This item is the archived peer-reviewed author-version of:

ATR-FTIR spectroscopy and chemometrics : an interesting tool to discriminate and characterize counterfeit medicines

Reference:

Custers Deborah, Cauwenbergh T., Bothy J.L., Courselle P., De Beer J.O., Apers Sandra, Deconinck E..- ATR-FTIR spectroscopy and chemometrics : an interesting tool to discriminate and characterize counterfeit medicines
Journal of pharmaceutical and biomedical analysis - ISSN 0731-7085 - 112(2015), p. 181-189
DOI: <http://dx.doi.org/doi:10.1016/j.jpba.2014.11.007>

ATR-IR spectroscopy and chemometrics: an interesting tool to discriminate and characterize counterfeit medicines

D. Custers^{1,2}, T. Cauwenbergh¹, J.L. Bothy¹, P. Courselle¹, J.O. De Beer¹, S. Apers^{2‡}, E. Deconinck^{1*‡}

¹ Division of Food, Medicines and Consumer Safety, Section Medicinal Products, Scientific Institute of Public Health (WIV-ISP), J. Wytsmanstraat 14, B-1050 Brussels, Belgium

² Research group NatuRA (Natural products and Food - Research and Analysis), Department of Pharmaceutical Sciences, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium

[‡] Equally contributing project leaders

Abstract

Counterfeit medicines pose a huge threat to public health worldwide. High amounts of counterfeit pharmaceuticals enter the European market and therefore detection of these products is essential. Attenuated Total Reflection Fourier-Transform infrared spectroscopy (ATR-FTIR) might be useful for the screening of counterfeit medicines since it is easy to use and no sample preparation is required. Furthermore, this approach might be helpful to customs to obtain a first evaluation of suspected samples. This study proposes a combination of ATR-FTIR and chemometrics to discriminate and classify counterfeit medicines. A sample set, containing 209 samples in total, was analyzed using ATR-FTIR and the obtained spectra were used as fingerprints in the chemometric data-analysis which included Principal Component Analysis (PCA), k-Nearest Neighbours (k-NN), Classification and Regression Trees (CART) and Soft Independent Modelling of Class Analogy (SIMCA). First it was verified whether the mentioned techniques are capable to distinguish samples containing different active pharmaceutical ingredients (APIs). PCA showed a clear tendency of discrimination based on the API present; k-NN, CART and SIMCA were capable to create suitable prediction models based on the presence of different APIs. However k-NN performs the least while SIMCA performs the best. Secondly, it was tested whether these three models could be expanded to discriminate between genuine and counterfeit samples as well. k-NN was not able to make the desired discrimination and therefore it was not useful. CART performed better but also this model was less suited. SIMCA, on the other hand, resulted in a model with a 100% correct discrimination between genuine and

counterfeit drugs. This study shows that chemometric analysis of ATR-FTIR fingerprints is a valuable tool to discriminate genuine from counterfeit samples and to classify counterfeit medicines.

Keywords: counterfeit medicines, Attenuated Total Reflection Fourier-Transform infrared spectroscopy, chemometrics, fingerprints

* Corresponding author: Eric.Deconinck@wiv-isp.be

Tel. +32 2 642 51 36

Fax. +32 2 642 53 27

1. Introduction

Counterfeit medicines pose a huge threat to public health worldwide [1]. Both developing and industrialized countries are exposed to pharmaceutical counterfeiting. A counterfeit medicine is defined by the World Health Organization (WHO) as “one which is deliberately and fraudulently mislabelled 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 active ingredients, with insufficient active ingredients or with fake packaging” [2]. These counterfeited medicines are often manufactured by uncontrolled or street laboratories without respecting Good Manufacturing Practices (GMP) [3]. They are not subjected to any form of quality control [4] and therefore their safety, efficacy and quality cannot be guaranteed [2]. Health risks, caused by counterfeit medicines, might be due to the absence of active ingredients, the presence of incorrect active ingredients, an incorrect dosage, the presence of high concentrations of potential toxic secondary components and fake packaging or documentation [5]. Assessing the actual extent of pharmaceutical counterfeiting is very difficult due to its illicit and clandestine character [5]. It is estimated that about 1% of the total medicines market of industrialized countries, such as European countries, the United States, Japan, etc., is covered by counterfeit medicines. In countries of the former Soviet Union about 20% of the medicines market is covered by counterfeit pharmaceuticals. This amount reaches even over 30% in African countries and parts of Asia and Latin-America. Furthermore, about 50% of all medicines bought online from websites which cover up their physical address, are suspected to be forged [1]. In fact, the extension of the Internet is one of the main reasons for the increasing threat posed by counterfeit drugs, especially in industrialized countries [6]. The types of medicines which are sold most as counterfeit in industrialized countries are commonly referred to as ‘life style drugs’ and comprise phosphodiesterase type 5 (PDE-5) inhibitors for the treatment of erectile dysfunction, weight-loss products (containing anorexics) and anabolic hormones [6; 7].

The high amount of counterfeit medicines entering the European market and the possible risk of these products to enter the legal medicine supply chain enforce the need to detect these forged pharmaceuticals and distinguish them from genuine ones [4]. For this purpose numerous analytical techniques were already described in literature. These techniques may be divided in two main groups: chromatographic and spectroscopic techniques. Chromatographic techniques are used for separation, identification and quantification of active substances. Commonly used chromatographic techniques are thin layer chromatography (TLC), which is easy and cheap, but also high pressure liquid chromatography (HPLC) coupled with UV detection or mass spectrometry, which are more expensive and sophisticated [8].

However, spectroscopic techniques are often preferred over chromatographic techniques to detect counterfeit medicines. This is due to the fact that spectroscopy is fast, no or less sample preparation is needed and some spectroscopic techniques are non-destructive [8]. Near infrared spectroscopy (NIR) [9-13], Fourier-Transform infrared spectroscopy [14; 15], Raman spectroscopy [13; 16; 17], nuclear magnetic resonance (NMR) [16; 18; 19] and X-ray diffraction [20] have already been applied successfully for the detection of counterfeit medicines. Furthermore Sacré et al. compared and combined Fourier-Transform infrared spectroscopy, Raman spectroscopy and Near infrared spectroscopy [21].

The use of Attenuated Total Reflection Fourier-Transform infrared spectroscopy (ATR-FTIR) to detect counterfeit medicines is a fairly new concept in literature. Ortiz et al. made use of ATR-FTIR to distinguish counterfeit samples from genuine Viagra® and Cialis® samples and to demonstrate the use of pharmaceutical powders from common origins in the production of counterfeit drugs from distinct seizures [22]. Champagne et al. used ATR-FTIR to screen raw materials, used in the formulation and production of dietary supplements, for adulteration [23]. Indeed, ATR-FTIR might be a useful technique for the screening of counterfeit samples due to the fact that this technique does not require sample preparation; samples can be measured by pressing them directly on the crystal

surface without the need of preparing KBr pellets. Therefore, this method is less time consuming and easy in use.

In this study, 209 samples were analyzed using ATR-FTIR spectroscopy. These samples include genuine Viagra® and Cialis® samples, generic products of Viagra®, counterfeit samples containing sildenafil, tadalafil or both and placebo samples. These spectra were used as fingerprints in the chemometric data-analysis. The chemometric data-analysis includes several techniques, i.e. Principal Component Analysis, k-Nearest Neighbours, Classification and Regression Trees and Soft Independent Modelling of Class Analogy, and aims to characterize counterfeit samples and to discriminate between genuine and counterfeit medicines.

2. Methods

2.1 Samples

All counterfeit samples were donated by the Federal Agency for Medicines and Health Products (FAMHP) in Belgium. Genuine samples of Viagra® and Cialis® and generic products of Viagra® were purchased in a local pharmacy.

2.2 Recording of ATR-FTIR spectra

All spectra were recorded in absorption mode using a Nicolet iS10 FT-IR spectrometer (Thermo Scientific, Madison, USA) accommodated with a Smart iTR accessory for ATR sampling and a deuterated triglycine sulfate (DTGS) detector. The Smart iTR accessory is equipped with a single bounce diamond crystal. A diamond crystal has the advantage of being inert and robust and therefore it is well suited for the analysis of hard and abrasive samples. The Smart iTR accessory was calibrated weekly using a polystyrene film.

As already mentioned, no sample preparation is necessary. Prior to spectrum recording, one tablet from each sample was crushed and homogenized using a pestle and mortar. Subsequently a small amount of homogenized sample was pressed directly on the diamond crystal by usage of the Smart iTR's high pressure clamp. Each spectrum was measured at a spectral resolution of 4 cm^{-1} and consisted of 16 co-added scans. Spectrum recording was performed in the range of $4000\text{ to }400\text{ cm}^{-1}$ with an estimated time for data collection of 23 seconds per spectrum. Spectral data were obtained using the OMNIC Software version 8.3 (Thermo Scientific, Madison, USA). After each measurement, the crystal was cleaned using a soft tissue soaked with methanol and left to dry in ambient air. Before each sample a blank was measured to check the crystal for contamination and carry over using the absorbance limits for contamination defined by the European Directorate for the Quality of Medicines and HealthCare (EDQM) [24]. Every hour a background spectrum was measured against air using identical instrumental conditions as the samples. All spectra were saved in CSV format which is suitable for input in statistical software.

Exemplary ATR-FTIR spectra obtained for both genuine and counterfeit samples are shown in Fig. 1.

2.3 Sample set

A sample set of 209 samples in total was analysed using the described ATR-FTIR technique and used for the chemometric data-analysis. This sample set contained 13 genuine Viagra[®] samples, 33 generic products of Viagra[®], 11 genuine Cialis[®] samples, 68 counterfeit samples containing sildenafil, 21 counterfeit samples containing tadalafil, 18 counterfeit samples containing both sildenafil and tadalafil and 45 placebo samples.

2.4 Chemometric methods

2.4.1 Data-preprocessing

As already mentioned before, ATR-FTIR spectra were measured in the range of 4000 to 400 cm^{-1} . However, this study focused on the fingerprint region ranging from 1800 to 650 cm^{-1} . In this spectral region primary differences in various compounds, such as the phosphodiesterase type 5 (PDE-5) inhibitors, are highlighted and characteristic groups show absorption. Fig. 2 shows the ATR-FTIR spectra of a sildenafil and tadalafil reference. The ATR-FTIR spectrum of sildenafil (Fig. 2a) shows a number of peaks which are characteristic for certain functional groups of this molecule: the 1676 cm^{-1} peak is the result of C=N stretching (1690 – 1640 cm^{-1}), the 1647 cm^{-1} band is correlated to N–H bending; the C=C bonds absorb at 1490 cm^{-1} ; the C–N bond present in the O=C–N functional group gives rise to a peak at 1402 cm^{-1} and the 1269 cm^{-1} peak can be attributed to the aryl C-N bonds. The ATR-FTIR spectrum of tadalafil (Fig. 2b) also shows a number of characteristic peaks: the C=O bonds result in the 1700 cm^{-1} peak and the C–C bonds from the ketone group are responsible for the 1280 and 1172 cm^{-1} bands [22].

2.4.2 Exploratory data-analysis

Principal Component Analysis was applied to test whether this technique can result in a clustering of samples which might be useful for the creation of classification models. These classification models were generated in a later phase by making use of supervised (modelling) techniques.

Principal Component Analysis (PCA) is a projection method which allows the projection of high dimensional data into a low dimensional space. This low dimensional space is defined by new orthogonal latent variables, commonly referred to as principal components (PCs). The result of such an analysis is a reduction in the number of variables by calculating linear combinations (= principal components) of these original variables. The first constructed principal component represents the highest variance in the data; the 2nd principal component explains the highest residual variance, around the first principal component. Hence this 2nd principal component is by definition orthogonal

to the first. The same principle is repeated for PC3 around the plane, defined by PC1 and PC2, etc. PCA analysis results in 2 matrices: a loading matrix and a score matrix. The loadings are a measure for the weight of each original variable, meaning that variables with higher loadings contribute more to the construction of a given principal component. PCA projects each object (= sample) on the created principal components. These projections are referred to as scores and provide information about the (dis)similarities among the objects [8; 25].

2.4.3 Selection of a test set for external validation of models

The data set was split into a training and test set using the Duplexx algorithm. The training set is used to generate classification models; the test set is selected to perform an external validation of the created classification models. When applying the Duplexx algorithm, the two most distant samples are assigned to the training set. The second pair of most distant samples is selected and included in the test set. Next, the sample, situated furthest from the pair of samples assigned to the training set, is selected and allocated to the training set. This procedure is repeated for the test set. After that the process is repeated, continuously alternating between the test and training set, until the desired number of samples in the test set is obtained [26].

As a result, the training set contains 167 samples. The test set comprises 42 samples.

2.4.4 Modelling techniques

After performing an exploratory data-analysis a number of modelling techniques were applied to test whether appropriate classification models could be obtained which also might serve to predict unknown samples.

2.4.4.1 k-Nearest Neighbours (k-NN)

k-NN is from a mathematical point of view a fairly simple technique to construct classification models. This method calculates the Euclidean distance between an unknown object and each of the objects of the training set. If the training set includes n samples, then n distances are calculated. Subsequently, the k nearest objects to the unknown object are selected and a majority rule is applied; the unknown object is assigned to the class to which the majority of the k neighbouring objects of the training set belong. The number of nearest neighbours (k) to be included in the construction of a classification model has to be determined by optimization [27]. A number of k-NN models are built using different values for k . The best model is selected based on the cross validation error obtained using a 10-fold cross validation procedure.

2.4.4.2 Classification and Regression Trees (CART)

CART is a supervised non-parametric technique, which can be used to solve both classification and regression problems. CART produces a classification tree (Fig. 3), used to solve classification problems, if the dependent variable is categorical. When the dependent variable is continuous, it creates a regression tree [28; 29].

A CART analysis consists of three steps. The first step is the creation of the maximum tree, starting at the tree-root containing all samples of the training set. This maximum tree is built using a binary split-procedure in which a mother group is split in two daughter groups. Every daughter group becomes a mother group in the next step of the splitting procedure. At each level, the split-procedure is based on one descriptor and its split value. To select the most appropriate descriptor and split value, an algorithm is used which considers all descriptors and all possible split values. The descriptor and split value are chosen in such a way that two daughter groups are obtained which show more homogeneous response variable values. In other words, the best descriptor and split

value result in the highest decrease in impurity between the mother group and daughter groups. In case of classification trees, the impurity can be defined by different split criteria. The split criterion used in this study is the Gini index. The split-procedure is repeated until the maximal tree is created. The maximal tree is the tree in which each end node (leaf) contains one object, or a predefined number of objects, or homogeneous groups [28; 29].

In the second step, the maximal tree is pruned since it is overgrown. This tree closely describes the used training set, resulting in overfitting. By pruning, i.e. cutting terminal branches, a number of smaller and less complex trees is derived from the maximal tree.

The final step consists of selecting the optimal tree, based on the evaluation of the predictive error. For this purpose cross-validation (CV) is often used. In this study a 10-fold CV was used [8; 28; 29].

2.4.4.3 Soft Independent Modelling of Class Analogy (SIMCA)

SIMCA is a supervised classification technique that models each class of samples separately by defining a number of principal components which are derived from PCA. This method starts by evaluating the optimal number of principal components, required to describe each training class individually, using a cross-validation procedure. Next, classification rules are constructed. Two critical values are taken into account: (1) one for the Euclidean distances towards the SIMCA model (often referred to as orthogonal distances) and (2) the Mahalanobis distances calculated in the space of the scores. These two critical values determine a restricted space around the samples belonging to one particular class. In other words, they are a measure for the boundaries of the model. The position of a new object is calculated using the scores and loadings of the created PCA model. If the object is situated within the restricted space of a training class, defined by the orthogonal and Mahalanobis distances, then the object is assigned to that particular class.

Confidence limits were set at 95%. SIMCA is a soft classification method, meaning that a sample can be assigned to one or more existing classes or to any [8].

2.4.5 Software

All data treatments were performed using Matlab version 8.0.0 (The Mathworks, Natick, USA). The algorithms of PCA and k-NN were part of the ChemoAC toolbox (Freeware, ChemoAC Consortium, Brussels, Belgium, version 4.0). The used CART algorithm was programmed according to the original CART algorithm proposed by Breiman et al. [30]. The toolbox for SIMCA was downloaded from the Matlab Central (www.mathworks.com/matlabcentral/fileexchange/30762-soft-independent-modelling-of-class-analogy-simca).

3. Results

3.1 Discrimination based on active substances

This study investigated whether ATR-FTIR in combination with chemometrics can distinguish between samples containing different active pharmaceutical ingredients (APIs). The obtained absorbance values in the spectra/fingerprints were used as explanatory variables. The class numbers, based on the presence of API, were included as response variables for the supervised techniques (Table 1).

3.1.1 Exploratory data-analysis

A simple PCA was performed, retaining only 2 principal components (PCs) since 95,5% of the total variance was explained (PC1 = 91,9% and PC2 = 3,6%). The resulting score plot (Fig. 4) shows a faint clustering. Four groups of samples can be separated to a certain extent: (1) samples containing

sildenafil of which most samples are situated in the top left corner, (2) placebo samples in the lower part of the score plot, (3) samples containing tadalafil which are mostly concentrated between the two large groups of placebo samples and samples containing sildenafil and (4) samples containing both sildenafil and tadalafil which are mostly clustered on the right part of the score plot. Study of the loading plot (figure not shown) did not reveal any absorbance values which could account for this clustering, thereby showing that the entire fingerprints account for the obtained discrimination.

Despite the fact that the obtained clustering is a little indistinct, there is a clear tendency of discrimination based on the API present. This tendency of distinction might serve as a foundation to create classification models. Therefore, it was tested whether classification models based on the API present could be built using supervised techniques. Four groups of samples can be distinguished according to the APIs present (Table 1).

3.1.2 k-Nearest Neighbours

k-NN analysis revealed that the best model was obtained by including 7 nearest neighbours. This model was characterized by a correct classification rate of cross validation of 76,05%, which indicated that 40 samples of the training set were either misclassified or unclassified. A closer investigation of the mis/unclassified samples is given in Table 2. This table shows that most misclassifications in the training set occur in class 3 (counterfeit samples containing both sildenafil and tadalafil) and class 4 (placebo samples). Eight samples belonging to class 3 were assigned to class 1. However, these misclassifications seem logic since most of these samples contain much higher doses of sildenafil compared to tadalafil (often 8 to 10 times more). The external validation of the obtained model generated similar results. Eight out of 42 test set samples were either misclassified or unclassified, resulting in a correct classification rate of external validation of 80,95%. An overview of the mis/unclassified samples of the test set is shown in Table 2.

3.1.3 CART

The graph representing the cross validation error in function of tree complexity of the obtained trees showed that the tree with complexity 5 (= tree containing 5 leaves) should be selected as the optimal tree (Fig. 5). This tree had a correct classification rate of cross validation of 88,62% due to the misclassification of 19 training set samples. Fig. 5 shows that most misclassifications occurred in class 3; seven of these samples were classified as class 1, four other samples were classified as class 2. As already mentioned in the section of the k-NN model, most of these counterfeit samples contain much higher doses of sildenafil than tadalafil. Therefore, the former misclassifications can be explained. The eight remaining misclassifications are also visualized in Fig. 5. The external validation resulted in a correct classification rate of 83,33%, which indicated that 7 out of 42 test set samples were misclassified. Two samples of class 1 were misclassified (1 sample as class 3, the other as class 4); three samples of class 2 were assigned to a wrong class (2 samples as class 1, the 3rd as class 4) and two samples belonging to class 4 were incorrectly classified (1 sample as class 1, the other as class 2).

3.1.4 SIMCA

SIMCA is a technique which models every class separately by selecting an appropriate number of principal components necessary to describe each class. This number of principal components was selected using leave-one-out cross validation. Twelve PCs were selected to model class 1 (samples containing sildenafil), thirteen PCs were kept to describe class 2 (samples containing tadalafil), ten PCs were retained for class 3 (samples containing both sildenafil and tadalafil) and for class 4 (placebo samples) sixteen PCs were selected. The obtained SIMCA model was characterized by a correct classification rate of cross validation of 94,01%, indicating that only 10 out of 167 training set

samples were misclassified. From both classes 1 and 2 only one sample was misclassified as class 4; seven samples of class 3 were misclassified (5 samples as class 1, 2 other samples as class 2 and 4) and class 4 contained one sample which was attributed to class 1. The external validation resulted in a correct classification rate of 90,48% due to the misclassification of 4 samples; 3 samples of class 2 which were assigned to class 1 and 1 sample of class 3 was misclassified as class 1.

Since three good models are acquired, it was verified whether these models could be extended to not only distinguish between different active ingredients, as they did thus far, but also to discriminate between genuine and counterfeit samples. Therefore, a new classification system was set up according to Table 3 and a new training and test set were selected.

3.2 Discrimination between genuine and counterfeit medicines based on active substances

3.2.1 k-NN

When extending the classification system, the best k-NN model was obtained using only 3 nearest neighbours. This model generated a correct classification rate of cross validation of 57,49%, indicating that 71 samples of the training set were either misclassified or unclassified. A closer investigation of the model revealed that one Viagra® sample, one Cialis® sample and 8 generic products of Viagra® were either unclassified or misclassified as counterfeit. Furthermore, four genuine Viagra® samples were classified as generic product and 3 generics were classified as genuine Viagra®. In addition a number of counterfeit samples were misclassified as genuine/generic; seven samples of class 4, three samples of class 6 and one sample of class 7 were assigned to class 3 (Viagra® generics) and one class 5 sample was misclassified as class 2 (genuine Cialis®). The remaining misclassifications occurred within the 4 classes of counterfeit samples. These results clearly show that k-NN is not suited for the aimed model expansion, despite the fact that the

external validation gave a better result. 76,19% of the test set samples was classified correctly, although 2 important misclassifications occurred: one Cialis[®] sample was classified as counterfeit (class 5) and one counterfeit with sildenafil (class 4) was considered a Viagra[®] generic (class 3).

3.2.2 CART

According to the graph representing the cross validation error in function of tree complexity, the tree containing 7 leaves was considered to be the optimal tree (Fig. 6). This tree was characterized by a correct classification rate of cross validation of 74,25%. Survey of the leaves revealed that 43 training set samples were misclassified. Unfortunately, a number of these misclassifications concerned genuine/generic samples which were considered counterfeit and counterfeit samples which were misclassified as genuine or generic. One genuine Viagra[®] sample (class 1) and 12 generic products of Viagra[®] (class 3) were classified as counterfeit (class 4). From both classes 1 and 2 one sample was attributed to class 6 (counterfeit samples containing sildenafil and tadalafil). One class 4 and one class 7 sample were considered to be genuine samples from class 1 and one placebo sample (class 7) was assigned to class 2. The remaining misclassifications occurred within the 4 classes of counterfeit samples. The external validation resulted in a correct classification rate of 78,57% which indicated that 9 test set samples were misclassified. One genuine Cialis[®] sample (class 2) was classified as a counterfeit from class 6; one placebo sample was considered to be a generic product of Viagra[®] (class 3) and one class 4 counterfeit was assigned to class 1 (genuine Viagra[®]). The other misclassifications occurred within the classes of counterfeit drugs.

3.2.3 SIMCA

Each class of samples is described separately by selecting a suitable number of principal components using leave-one-out cross validation. Seven PCs were retained to describe the genuine Viagra[®] class

(class 1); five PCs were selected for class 2 (genuine Cialis[®]); to model class 3 (generic products of Viagra[®]) 13 PCs were kept; 11 PCs were selected for class 4; ten PCs were retained for class 5; to describe class 6 nine PCs were used and 17 PCs were kept to model class 7. The obtained SIMCA model generated a correct classification rate of cross validation of 90,42% since only 16 out of 167 training set samples were misclassified. Both classes 1 and 2 were classified correctly. Five generic products (class 3) were classified as class 1. All remaining misclassifications occurred solely within the 4 classes of counterfeit samples, signifying that no counterfeit samples were classified as genuine or generic. The external validation confirmed these observations since the test set resulted in a correct classification rate of 90,48%. None of the genuine/generic samples were misclassified. All 4 misclassifications concerned counterfeit samples which were wrongly attributed to one of the 4 counterfeit classes. These results show that for both training and test set a 100% correct discrimination between genuine/generic and counterfeit samples was obtained.

Since only SIMCA was able to differentiate genuine/generic samples from counterfeit ones, it was verified whether this model could be improved even further by separating the data set according to the API present. The large data set, used thus far, was divided in 2 smaller data sets: (1) Viagra like data set and (2) Cialis like data set. For both data sets a training and test set were selected using the Duplexx algorithm.

3.3 Viagra like samples

Table 4 gives an overview of the classification of the Viagra like data set, along with the number of samples in each class in the entire data set, the training set and the test set.

Like the two previously created SIMCA models the number of principal components to describe each class was selected using leave-one-out cross validation. Classes 1, 2 and 3 were described by 11 PCs

each; ten PCs were selected for class 4 and 12 PCs were retained to model class 5. The acquired SIMCA model resulted in a correct classification rate of cross validation of 92,20%. Eleven training set samples were misclassified. Class 1 (genuine Viagra[®]) was classified correctly; 5 generic products (class 2) were classified as class 1; two samples of class 3 were wrongly classified (one as class 4 and the other as class 5); two samples of class 4 were assigned to an incorrect class (one as class 3 and the 2nd as class 5) and 2 placebo samples (class 5) were wrongly attributed to class 3. The external validation resulted in a correct classification rate of 94,44% since the test set only showed 2 misclassified samples: one class 3 sample which was grouped in class 4 and one sample of class 5 which was classified as belonging to class 3.

3.4 Cialis like samples

Table 5 gives an overview of the classification of the Cialis like data set, along with the number of samples in each class in the entire data set, the training set and the test set.

For each class separately, the number of principal components was selected using leave-one-out cross validation. Five PCs were selected for class 1; twelve PCs were kept to model class 2; class 3 was described using six PCs and for class 4 eight PCs were retained. As a result, a SIMCA model with a correct classification rate of cross validation of 92,10% was acquired. This model was characterized by 6 misclassified samples, all belonging to class 3 and misclassified as class 4 samples. The external validation gave similar results with a correct classification rate of 84,21%. Three samples were misclassified: two samples of class 2 were misclassified as class 4 and one class 3 sample was attributed to class 4. The genuine Cialis[®] samples (class 1) were classified correctly.

4. Discussion and conclusions

High amounts of counterfeit medicines enter the European market and detection of such pharmaceuticals at customs is not always simple due to misleading packaging. Therefore easy to use equipment and techniques to perform an initial screening of these samples are needed. For this purpose Attenuated Total Reflection Fourier-Transform infrared spectroscopy (ATR-FTIR) is an interesting technique since it is user-friendly and no sample preparation is needed.

This study proposes the use of ATR-FTIR and chemometrics to discriminate and classify counterfeit medicines. ATR-FTIR could be an interesting method for customs to obtain a first evaluation of suspected samples. For this study a total of 209 samples were analysed using ATR-FTIR. This sample set comprises different types of samples, e.g. genuine Viagra[®], genuine Cialis[®], generic products of Viagra[®], counterfeit samples containing sildenafil and/or tadalafil and placebo samples.

The exploratory analysis, using PCA, shows a clear tendency of discrimination based on the API present. This trend of discrimination was further explored in the supervised data analysis which resulted in 3 models differentiating between four groups of samples (Table 1). The results of the supervised analysis are summarized in Table 6. Not all models perform equally well; k-NN performs the least, CART results in a better model compared to k-NN but SIMCA surpasses both models and performs the best.

In overall, 3 suitable models are obtained. Therefore it was verified whether these 3 models could not only distinguish between different APIs but also discriminate genuine from counterfeit samples. This new classification system consists of seven classes and is presented in Table 3.

In this case, despite the relative good external validation (Table 6), k-NN is not able to discriminate between genuine and counterfeit samples and therefore this model is not useful.

For CART, conclusions are more complicated. It could be argued that the misclassification of genuine samples as counterfeit ones is unacceptable. However, every seized suspected medicine is kept from the market until a thorough analysis of the sample identifies its true nature. A medicine, which after

analysis turns out to be genuine, will be launched again. Such an approach is, from a public health point of view, much safer. A number of genuine samples which are misclassified as counterfeit pose much less risks to public health than counterfeit samples which are misclassified as genuine; therefore the latter is absolutely unacceptable. Following this reasoning, the extended CART model is not at all bad despite the large number of generic samples which are classified as counterfeit. However, a number of counterfeit samples are classified as genuine or generic, making this tree less suited.

SIMCA, on the other hand, generates a very suitable model which results in a model with a 100% correct discrimination between genuine and counterfeit drugs. No genuine samples (including the generic products of Viagra®) were classified as counterfeit and vice versa. Although 5 generic Viagra® samples are classified as genuine Viagra®, these misclassifications do not pose any problems since both genuine and generic medicines have to comply with the same quality standards. In addition, these 5 generics are also produced by Pfizer and their composition is very similar to that of genuine Viagra®.

The separation of the data set in 2 smaller data sets (Viagra and Cialis like samples) also generates two very suitable SIMCA models. Both models acquired a 100% correct discrimination between genuine and counterfeit medicines. When comparing the correct classification rates of external validation of the SIMCA models (Table 6) it is clear that the two models acquired after separation of the data set are not significantly better (or even worse in case of the Cialis like data set) than the former model obtained with the large data set. Therefore, all samples can be included in the chemometric analysis using one large data set with an adjusted classification system.

This study shows that chemometric analysis of ATR-FTIR fingerprints is a valuable tool to discriminate genuine from counterfeit samples and to classify counterfeit medicines. Furthermore, this approach might be helpful to customs to obtain a first evaluation of suspected samples. Based on the obtained results it can be stated that SIMCA generates the best predictive models for both

classification systems. CART and k-NN are only useful for the discrimination based on the APIs present. However, it should be noticed that when working with other groups of counterfeit drugs, e.g. slimming products, perhaps other chemometric techniques might need to be applied. Therefore, each group of counterfeit medicines has to be regarded separately and the best methods have to be explored for each individual group.

Reference List

- [1] International Medical Products Anti-Counterfeiting Taskforce (IMPACT). Counterfeits drugs kill. May 2008. (<http://www.who.int/impact/FinalBrochureWHA2008a.pdf>) (accessed 09.07.14)
- [2] World Health Organization (WHO). General information on counterfeit medicines. (<http://www.who.int/medicines/services/counterfeit/overview/en/>) (accessed 09.07.14)
- [3] P. Y. Sacre, E. Deconinck, M. Daszykowski, P. Courselle, R. Vancauwenberghe, P. Chiap, J. Crommen, J. O. De Beer, Impurity fingerprints for the identification of counterfeit medicines – a feasibility study, *Anal. Chim. Acta* 701 (2011) 224-231.
- [4] G. Jackson, S. Patel, S. Khan, Assessing the problem of counterfeit medications in the United Kingdom, *Int. J. Clin. Pract.* 66 (2012) 241-250.
- [5] European Alliance for Access to Safe Medicines (EAASM). The Counterfeiting Superhighway. 2008. ([http://www.eaasm.eu/cache/downloads/dqqt3sge9hwssgcgc440g40/455_EAASM_counterfeiting%20report_020608\(1\).pdf](http://www.eaasm.eu/cache/downloads/dqqt3sge9hwssgcgc440g40/455_EAASM_counterfeiting%20report_020608(1).pdf)) (accessed 09.07.14)
- [6] E. Deconinck, M. Canfyn, P. Y. Sacre, P. Courselle, J. O. De Beer, Evaluation of the residual solvent content of counterfeit tablets and capsules, *J. Pharm. Biomed. Anal.* 81-82 (2013) 80-88.
- [7] E. Deconinck, M. Canfyn, P. Y. Sacre, S. Baudewyns, P. Courselle, J. O. De Beer, A validated GC-MS method for the determination and quantification of residual solvents in counterfeit tablets and capsules, *J. Pharm. Biomed. Anal.* 70 (2012) 64-70.
- [8] E. Deconinck, P. Y. Sacre, P. Courselle, J. O. De Beer, Chemometrics and chromatographic fingerprints to discriminate and classify counterfeit medicines containing PDE-5 inhibitors, *Talanta* 100 (2012) 123-133.
- [9] F. R. da Silva, F. S. da Costa, P. Valderrama, P. H. Marco, K. M. de Lima, Non-destructive detection of adulterated tablets of glibenclamide using NIR and solid-phase fluorescence spectroscopy and chemometric methods, *J. Pharm. Biomed. Anal.* 66 (2012) 85-90.
- [10] I. Storme-Paris, H. Rebiere, M. Matoga, C. Civade, P. A. Bonnet, M. H. Tissier, P. Chaminade, Challenging near infrared spectroscopy discriminating ability for counterfeit pharmaceuticals detection, *Anal. Chim. Acta* 658 (2010) 163-174.
- [11] F. E. Dowell, E. B. Maghirang, F. M. Fernandez, P. N. Newton, M. D. Green, Detecting counterfeit antimalarial tablets by near-infrared spectroscopy, *J. Pharm. Biomed. Anal.* 48 (2008) 1011-1014.
- [12] M. J. Vredenbregt, L. Blok-Tip, R. Hoogerbrugge, D. M. Barends, D. de Kaste, Screening suspected counterfeit Viagra and imitations of Viagra with near-infrared spectroscopy, *J. Pharm. Biomed. Anal.* 40 (2006) 840-849.
- [13] F. Been, Y. Roggo, K. Degardin, P. Esseiva, P. Margot, Profiling of counterfeit medicines by vibrational spectroscopy, *Forensic Sci. Int.* 211 (2011) 83-100.

- [14] B. J. Venhuis, G. Zomer, M. Hamzink, H. D. Meiring, Y. Aubin, D. de Kaste, The identification of a nitrosated prodrug of the PDE-5 inhibitor aildenafil in a dietary supplement: a Viagra with a pop, *J. Pharm. Biomed. Anal.* 54 (2011) 735-741.
- [15] B. J. Venhuis, G. Zomer, D. de Kaste, Structure elucidation of a novel synthetic thiono analogue of sildenafil detected in an alleged herbal aphrodisiac, *J. Pharm. Biomed. Anal.* 46 (2008) 814-817.
- [16] S. Trefi, C. Routaboul, S. Hamieh, V. Gilard, M. Malet-Martino, R. Martino, Analysis of illegally manufactured formulations of tadalafil (Cialis) by ¹H NMR, 2D DOSY ¹H NMR and Raman spectroscopy, *J. Pharm. Biomed. Anal.* 47 (2008) 103-113.
- [17] M. de Veij, A. Deneckere, P. Vandenabeele, D. de Kaste, L. Moens, Detection of counterfeit Viagra with Raman spectroscopy, *J. Pharm. Biomed. Anal.* 46 (2008) 303-309.
- [18] I. Wawer, M. Pisklak, Z. Chilmonczyk, ¹H, ¹³C, ¹⁵N NMR analysis of sildenafil base and citrate (Viagra) in solution, solid state and pharmaceutical dosage forms, *J. Pharm. Biomed. Anal.* 38 (2005) 865-870.
- [19] I. McEwen, A. Elmsjö, A. Lehnström, B. Hakkarainen, M. Johansson, Screening of counterfeit corticosteroid in creams and ointments by NMR spectroscopy, *J. Pharm. Biomed. Anal.* 70 (2012) 245-250.
- [20] J. K. Maurin, F. Plucinski, A. P. Mazurek, Z. Fijalek, The usefulness of simple X-ray powder diffraction analysis for counterfeit control - the Viagra example, *J. Pharm. Biomed. Anal.* 43 (2007) 1514-1518.
- [21] P. Y. Sacre, E. Deconinck, De Beer T., P. Courselle, R. Vancauwenberghe, P. Chiap, J. Crommen, J. O. De Beer, Comparison and combination of spectroscopic techniques for the detection of counterfeit medicines, *J. Pharm. Biomed. Anal.* 53 (2010) 445-453.
- [22] R. S. Ortiz, K. C. Mariotti, B. Fank, R. P. Limberger, M. J. Anzanello, P. Mayorga, Counterfeit Cialis and Viagra fingerprinting by ATR-FTIR spectroscopy with chemometry: can the same pharmaceutical powder mixture be used to falsify two medicines?, *Forensic Sci. Int.* 226 (2013) 282-289.
- [23] A. B. Champagne, Emmel K.V., Rapid screening test for adulteration in raw materials of dietary supplements, *Vib. Spectrosc.* 55 (2011) 216-223.
- [24] European Directorate for the Quality of Medicines and HealthCare (EDQM). Qualification of equipment, Annex 4: Qualification of IR spectrophotometers (PA/PH/OMCL (07) 12 DEF CORR). 2007.
- [25] 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, Amsterdam, 1997.
- [26] I. Stanimirova, A. Kubik, B. Walczak, J. W. Einax, Discrimination of biofilm samples using pattern recognition techniques, *Anal. Bioanal. Chem.* 390 (2008) 1273-1282.
- [27] B. G. M. Vandeginste, D. L. Massart, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics - Part B*, Elsevier Science, Amsterdam, 1998.

- [28] E. Deconinck, T. Hancock, D. Coomans, D. L. Massart, Y. V. Heyden, Classification of drugs in absorption classes using the classification and regression trees (CART) methodology, *J. Pharm. Biomed. Anal.* 39 (2005) 91-103.
- [29] E. Deconinck, M. H. Zhang, D. Coomans, Y. Vander Heyden, Classification tree models for the prediction of blood-brain barrier passage of drugs, *J. Chem. Inf. Model.* 46 (2006) 1410-1419.
- [30] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, *Classification and Regression Trees*, Wadsworth & Brooks, Monterey, 1984.

CAPTIONS

Fig. 1. Examples of ATR-FTIR spectra: (A) genuine Viagra[®] sample, (B) counterfeit sample containing sildenafil, (C) genuine Cialis[®] sample, (D) counterfeit sample containing tadalafil, (E) counterfeit sample containing both sildenafil and tadalafil, (F) placebo sample

Fig. 2a. ATR-FTIR spectrum of a sildenafil reference.

Fig. 2b. ATR-FTIR spectrum of a tadalafil reference.

Fig. 3. General structure of a classification tree obtained by CART, x_i is the selected descriptor and a_i is the selected split value [28].

Fig. 4. Score plot obtained with Principal Component Analysis.

Fig. 5. Classification tree obtained using the Gini index as split criterion. Each split is described by the selected wavenumber and its split value for absorbance. Each leaf is defined by the number of the class which is highest represented in the respective leaf. Each graph represents the number of training set samples in each leaf in function of the classes they belong to.

Fig. 6. Classification tree obtained after extending the classification system using the Gini index as split criterion. Each split is described by the selected wavenumber and its split value for absorbance.

Each leaf is defined by the number of the class which is highest represented in the respective leaf.

Each graph represents the number of training set samples in each leaf in function of the classes they belong to.

Fig. 1

Fig. 2a

Fig. 2b

Fig. 3

Fig. 4

Fig. 5

Table 1

Overview of the classification of samples based on the API present.

Class number	Type of samples	Total no. of samples per class	No. of samples in training set	No. of samples in test set
1	samples containing sildenafil	114	92	22
2	samples containing tadalafil	32	26	6
3	samples containing both sildenafil & tadalafil	18	16	2
4	placebo	45	33	12

Table 2

Overview of the unclassified and misclassified samples (with the class to which they were assigned) of each class for both training and test set by k-NN.

	Misclassified	Unclassified
Training set		
Class 1	1 sample (class 2) 2 samples (class 3) 4 samples (class 4)	2 samples
Class 2	2 samples (class 1) 2 samples (class 3) 3 samples (class 4)	/
Class 3	8 samples (class 1) 1 sample (class 2)	2 samples
Class 4	8 samples (class 1) 2 samples (class 2) 2 samples (class 3)	1 sample
Test set		
Class 1	1 sample (class 2) 2 samples (class 4)	1 sample
Class 2	1 sample (class 4)	1 sample
Class 4	2 samples (class 1)	/

Table 3

Overview of the new classification system to extend the obtained models in order to discriminate between genuine and counterfeit samples and to distinguish between different APIs.

Class number	Type of samples	Total no. of samples per class	No. of samples in training set	No. of samples in test set
1	Viagra [®] (genuine)	13	12	1
2	Cialis [®] (genuine)	11	9	2
3	generic products of Viagra [®]	33	31	2
4	counterfeit samples containing sildenafil	68	49	19
5	counterfeit samples containing tadalafil	21	17	4
6	counterfeit samples containing sildenafil & tadalafil	18	16	2
7	Placebo	45	33	12

Table 4

Overview of the classification of the Viagra like data set.

Class number	Type of samples	Total no. of samples per class	No. of samples in training set	No. of samples in test set
1	Viagra [®] (genuine)	13	12	1
2	generic products of Viagra [®]	33	30	3
3	counterfeit samples containing sildenafil	68	51	17
4	counterfeit samples containing sildenafil & tadalafil	18	16	2
5	placebo	45	32	13
Total		177	141	36

Table 5

Overview of the classification of the Cialis like data set.

Class number	Type of samples	Total no. of samples per class	No. of samples in training set	No. of samples in test set
1	Cialis [®] (genuine)	11	9	2
2	counterfeit samples containing tadalafil	21	18	3
3	counterfeit samples containing sildenafil & tadalafil	18	15	3
4	placebo	45	34	11
Total		95	76	19

Table 6

Overview of the performance of the different models for the external test set expressed as correct classification rate.

Technique	Discrimination on API (4 groups)	Discrimination genuine/counterfeit and API (7 groups)	Viagra like samples	Cialis like samples
k-NN	80.9%	76.2%	/	/
CART	83.3%	78.6%	/	/
SIMCA	90.5%	90.5%	94.4%	84.2%