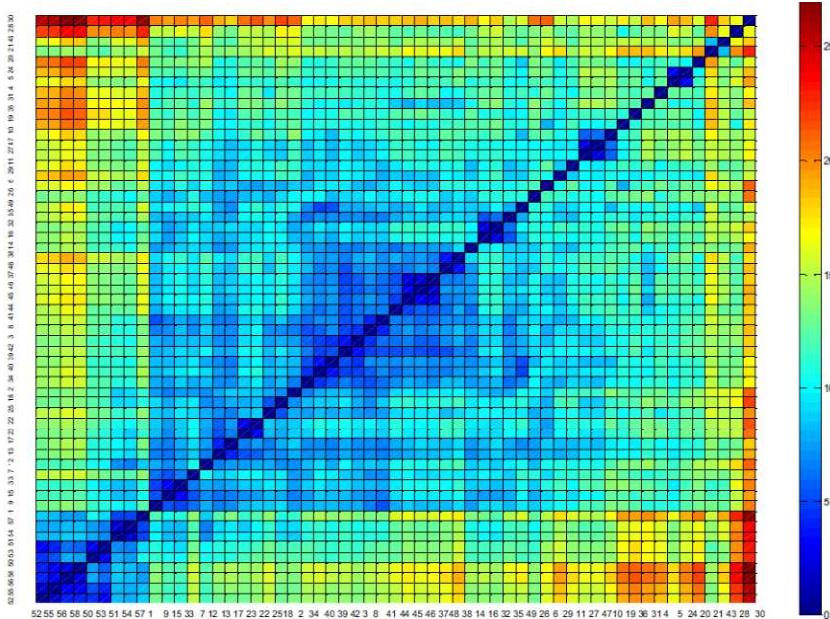


Figure 5: (a) dendrogram representing the hierarchical clustering obtained for the Viagra® data set, using the Euclidean distance as similarity measure and centroid linkage. (b) Color map representing the similarity matrix.

Cialis®-like samples

The results obtained for this data set are less clear than for the Viagra® data set. The only technique/similarity measure combination that resulted in a clustering of the genuine samples was the Ward method using the Manhattan distance. The dendrogram (Figure 6) shows that the genuine samples are clustered on the right side of the structure. When the other samples/groups are investigated the clustering seems more or less random. It seems that the differences between the samples belonging to the different classes are less important than for the Viagra® data set. The bad clustering of the different counterfeit classes could also be due to the fact that the classes are represented by fewer samples. In fact only class 7 is represented by a significant number of samples, while the other classes are rather small. This could also explain why only PLS, which is a supervised technique, is able to reveal differences and patterns.

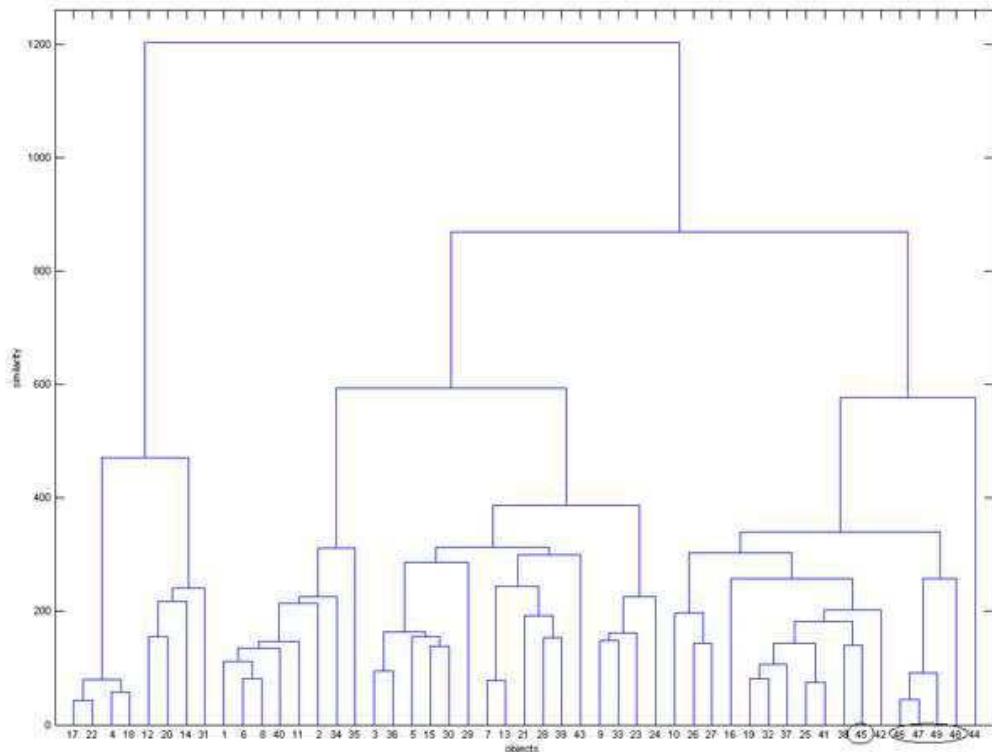


Figure 6: dendrogram representing the hierarchical clustering obtained for the Cialis® data set, using the Manhattan distance as similarity measure and Ward's method. The genuine samples are indicated by the ellipses

4.2. Modeling

4.2.1. k-NN

In a first step k-NN models were built using different numbers of nearest neighbors. The choice of the selected model was based on the cross validation error obtained using a leave-one-out cross validation procedure.

Viagra®-like samples

For the Viagra® data set the lowest cross-validation error was obtained for the model based on three nearest neighbors. A cross validation error, evaluated using Leave-one-out cross validation (LOOCV), was obtained of 0.7955. Closer investigation showed that all genuine samples and no counterfeit samples are classified as genuine. Four counterfeit samples were misclassified and five unclassified. The results obtained during external validation showed more or less the same trend. For the discrimination between genuine and counterfeit samples a 100 % correct classification is obtained. Three counterfeit samples are unclassified and one is misclassified (sample of class 2 classified as class 7).

Cialis®-like samples

For the Cialis® data set also the model based on three nearest neighbors was selected. The model has a cross validation error of 0.8108, evaluated using LOOCV and showed a 100% correct classification for the discrimination between counterfeit and genuine samples. Within the group of the counterfeit samples 7 samples were misclassified. The validation with the external test set also showed a 100% correct classification for the discrimination between genuine and counterfeit. Four counterfeit samples were misclassified (two samples of class 9 classified as class 3, a sample of class 8 as class 7 and one of class 10 as class 7).

4.2.2. CART

Classification trees were built with the two data sets. In a first step the maximal tree was built and pruned, followed by the selection of the optimal tree using a 10-fold cross validation procedure. For each data set two classification trees were built using respectively the Gini and the Twoing index. The information index was not used since it was not considered useful for solving the considered classification problem.

Viagra®-like samples

From the graph representing the cross-validation error in function of the tree complexity of the trees obtained using the Gini index as split criterion, the tree with complexity 4 could be selected as optimal tree. The tree (figure not shown) has a good homogeneity of its leaves with complete homogene leaves for the genuine class (class 1) and the class of the professional imitations (class 7). The two other leaves contain respectively the samples of the non professional imitations (class 8) with a sample of class 3 and one of class 9 and the samples of class 2. The cross validation error was 0.32 or 6.8%. During cross validation all genuine and no counterfeits samples were classified as genuine. Even if based on the cross validation the tree has good predictive properties, the external validation showed the

opposite. 50 % of the samples of the external test set were misclassified and one genuine sample was classified as counterfeit, which is unacceptable.

For the trees built with the Twoing index the tree with complexity 5 was selected as optimal tree (Figure 7). This tree has a good homogeneity for its leave with complete homogen leaves for the genuine class (class 1) and the classes 7 and 8 (separated over two leaves). The fifth class contains the 3 less represented classes of the data set: class 2, 3 and 9. The cross validation error was 0.23 or 4.5%. During external validation all genuine samples were classified as genuine and all counterfeits as counterfeits. Within the counterfeits 4 samples were misclassified (two samples of class 7 classified as class 2, one sample of class 8 as class 2 and one of class 2 as class 8).

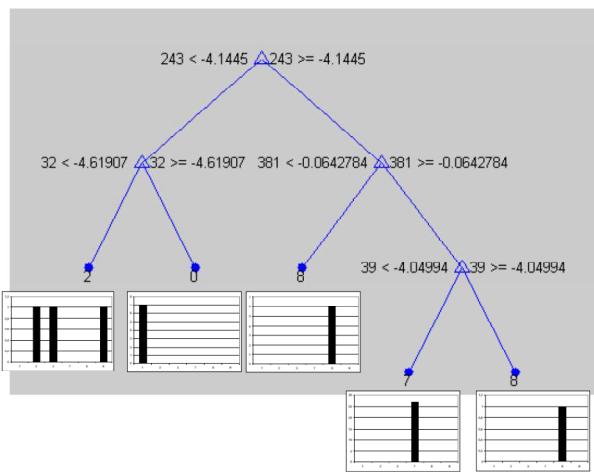


Figure 7: Classification tree obtained for the Viagra® data set with the Twoing index as split criterion. Each split is defined by the selected time point and its split value for the absorbance. Each leaf is defined by the class number of the class most represented in the leaf and the graph gives the distribution (homogeneity) of the different samples in the leaf.

Cialis®-like samples

Exactly the same approach was followed for the Cialis® data set. For both split criteria applied, trees were obtained with a quite good homogeneity of their leaves but external validation showed that neither the Gini nor the Twoing index resulted in a tree that was able to distinguish between counterfeit and genuine samples. This shows again that the strategy necessary to discriminate between genuine and counterfeit medicines will be different depending on the group of products considered. This can be due to the nature of the active ingredients and excipients, but also to the diversity of producers of a type of counterfeit medicines. It is clear that Cialis® is less counterfeited than Viagra®, due to the popularity of the brand name of Viagra®. Counterfeiting a popular brand name is just more lucrative.

4.2.3. SIMCA

Viagra®-like samples

Before applying SIMCA to the data set it was decided to remove the samples of the poor represented classes, i.e. the samples of class 2, 3 and 9. Since SIMCA models every class separately a minimum number of samples is necessary. The model is therefore based on only the genuine class and the classes 7 and 8. The number of principal components for each class was selected using leave one out cross validation. For the genuine class five PCs were selected for the model, for class 7 fifteen PCs and for class 8 six. The cross validation gave a correct classification rate of 97.6% which correspond to one sample of class 8 misclassified as class 7. For the discrimination between genuine and counterfeit samples a 100% correct classification rate is obtained. The external validation also showed a 100% correct classification for the discrimination between genuine and counterfeit. Two samples of the thirteen remaining samples in the test set were misclassified (one of class 8 classified as class 7 and one of class 7 classified as class 8).

Cialis®-like samples

For the same reasons as for the previous data set the sample representing class 10 was removed before applying SIMCA. Based on the leave one cross validation three principal components were selected for the genuine class (class 1), class 3 and class 9. For class 7 sixteen PCs were selected and for class 8 two. For this group of models a correct classification rate from the cross validation of 97.9% was obtained. This corresponds to one sample of class 8 misclassified as class 7. For the discrimination genuine/counterfeit all samples were correctly classified. Also the external validation resulted in a 100% correct classification for the discrimination between counterfeit and genuine. Three counterfeit (one of class 3 and two of class 9 classified as class 7) samples of the external test set were misclassified.

4.2.4. SVM

Two types of SVM models were built for both data sets. A least square SVM algorithm [43,44] was applied two times. Once using a radial basis function Kernel function (RBF-Kernel) and once using a Kernel function based on the standard inner product (lin-Kernel). Both functions are adapted for implicit higher dimension mapping. The selection of the models was based on a 10-fold cross validation procedure.

Viagra®-like samples

The best model was obtained by using the RBF-Kernel function for mapping the data. For the external test set a 100% correct classification rate was obtained for the discrimination between genuine and counterfeit samples, while the counterfeit sample of class 2 was misclassified as class 7. Also the model using the lin-Kernel function gave satisfying results with a 100% correct classification rate for the discrimination between genuine and counterfeit samples and only the sample of class 2 was misclassified. In comparison with the first model 2 counterfeit samples, one of class 7 and one of class 8, were unclassified.

Cialis®-like samples

For the Cialis® data set the SVM model using the RBF Kernel function gave unsatisfying results. The genuine samples of the external test set were considered as unclassified, while five counterfeit samples of the test set were misclassified. The model obtained with the lin-Kernel function performed better and gave a 100% correct classification rate for the discrimination between genuine and counterfeit samples. Only 3 counterfeit samples (a sample of class 8, 9 and 10) were misclassified.

5. Discussion and conclusions

The analysis of the two considered data sets was started with exploratory chemometric tools in order to reveal structure in the data set and differences between the different samples. For the differences between the samples, the analysis was focused on differences between genuine and counterfeit samples and on differences between the different groups of counterfeit samples.

For the Viagra® data set, a clear differentiation between genuine and counterfeit samples could be observed with all chemometric techniques applied (PCA, D-PLS, PP and hierarchical clustering). Despite this, only PP with the Yenyukov index and hierarchical clustering using centroid linkage with the Euclidean distance are able to reveal differences among the different counterfeit samples. Where PP indicates only a tendency to cluster between groups 7 and 8, next to the clear differentiation between genuine and counterfeits, hierarchical clustering clearly clusters the samples in genuine, the three smaller groups, class 7 and class 8. This indicated that there are clear differences among the samples and that it should be possible to model these differences. When investigating the loadings in the projection techniques, it was observed that no specific regions could be indicated as responsible for the differentiations between genuine and counterfeit samples and among counterfeit samples. This could mean that it is the complete profile that enables the differentiations.

For the Cialis® data set it was more tedious to find a structure within the data. Only D-PLS, which is a supervised technique and hierarchical clustering using Ward's method with the Manhattan distance were able to show a differentiation between genuine and counterfeit samples. The reason for a more difficult clustering in this data set is probably due to the nature of the different samples. The counterfeit samples of Cialis® are less diverse than the ones of Viagra®. They have practically all the same color and size, where for the Viagra® data set there is a wide range of shapes, sizes and colors. Though based on the results of the exploratory analysis, it should be possible to at least model the differences between genuine and counterfeits. Investigation of the loadings of the PLS analysis lead to the same conclusion as for the Viagra® data set. The complete chromatographic profile seems responsible for the differentiation between genuine and counterfeit samples.

Based on the results of the exploratory analysis four different modeling techniques were applied and compared for their predictive performance for the discrimination between genuine and counterfeit samples and for the classification in the different groups defined by RIVM. Table 3 summarizes the results, obtained for the external test set, for the selected models of the techniques applied.

Table 3: Overview of the performance of the different models for the external test sets expressed as correct classification rates (CCR)

| Technique | Viagra data set | | | Cialis data set | | |
|-----------|-------------------------|---------------|--------------|-------------------------|---------------|--------------|
| | CCR Genuine/counterfeit | CCR overall | | CCR Genuine/counterfeit | CCR overall | |
| | | misclassified | unclassified | | misclassified | unclassified |
| kNN | 100% | 1/14 | 3/14 | 100% | 4/12 | 0/12 |
| CART | 100% | 4/14 | 0/14 | 0% | 8/12 | 0/12 |
| SIMCA | 100% | 2/13 | 0/13 | 100% | 3/12 | 0/12 |
| LS-SVM | 100% | 1/14 | 0/14 | 100% | 3/12 | 0/12 |

For the Viagra® data set all models selected are able to differentiate between genuine and counterfeit medicines and have a 100% correct classification rate for this distinction in cross validation as well as for external validation. For the classification of the samples in the different groups of counterfeit, the LS-SVM model performs the best with only one sample misclassified in the external test set. Here it should also be mentioned that SIMCA gave also good classification results with only two misclassified samples for the external test set. The disadvantage here is that it is not possible to have a correct modeling of very small classes, since each class is modeled separately. Though it could be expected, that when the data set would be bigger, with higher representation of the smaller classes, SIMCA would perform as good as LS-SVM.

For the Cialis® data set kNN, SIMCA and LS-SVM gave very comparable results, while CART failed completely in modeling the differences among the samples. kNN, SIMCA and LS-SVM all had 100% correct classification rates for the classification in genuine and counterfeit and SIMCA and LS-SVM had both 3 counterfeit samples misclassified.

To conclude it can be said that impurity profiles or chromatographic fingerprints could be a valuable approach in the detection of counterfeit medicines, as well as in the classification of these counterfeit samples in different categories linked to their risk for public health. The possibility to differentiate genuine and counterfeit samples as well as the different groups of counterfeit samples was demonstrated with different exploratory chemometric techniques.

To build predictive models based on chromatographic fingerprints LS-SVM performed best considering the results for both the Viagra® and the Cialis® data set. Though, based on the results for SIMCA it should be mentioned that SIMCA can be a good compromise, since despite the slightly worse predictive properties, SIMCA is easier to calculate, understand and to interpret.

Off course it should be mentioned that the approach proposed here is only valid for the group of Viagra® and Cialis® samples. The applicability of the approach to other groups of medicines like counterfeit antibiotics or slimming products can be different both in

analytical as in chemometric techniques. Each group of counterfeit medicines should be considered as a separate problem and the best approach should therefore be selected / investigated for each group of medicines.

References

- [1] European Alliance For Acces to Safe Medicines: www.eaasm.eu
- [2] C. Clift, Combating Counterfeit, Falsified and Substandard Medicines Defining the Way Forward?, Chatham House Briefing Papers, November 2010
- [3]http://www.emea.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004481.pdf
- [4] www.coe.int/medicrime
- [5] WHO, sixty-second world health assembly item 12.9, counterfeit medical products, april 2009. http://aps.who.int/gb/ebhwa/pdf_files/A62/A62_13-en.pdf
- [6] B.J. Venhuis, D.M. Barends, M.E. Zwaagstra, D. de Kaste, Recent developments in counterfeit and imitations of Viagra, Cialis and Levitra, RIVM report 370030001/2007, Bilthoven, 2007.
- [7] M.J. Vredenbregt, L. Blok-Tip, R. Hoogerbrugge, D.M. Barends, D. de Kaste, J. Pharm. Biomed. Anal. 40 (2006) 840-849.
- [8] T.S. Reddy, A.S. Reddy, P. Devi, J. Planar Chromatogr. - Modern TLC 19 (2006) 427-431
- [9] J. Reepmeyer, J. Woodruff, J. Pharm. Biom. Anal. 44 (2007) 887-893.
- [10] J. Reepmeyer, D. André d'Avignon, J. Pharm. Biomed. Anal. 49 (2009) 145-150.
- [11] S. Gratz, M. Zeller, D. Mincey, C. Flurer, J. Pharm. Biomed. Anal 50 (2009) 228-231.
- [12] J. Reepmeyer, J. Woodruff, J. Chromatogr. A 1125 (2006) 67-75.
- [13] Y-H. Lam, W-T. Poon, C-K. Lai, A Y-W. Chan, T W-L. Mak, J. Pharm.Biomed. Anal. 46 (2008) 804-807.
- [14] X. Ge, M-Y. Low, P. Zou, L. Lin, S.O.S. Yin, B.C. Bloodworth, H-L. Koh, J. Pharm. Biomed. Anal. 48 (2008) 1070-1075.
- [15] H.J. Park, H.K. Jeong, M.I. Chang, M.H. Im, J.Y. Jeong, D.M. Choi, K. Park, M.K. Hong, J. Youm, S.B. Han, D.J. Kim, J.H. Park, S.W. Kwon, Food Addit. Contam. 24 (2007) 122-129.
- [16] L. Blok-Tip, B. Zomer, F. Bakker, K.D. Hartog, M. Hamzink, J. ten Hove, M. Vredenbregt, D. de Kaste, Food Addit. Contam. 21 (2004) 737-748.
- [17] P. Zou, P. Hou, S.S-Y. Oh, Y.M. Chong, B.C. Bloodworth, M-Y. Low, H-L. Koh, J. Pharm. Biomed. Anal. 47 (2008) 279-284.

- [18] B.J. Venhuis, G. Zomer, D. de Kaste, *J. Pharm. Biomed. Anal.* 46 (2008) 814-817.
- [19] B.J. Venhuis, G. Zomer, M. Hamzink, H.D. Meiring, Y. Aubin, D. de Kaste, *J. Pharm. Biomed. Anal.* 54 (2011) 735-741.
- [20] B.J. Venhuis, G. Zomer, M.J. Vredenbregt, D. de Kaste, *J. Pharm. Biomed. Anal.* 51 (2010) 723-727.
- [21] A. Häberli, P. Girard, M-Y. Low, X. Ge, *J. Pharm. Biomed. Anal.* 53 (2010) 24-28.
- [22] H-M. Lee, C.S. Kim, Y.M. Jang, S.W. Kwon, B-J. Lee, infrared spectroscopy and nuclear magnetic resonance spectroscopy *J. Pharm. Biomed. Anal.* 54 (2011) 491-496.
- [23] P-Y Sacré, E. Deconinck, T. De Beer, P. Courselle, R. Vancauwenberghe, P. Chiap, J. Crommen, J.O. De Beer, *J. Pharm. Biomed. Anal.* 53 (2010) 445-453.
- [24] M. de Veij, A. Deneckere, P. Vandenabeele, D. de Kaste, L. Moens, *J. Pharm. Biomed. Anal.* 46 (2008) 303-309.
- [25] S. Trefi, C. Routaboul, S. Hamieh, V. Gilard, M. Malet-Martino, R. Martino, *J. Pharm. Biomed. Anal.* 47 (2008) 103-113.
- [26] J.K. Maurin, F. Pluci ski, A.P. Mazurek, Z. Fija ek, *J. Pharm. Biomed. Anal.* 43 (2007) 1514-1518.
- [27] A.S. Amin, M.E. Moustafa, R. El-Dosoky, *J. AOAC Int.* 92 (2009) 125-130.
- [28] A.L. Rodomonte, M.C. Gaudiano, E. Antoniella, D. Lucente, V. Crusco, M. Bartolomei, P. Bertocchi, L. Manna, L. Valvo, N. Muleri, *J. Pharm. Biomed. Anal.* 53 (2010) 215-220.
- [29] I. Wawer, M. Pisklak, Z. Chilmonczyk, *J. Pharm. Biomed. Anal.* 38 (2005) 865-870.
- [30] C.-J. Xu, Y.-Z. Liang, F.-T. Chau, Y. Vander Heyden, *J. Chromatogr. A* 1134 (2006) 253-259.
- [31] S.-K. Yan, W.-F. Xin, G.-A. Luo, Y.-M. Wang, Y.-Y Cheng, *J. Chromatogr. A* 1090 (2005) 90-97.
- [32] C. Han, Y. Shen, J. Chen, F.S. Lee, X. Wang, *J. Chromatogr. B* 862 (2008) 125-131.
- [33] P. Hu, Q.-L. Liang, G.-A. Luo, Z.-Z. Zhao, Z.-H. Jiang, *Chem. Pharm. Bull.* 53 (2005) 677-683.
- [34] Y. Ni, Y. Lai, S. Brandes, S. Kokot, *Anal. Chim. Acta* 647 (2009) 149-158.
- [35] F. Xiaohui, W. Yi, C. Yiyu, *J. Pharm. Biomed. Anal.* 40 (2006) 591-597.
- [36] P.Y. Sacré, E. Deconinck, M. Daszykowski, P. Courselle, R. Vancauwenberghe, P. Chiap, J. Crommen, J. O. De Beer, *Anal. Chim. Acta* 701 (2011) 224-231.

- [37] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers- Verbeke: Handbook of Chemometrics and Qualimetrics-Part A. Elsevier Science, Amsterdam, 1997 .
- [38] C. Croux, A. Ruiz-Gazen, COMPSTAT : Proceedings in computational statistiques 1996, Physica-Verlag, Heidelberg, pp. 211-217.
- [39] D.L. Massart, L. Kaufman, The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis, Wiley, New York, 1983.
- [40] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and regression trees; Wadsworth & Brooks, Monterey, 1984.
- [41] E. Deconinck, T. Hancock, D. Coomans, D.L. Massart, Y. Vander Heyden, J. Pharm. Biomed. Anal. 39 (2005) 91-103.
- [42] E. Deconinck, P.-Y. Sacré, D. Coomans, J. De Beer. J. Pharm. Biomed. Anal. 57 (2012) 68-75.
- [43] J. Luts, F. Ojeda, R. Van De Pls, B. De Moor, S. Van Huffel, J. A. K Suykens, Anal. Chim. Acta 665 (2010) 129-145.
- [44] K. De Brabanter, P. Karsmakers, F. Ojeda, C. Alzate, J. De Brabanter, K. Pelckmans, B. De Moor, J. Vandewalle, J.A.K. Suykens, LS-SVMLab Toolbox User's Guide version 1.8, www.esat.kuleuven.be/sista/lssvmlab.
- [45] European Directorate for the Quality of Medicines, draft monography of sildenafil citrate, Pharmeuropa 23 (2011), 381–383.
- [46] European Directorate for the Quality of Medicines, draft monography of tadalafil, Pharmeuropa 22 (2010), 328–332.
- [47] M. Daszykowski, B. Walczak, Trends Analyt. Chem. 25 (2006) 1081-1096. [48] M. Daszykowski, B. Walczak, J. Chromatogr. A 1176 (2007) 1-11.
- [49] R.D. Snee, Technometrics 19 (1977) 415-428.
- [50] R.W. Kennard, L.A. Stone, Technometrics 11 (1969) 137-148.