Talanta 123 (2014) 78–88

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Headspace–gas chromatographic fingerprints to discriminate and classify counterfeit medicines

D. Custers [a,b], M. Canfyn [a], P. Courselle [a], J.O. De Beer [a], S. Apers [b], E. Deconinck [a,*]

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

[b] Laboratory of Pharmacognosy and Pharmaceutical Analysis, Department of Pharmaceutical Sciences, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium

ARTICLE INFO

Article history:
Received 31 October 2013
Received in revised form
15 January 2014
Accepted 20 January 2014
Available online 31 January 2014

Keywords:
Counterfeit medicines
Residual solvents
Chemometrics
Fingerprints

ABSTRACT

Counterfeit medicines are a global threat to public health. These pharmaceuticals are not subjected to quality control and therefore their safety, quality and efficacy cannot be guaranteed. Today, the safety evaluation of counterfeit medicines is mainly based on the identification and quantification of the active substances present. However, the analysis of potential toxic secondary components, like residual solvents, becomes more important. Assessment of residual solvent content and chemometric analysis of fingerprints might be useful in the discrimination between genuine and counterfeit pharmaceuticals. Moreover, the fingerprint approach might also contribute in the evaluation of the health risks different types of counterfeit medicines pose. In this study a number of genuine and counterfeit Viagra® and Cialis® samples were analyzed for residual solvent content using headspace–GC–MS. The obtained chromatograms were used as fingerprints and analyzed using different chemometric techniques: Principal Component Analysis, Projection Pursuit, Classification and Regression Trees and Soft Independent Modelling of Class Analogy. It was tested whether these techniques can distinguish genuine pharmaceuticals from counterfeit ones and if distinct types of counterfeits could be differentiated based on health risks. This chemometric analysis showed that for both data sets PCA clearly discriminated between genuine and counterfeit drugs, and SIMCA generated the best predictive models. This technique not only resulted in a 100% correct classification rate for the discrimination between genuine and counterfeit medicines, the classification of the counterfeit samples was also superior compared to CART. This study shows that chemometric analysis of headspace–GC impurity fingerprints allows to distinguish between genuine and counterfeit medicines and to differentiate between groups of counterfeit products based on the public health risks they pose.

1. Introduction

Counterfeit medicines pose a huge threat to public health worldwide [1]. Not only developing countries are threatened, also 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 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 active ingredients, with insufficient active ingredients or with fake packaging" [2].

These forged medicines are mostly 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 presence of incorrect active ingredients, the absence of 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]. Moreover the size of the problem differs from region to region. It is estimated that about 1% of the total medicines market of industrialized countries, such as the United States, European countries, Japan, etc., consists of counterfeit medicines. In countries of the former Soviet Union about 20% of the medicines market is covered by counterfeit pharmaceuticals. This number reaches even more than 30% in African countries and parts of Asia and Latin-America. Furthermore, it is also estimated that approximately 50% of all medicines, bought online from websites which cover up their

* Corresponding author. Tel.: +32 2 642 51 36; fax: +32 2 642 53 27.
E-mail address: Eric.Deconinck@wiv-isp.be (E. Deconinck).

0039-9140/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2014.01.020

physical address, are fake [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 most sold as counterfeit in industrialized countries are commonly referred to as 'life style drugs' and comprise phosphodiesterase type 5 (PDE-5) inhibitors, slimming products (containing anorexics) and anabolic hormones [6,7].

In most literature, the characterization of counterfeit medicines is based on the identification and quantification of the active substances present. Indeed, potential toxic secondary components, such as impurities, residual solvents, etc., are often not taken into account. As a result, a product can be regarded as relatively save for it might contain the right active substances in the correct dosage, while in actual fact high concentrations of potential toxic secondary components could be present. Since counterfeiters probably use inferior primary substances and manufacture these medicines without respecting any quality norm, the analysis of these secondary components becomes more important. The evaluation of residual solvents is fundamental for quality control of genuine medicines, especially for medicines intended for chronic use. Consequently, residual solvents are of great interest for the characterization of counterfeit medicines [7].

The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) defines residual solvents as "organic volatile chemicals that are used or produced in the manufacture of drug substances or excipients, or in the preparation of drug products" [8]. Many of these solvents are known to be harmful to humans or the environment [9]. Furthermore, these chemicals have no therapeutic benefit and they may facilitate decomposition of pharmaceuticals [10]. Since it is not possible to completely remove residual solvents from drug substances and excipients, it is important that these impurities are eliminated to the extent possible in order to meet quality norms. ICH issued guidelines which not only recommend the use of less toxic solvents; they also recommend acceptable amounts for residual solvents in order to ensure the patient's safety [8]. These guidelines have been adopted by the European Pharmacopoeia, the United States Pharmacopoeia and the Japanese Pharmacopoeia [6].

ICH defines 4 classes of residual solvents. Class I solvents (e.g., benzene, 1,2-dichloroethane, etc.) are to be avoided because of their high toxicity or harmful environmental effect. Concentration limits vary between 2 and 8 ppm. 1,1,1-trichloroethane is classified as class I solvent because of its environmental hazard. Its concentration limit is set at 1500 ppm. Class II (e.g., acetonitril, methanol, toluene, etc.) consists of solvents which should be limited due to low toxicity [8]. Their limits range between 50 and 5000 ppm [7]. Class III solvents (e.g., ethanol, acetic acid, acetone, etc.) are considered to have low toxic potential and they are limited to 5000 ppm. Class IV (e.g., isopropyl ether, trifluoroacetic acid, etc.) is composed of solvents for which no adequate toxicological data are available [8].

The increasing interest in residual solvent assessment has led to the development of a large number of analytical techniques intended for the determination of these chemicals [9]. In general, most of these techniques are based on gas chromatography (GC) [7]. The European Pharmacopoeia mentions two gas chromatographic methods using static headspace injection and a flame ionisation detector. A mass spectrometer or, if needed, an electron-capture detector for the determination of chlorinated residual solvents may also be used. These two methods allow: (1) the identification of class I, II and III solvents; (2) to carry out a limit test for class I and II solvents and (3) to quantify class II solvents, if the limits are higher than 1000 ppm, and class III solvents [11]. Besides these two techniques, several other GC methods are described in literature using different injection techniques, such as split/splitless injection, headspace and solid-phase micoextraction [9,12–15]. Other techniques for residual solvent determination, used as alternatives to gas chromatography, are loss on drying, thermogravimetric analysis, differential scanning calorimetry, IR spectroscopy and NMR spectrometry. Many of these techniques have the disadvantage of being non-specific or they are characterized by high detection limits, making them often less suitable for residual solvents assessment [16]. Both groups of techniques, gas chromatography and alternatives, are reviewed by B'Hymer [16] and Grodowska et al. [17]. Even though many different analytical methods are available, gas chromatography remains the most powerful technique for residual solvent analysis [17]. The combination of headspace injection with GC–MS has also the advantage of a limited sample preparation effort, allowing fast analysis. Our group developed and validated its own GC technique for the identification and quantification of residual solvents [7]. This technique has the advantages of being fast and suitable for routine analysis of pharmaceuticals.

Despite the fact that GC is the most suited technique for residual solvent analysis, the use of GC impurity fingerprints is a fairly new concept in literature. The fingerprint approach is already extensively used in the field of Pharmacognosy for the identification and quality control of plants. This approach might be interesting for the identification of potential toxic secondary components in counterfeit medicines. A fingerprint is a characteristic profile which visualizes the composition of a sample. It can be obtained by usage of chromatographic, spectroscopic or electrophoretic techniques. However, chromatographic fingerprints are the most interesting fingerprints. By spreading information about the composition of a sample over time, they provide information about individual compounds [18].

In this paper the chromatograms, obtained by the headspace–GC–MS analysis of a set of genuine and counterfeit Viagra® and Cialis® samples, were used as fingerprints. These fingerprints were analyzed using different chemometric techniques. The purpose of this data-analysis was to test whether these techniques allow for the distinction between genuine and counterfeit medicines, based on the obtained fingerprints. Furthermore, it was tested if these methods can also discriminate between different counterfeit medicines based on the public health risk they pose.

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® were kindly provided by Pfizer SA/NV (Belgium). Eli Lilly SA/NV (Benelux) kindly provided genuine samples of Cialis®.

2.2. Chemicals and reagents

2-Propanol, dichloromethane, acetone, ethanol absolute, acetonitril (all HPLC grade) and ethylacetate (pesti-S) were purchased from Biosolve (Valkenswaard, The Netherlands). Chloroform (for gas chromatography), benzene, tetrachloromethane (CCl_4) (for spectroscopy) and ethylbenzene (for gas chromatography) were purchased from Merck (Darmstadt, Germany). Toluene and cyclohexane were purchased from VWR prolabo (Fontenay-Sous-Bois, France). These solvents were used as reference standards. Dimethyl sulfoxide, which was used as solvent for the samples, was purchased from Merck.

2.3. Instrumental conditions

The samples were injected on a GC–MS system using a G1888 headspace sampler (Agilent Technologies, Palo Alto, USA). The analyses were performed on an Agilent 6890N gas chromatograph coupled to an Agilent 5973N mass spectrometer. By application of Agilent MSD ChemStation data acquisition and data handling software full automation was achieved.

The samples were incubated in a 10 ml headspace vial and shaken at 120 °C for 15 min. Next, 1 ml of the vapor phase was injected into the GC–MS system in a split injection mode (split ration 6.8:1). The temperatures of the headspace loop, the transfer line and the EPC volatiles interface were set at 135, 145 and 160 °C, respectively. The solvents were separated on a Phenomenex 624 capillary column (60 m × 0.32 mm; 1.8 μm film thickness) (Phenomenex, Torrance, USA). The oven temperature was programmed from 60 °C, which was held for 5 min, to 270 °C at 25 °C/min. 270 °C was held for 10 min, making a total runtime of 23.4 min. The temperatures of the injection port, the ion source, the quadrupole and the interface were set at 160, 230, 150 and 280 °C, respectively. The mass spectrometer was set at full scan mode for the identification of the solvents present in the samples. For quantification and validation, the mass spectrometer was operated in single ion monitoring mode (100 ms dwell times) [7]. The chromatograms, obtained in full scan mode, were subsequently used as fingerprints in the chemometric data-analysis. For more information about the validation of this GC-technique, the reader is referred to reference [7].

2.4. Sample preparations

2.4.1. Preparation of internal standard solution

Two stock solutions in dimethyl sulfoxide were prepared; one stock solution containing 1000 ppm acetonitril, the second containing 1000 ppm cyclohexane. 1 ml from each solution was diluted to 100.0 ml with dimethyl sulfoxide and used as internal standard. Cyclohexane was used for the quantification of tetrachloromethane, benzene, toluene and ethylbenzene. Acetonitril served as internal standard for the remaining solvents. These internal standards were selected, based on an initial screening of counterfeit samples, due to the fact that both solvents were not detected in the screened samples or only as traces [7].

For more information about the preparation of other standard solutions, the reader is referred to reference [7].

2.4.2. Preparation of samples

Tablets were broken in two before addition of 5 ml of dimethyl sulfoxide. Breakage of tablets was necessary because of the coating of some tablets, which might prevent the recovery of residual solvents. Capsules were opened before adding the solvent. 500 μl of the internal standard was added to these solutions [7].

2.5. Sample sets

Two sample sets, one Viagra® and one Cialis® sample set, were analyzed using the described GC-method and used for the chemometric data-analysis. The Viagra® sample set contained 5 genuine Viagra® samples and 31 counterfeit samples. For the Cialis® sample set 5 genuine and 35 counterfeit samples were analyzed, making a total of 36 samples for the Viagra® sample set and 40 samples for the Cialis® sample set.

2.6. Chemometric methods

2.6.1. Data-preprocessing

All fingerprints were cut at the beginning (at 3.4 min) and at the end (at 13 min) since the fingerprints did not contain any useful information before 3.4 and after 13 min. After cutting, the fingerprints were normalized to obtain the same scaling for the abundances at different time points.

2.6.2. Exploratory techniques

Two unsupervised chemometric techniques, i.e., Principal Component Analysis (PCA) and Projection Pursuit (PP), were applied to test whether these techniques can discriminate between genuine and counterfeit medicines. At the same time, it was explored if the obtained clustering of the samples provides a foundation to create classification models using supervised techniques. Both techniques were tested on the Viagra® and Cialis® data set.

2.6.2.1. Principal Component Analysis. Principal Component Analysis is a projection method that allows to project high dimensional original variables into a low dimensional space. This low dimensional space is defined by new orthogonal latent variables, commonly referred to as principal components. The result of such an analysis is a reduction in the number of variables by calculating linear combinations (=principal components) of these variables. The first constructed principal component explains the highest variance in the data. The loadings of the original variables are a measure for the contribution of each variable in the construction of a given principal component. Each object (=sample) is then projected on the created principal components. These projections are referred to as scores and provide information about (dis)similarities among the objects [19].

2.6.2.2. Projection Pursuit. Projection Pursuit is a projection method as well. This technique also creates a low dimensional space, defined by latent variables, in which high dimensional data are projected. With PP, these latent variables are called projection pursuit features (PPFs). The main difference with PCA is the way in which these latent variables are defined. PPFs are constructed by maximizing a projection index which is a measure for the inhomogeneity of the data [19]. In this study two projection indices were tested, i.e., entropy and yenyukov.

2.6.3. Selection of a test set for external validation of models

The chromatographic fingerprints of both the Viagra® and the Cialis® data set were split into a training set and a test set. The training set is used to generate classification models; the test set is selected to perform an external validation of the created classification models. Two algorithms can be applied for the selection of the test set: (1) the Kennard and Stone algorithm and (2) the Duplexx algorithm [20]. Both algorithms were tested on the Viagra® and Cialis® data set. The algorithm which resulted in the best test and training set was selected.

The Kennard and Stone algorithm starts by selecting the sample which is situated closest to or farthest from the data mean. This first selected sample is assigned to the training set. The second sample to be selected is the one situated furthest away from the first sample and is also assigned to the training set. The third sample to be selected and assigned to the training set is the one most remote from the two earlier selected samples. This procedure is repeated until a predefined number of samples is not allocated to the training set. The test set is composed of these non-assigned samples.

When applying the Duplexx algorithm, the two most distant samples are assigned to the training set. The second pair of most

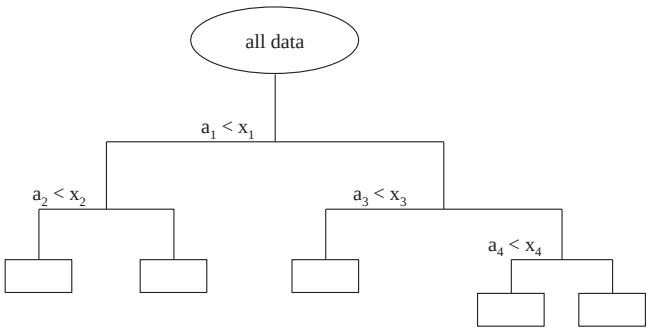

Fig. 1. General structure of a tree obtained by CART, x_i is the selected descriptor and a_i is the selected split value [22].

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 [20].

The Kennard and Stone algorithm has the advantage of generating a training set which covers all possible sources of data variance. The Duplexx algorithm, on the other hand, warrants both the training and test set to be representative.

2.6.4. Modelling techniques

2.6.4.1. 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. 1), used to solve classification problems, if the dependent variable is categorical. When the dependent variable is continuous, it creates a regression tree which is useful to solve regression problems [21,22].

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 which result in the highest decrease in impurity between the mother group and daughter groups is chosen. Decrease in impurity means that all samples in a daughter group become more homogeneous in the response variable values. 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 splitting 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 [21,22].

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. Cross-validation (CV) is often used to evaluate the predictive error. In this study a 10-fold CV was used [19,21,22].

2.6.4.2. Soft Independent Modelling by Class Analogy (SIMCA). SIMCA is a supervised classification technique that models each class of samples separately by defining a number of principal components. These principal components are derived from PCA. SIMCA 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 scores. These two critical values determine a restricted space around the samples of the training set. 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 [19].

2.6.5. Calculations

All data treatments were performed using Matlab version 8.0.0 (The Mathworks, Natick, MA, USA). The algorithms PCA, PP and CART were part of the ChemoAC toolbox (Freeware, ChemoAC Consortium, Brussels, Belgium, version 4.0). 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

The obtained abundances (MS) for the components present were used as explanatory variables and the class numbers, based on the presence and concentration of residual solvents, as response variables for the supervised techniques. Exemplary chromatograms, obtained in full scan mode for genuine and counterfeit samples of both Viagra® and Cialis®, are shown in Fig. 2.

3.1. Exploratory data analysis

Exploratory data analysis of the Viagra® data set showed that PCA gave the best results. It was chosen to limit the number of PCs to 3 since the total variance explained was 98.6% (PC1=97.7%, PC2=0.6% and PC3=0.3%). The resulting score plot of the PCA is shown in Fig. 3a. A clear distinction could be observed between the genuine and counterfeit samples. Study of the loading plot (figure not shown) did not reveal any time points which could account significantly for this discrimination, thereby showing that the entire fingerprints account for the discrimination. Despite the fact that PCA resulted in a manifest separation between genuine and counterfeit, the counterfeit samples were clustered together. Therefore no pattern could be observed.

The exploratory data analysis was also performed on the Cialis® data set. For this set of data, PCA yielded the best results as well. Similar to the Viagra® data set 3 PCs were retained, explaining 98.5% of the total variance (PC1=97.5%, PC2=0.6% and PC3=0.4%). The score plot (Fig. 3b) shows that the genuine samples were isolated from the counterfeit samples, indicating that PCA clearly distinguished between genuine samples and counterfeit ones. In contrast to the Viagra® data set, PCA gave slight clustering for the Cialis® data. Unfortunately, this clustering did not provide any useful foundation for the creation of classification models. No time points were found to be of significance for

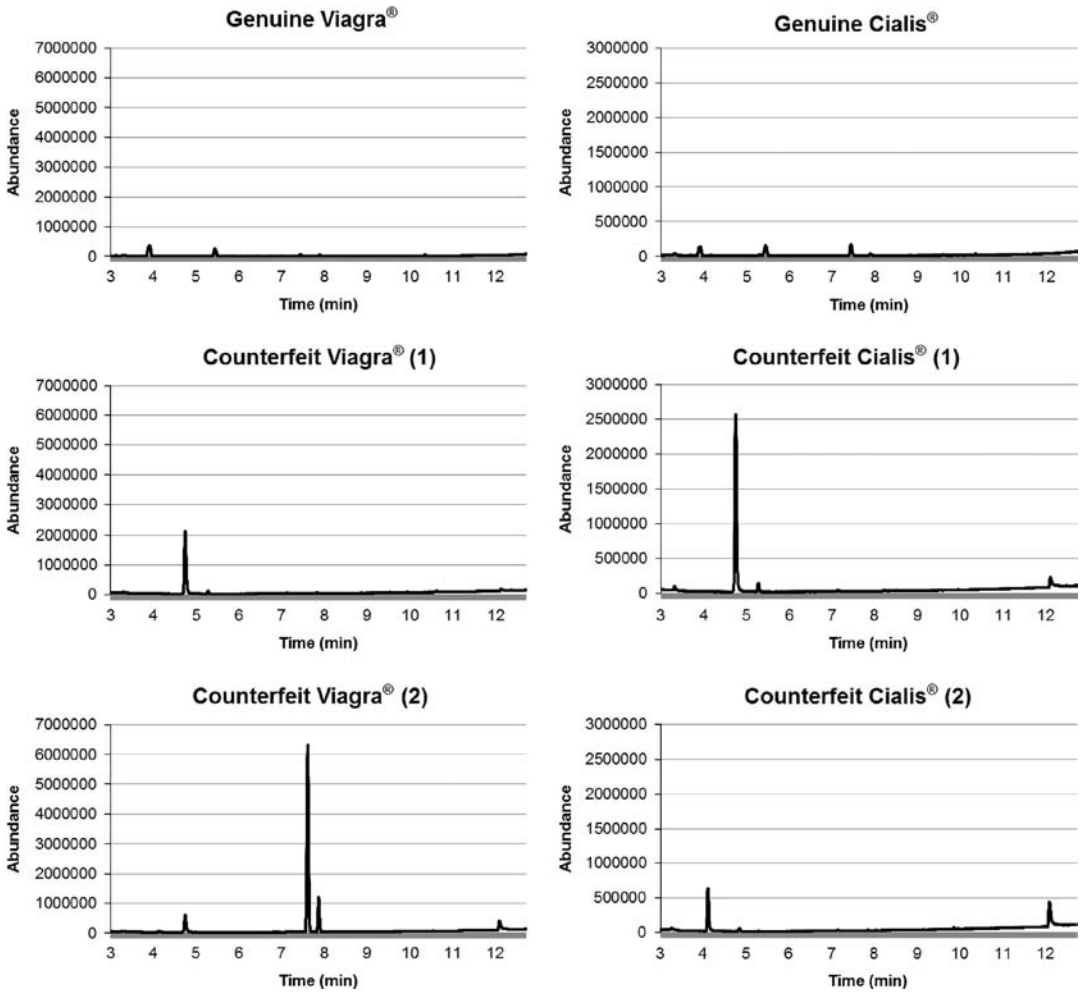

Fig. 2. Examples of full scan chromatograms obtained for genuine and counterfeit Viagra® and Cialis® samples.

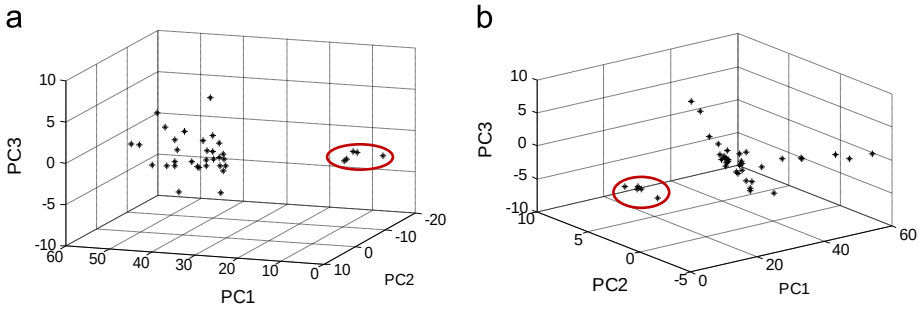

Fig. 3. (a) Score plot obtained with Principal Component Analysis for the Viagra® data set. The genuine samples are indicated by the circle. (b) Score plot obtained with Principal Component Analysis for the Cialis® data set. The circle indicates the genuine samples.

the discrimination between genuine and counterfeit samples (loading plot not shown). Therefore the overall fingerprints are taken into account for discrimination.

For both data sets, the results obtained by Projection Pursuit were less optimal (figures not shown). In most cases a distinction could be made between genuine and counterfeit samples, except for the PP analysis of the Cialis® data set using yenyukov as projection index. However, the obtained discrimination between counterfeit and genuine samples was less clear than the distinction obtained using PCA. Also, no clear clustering was acquired using Projection Pursuit.

3.2. Selection training and test set

Before any modelling technique can be applied, a test and training set have to be selected. Two algorithms were tested: (1) Kennard and Stone and (2) Duplexx. 20% of all data were assigned to the test set, resulting in a test set of 7 samples for the

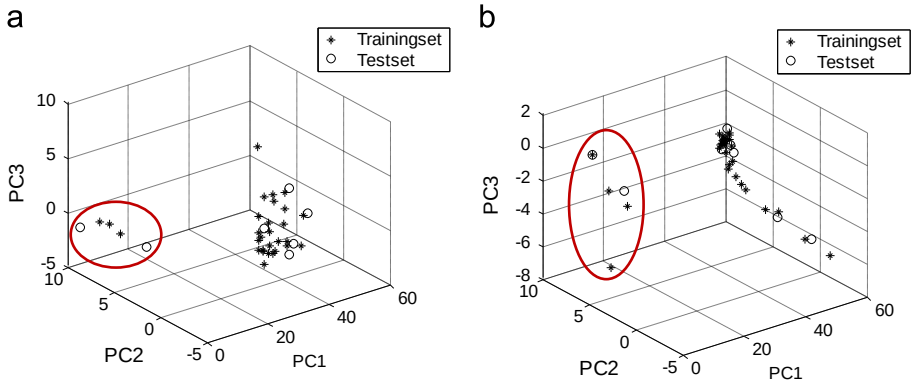

Fig. 4. (a) Score plot of the PCA, performed after selection of the test set for the Viagra® data. The acquired test set shows a good spreading over the entire data set and contains 2 genuine and 5 counterfeit samples. The genuine samples are indicated by the circle. (b) Score plot of the PCA, performed after selection of the test set for the Cialis® data. The obtained test set shows good spreading over the entire data set and contains 2 genuine and 6 counterfeit samples. The genuine samples are indicated by the circle.

Table 1
Overview of the considered residual solvents, along with ICH class, limit of content and limit of quantification [6,8].

Residual solvent	ICH class	Limit of content (ppm)	LOQ (ppm)
Ethanol	3	5000	0.114
2-Propanol	3	5000	0.001
Acetone	3	5000	0.001
Ethylacetate	3	5000	0.421
Chloroform	2	60	0.002
Tetrachloromethane	1	4	0.001
Dichloromethane	2	600	0.002

Viagra® data set and a test set containing 8 samples for the Cialis® set. After selection of the test set a PCA was performed to make sure the test set is representative for the overall data set.

For the Viagra® data set, the most appropriate test set was acquired using the Kennard and Stone algorithm, with the first selected sample that one being situated closest to the data mean (Fig. 4a).

For the Cialis® data set, the Duplexx algorithm resulted in the best separation in test and training set (Fig. 4b).

3.3. Creation of classes

Since PCA and PP did not provide any clustering which could serve as a base for the modelling techniques, the samples had to be assigned to classes in an arbitrary way. The creation of these arbitrary classes was based on the presence and content of 7 residual solvents which were detected in the samples above the quantification limits (LOQ) of the described GC–MS method. Table 1 gives an overview of the considered residual solvents along with their ICH class, limits of content and estimated limits of quantification. Tables 2 and 3 give an overview of the residual solvent content of the screened Viagra® and Cialis® samples.

Screening of genuine Viagra® samples with the described GC–MS method indicated the presence of toluene (ICH class II), ethylacetate and 2-butanone (both ICH class III). The content of these solvents was smaller than the limits of quantification of the method and therefore well below the limits set by ICH. Genuine samples of Cialis® were also analyzed. The method demonstrated the presence of only tetrahydrofuran (ICH class II). Similarly to the Viagra® genuines, the residual solvent is below the limit of quantification and hence below the ICH limits [6].

An arbitrary classification system was set up for both sample sets by assigning all samples containing only ICH class III solvents to one group. Samples containing class II solvents constitute a second group and samples containing class I solvents were assigned to a third group. Genuine samples constitute another separate group. Tables 2 and 3 show that the majority of the counterfeit samples in both sample sets contain class III solvents only. As both sample sets are quite small, the created groups have to contain more or less an equal amount of samples in order to obtain classification models with good predictive properties. Therefore the group of samples, containing only class III solvents, was split up in two. A residual solvent content of 100 ppm is considered acceptable for qualitative products. Consequently, it was verified whether this limit could serve as a threshold to split this large group of samples. Indeed, survey of the screening results revealed that 2 more or less equivalent groups could be obtained by assigning samples containing class III solvents in a total amount lower than 100 ppm (regardless if the samples contained one or more class III solvents) to one group. The other samples, containing more than 100 ppm of class III solvents, were put in another group. Table 4 gives an overview of the resulting classification.

3.4. CART

A classification tree was built for both data sets, using the Gini index. First the maximal tree was built and pruned. Then the optimal tree was selected using a 10-fold cross validation.

3.4.1. Viagra® like samples

The graph representing the cross validation error in function of tree complexity of the obtained trees showed that the tree with complexity 2 (meaning the tree contains 2 leaves) should be selected as the optimal tree (Fig. 5a). This tree had a cross validation error of 0.45. Study of the leaves revealed that all genuine samples were classified as counterfeits, making this tree an inappropriate model to describe the Viagra® data set. However, the cross validation graph also showed that a tree with complexity 5 (Fig. 5a) could be built which is characterized by the same cross validation error as the two-leaf tree. In contrast to the two-leaf tree, the five-leaf tree showed good homogeneous leaves (shown by the graphs in Fig. 5a). Four out of five leaves show complete homogeneity, indicating that no sample was misclassified. However, the leaf representing the genuine samples does not show complete homogeneity, since a counterfeit sample is classified as genuine. All genuine samples were classified as genuine. 25 out of 26 counterfeit samples were classified correctly. The misclassified counterfeit sample is a sample belonging to class 3, which was classified as genuine. This particular sample contains 2-propanol and acetone (both ICH class III solvents) in a total amount of 2833 ppm which is

Table 2
Results for the screening of 31 counterfeit samples of Viagra® [6].

Sample no.	Ethanol (ppm)	2-Propanol (ppm)	Acetone (ppm)	Ethylacetate (ppm)	Tetrachloromethane (ppm)	Dichloormethane (ppm)
1	188.5	2.0	3.4	1.8	–	–
2	–	414.1	18.9	–	–	5.6
3	–	17.5	3.3	–	–	–
4	–	184.9	20.5	–	–	–
5	–	926.5	16.3	–	–	4.8
6	–	2826.4	6.8	–	–	–
7	–	166.0	19.8	–	–	3.5
8	–	16.5	84.8	–	–	–
9	–	128.6	336.5	1.9	–	–
10	–	12.0	35.6	–	–	–
11	–	17.9	55.0	–	–	–
12	–	–	16.1	–	–	6.8
13	6.7	221.8	20.7	–	–	1.1
14	–	563.8	22.3	–	–	5.9
15	–	5.2	2.0	–	–	–
16	–	22.9	0.7	–	–	–
17	–	1112.7	7.7	–	–	–
18	858.6	4.8	7.4	10.6	–	–
19	25.5	116.5	63.8	–	7.8	–
20	4.4	7.1	88.3	–	–	–
21	–	5.1	61.7	–	–	–
22	27.4	140.6	5.3	–	47.0	–
23	–	2205.9	220.5	–	–	23.9
24	6.8	777.0	8.1	–	–	1.6
25	–	384.3	15.3	–	–	–
26	–	91.1	10.3	–	–	–
27	–	333.7	4.9	–	–	3.3
28	139.7	15.6	0.7	–	–	–
29	–	23.2	217.0	–	–	–
30	4.2	3.3	1.6	–	–	–
31	805.6	4.1	3.9	4.8	–	–

Table 3
Results for the screening of 35 counterfeit samples of Cialis® [6].

Sample no.	Ethanol (ppm)	2-Propanol (ppm)	Acetone (ppm)	Ethylacetate (ppm)	Chloroform (ppm)	Tetrachloromethane (ppm)	Dichloormethane (ppm)
1	319.5	–	–	4.4	–	–	–
2	295.5	17.5	–	2.3	–	–	–
3	–	48.4	2.3	–	–	–	–
4	–	14.3	2.5	–	–	–	–
5	–	0.2	–	–	–	–	–
6	–	433.8	30.4	–	–	–	–
7	–	4.5	–	–	–	–	–
8	4.7	722.9	54.0	–	–	–	–
9	–	4.4	–	–	–	–	–
10	–	68.5	3.5	–	–	–	–
11	177.1	–	–	–	–	–	–
12	3.7	1.1	–	–	–	–	–
13	–	34.0	–	–	–	–	–
14	–	30.4	0.8	–	–	–	–
15	–	8.8	–	–	–	–	–
16	–	150.5	9.7	–	–	–	–
17	–	2157.4	162.3	–	–	–	28.7
18	4.3	2072.4	155.2	11.4	–	–	–
19	179.5	–	–	–	–	–	–
20	–	–	–	–	–	–	20.0
21	–	200.7	13.0	–	–	–	–
22	–	639.4	4.9	–	–	–	7.9
23	–	357.1	49.5	–	–	–	22.5
24	3.8	155.0	20.6	–	–	–	–
25	3.2	2.2	–	–	–	–	–
26	–	97.0	7.3	–	–	–	–
27	480.2	1.3	1.8	6.0	–	–	–
28	–	114.0	7.6	–	–	–	–
29	327.4	2.1	–	3.6	26.6	–	–
30	–	2.5	–	–	–	–	–
31	–	91.4	6.0	–	–	–	–
32	–	1670.8	8.6	–	–	–	–
33	–	1.7	–	–	–	–	–
34	–	2.7	–	–	–	–	–
35	403.2	86.5	5.8	–	–	17.3	–

Table 4
Overview of the obtained classification system and the number of samples of the Viagra® and Cialis® sample set belonging to each class.

Class no.	Description	Number of Viagra® samples	Number of Cialis® samples
1	Genuine samples	5	5
2	Samples contain *less* than 100 ppm ICH Class III solvents	8	17
3	Samples contain *more* than 100 ppm ICH Class III solvents	12	12
4	Samples contain ICH Class II solvents (content not taken into account)	9	5
5	Samples contain ICH Class I solvents (content not taken into account)	2	1

far below the international accepted limit of 5000 ppm. The external validation showed 2 misclassified samples. The genuine samples, which were part of the selected test set, were classified as genuine. Tree out of five counterfeit samples were also classified correctly; one class 3 counterfeit was classified as belonging to class 2 and one sample of class 3 was classified as a class 4 sample. The misclassification of the first sample might be explained by the relatively low content of residual solvents. This sample only contains ICH class III solvents in a total amount of 196 ppm, while the boundary between classes 2 and 3 is set at 100 ppm. The misclassification of the second sample as a class 4, instead of class 3, is difficult to explain.

3.4.2. Cialis® like samples

CART analysis was also performed for the Cialis® data set. A tree with complexity 4 turned out to be the optimal tree (figure not shown). Survey of the leaves showed that 3 out of 32 training set samples were misclassified. However, two of these misclassified samples were class 4 samples (=samples containing ICH class II solvents) that were misclassified as genuine, which is unacceptable. The external validation confirmed the inapplicability of this model; 75% of the samples of the test set (6 out of 8) were misclassified. Despite the fact that all genuine samples were classified correctly, this model clearly shows low predictive properties.

In total 9 samples were misclassified. Remarkably, five of these belong to class 4 and one belongs to class 5, meaning that all class 4 and class 5 samples were misclassified (Table 4). This suggests that class 5 might be too small to be modelled correctly. An explanation for the inability to model class 4 might be more complicated since both classes 1 and 4 are represented by the same amount of samples whilst class 1 is modelled properly. The manufacturing of genuine medicines occurs in strictly controlled circumstances and their quality is guaranteed. Therefore these fingerprints are very similar to each other. Furthermore they also differ vastly from the fingerprints of counterfeit samples, which explains why class 1 can be well modelled despite being represented by a low number of samples. Counterfeit medicines on the other hand are produced in uncontrolled circumstances without any quality guarantee. Consequently, differences among the fingerprints of one and the same class might be larger, which might explain the difficulty to model this particular class. Therefore, the CART analysis was repeated without classes 4 and 5, taking only classes 1, 2 and 3 into account. A new test set was selected, using the Duplexx algorithm.

According to the graph representing the cross validation error in function of tree complexity (Fig. 5b) a tree with complexity 1, with a cross validation error of 0.51, should be selected as optimal tree. Obviously, a tree containing only one leaf has very poor predictive properties. However, a tree with 4 leaves could also be built (Fig. 5b). This new tree showed relatively good homogeneous leaves (shown by the graphs in Fig. 5b) and had a cross validation error of 0.48. Two leaves were characterized by complete homogeneity which indicated that all samples in these leaves were classified correctly. The two remaining leaves each showed one misclassified sample. These two misclassified samples belong to class 3; one was misclassified as genuine, the other as class 2. The first misclassified sample (as a genuine) only contains 2-propanol and acetone in a total amount of 464.2 ppm, which is far below the international accepted limit; the second misclassified sample (as class 2) contains a total amount of 121.6 ppm ICH class III solvents, which is close to the set boundary of 100 ppm between classes 2 and 3. After applying the external cross validation, 2 out of 7 test set samples were misclassified. Two genuine samples were classified correctly. Two out of five counterfeit samples in the test set were misclassified. The first sample is a class 2 sample, misclassified as a class 3 sample. Since this particular sample contains 34 ppm of 2-propanol the misclassification is difficult to explain. The second sample (belonging to class 3) contains solvents from ICH class III in a total amount of 1679.4 ppm and therefore the misclassification as a class 2 sample is also difficult to explain.

3.5. SIMCA

3.5.1. Viagra® like samples

SIMCA is a technique which selects a number of principal components to describe each separate class. The number of principal components for each class was selected using leave-one-out cross validation. Two PCs were retained to describe the genuine class; five PCs were selected for class 2; to model class 3 six PCs were kept; for class 4 seven PCs were selected and for class 5 only one PC was retained. The obtained SIMCA model was characterized by a correct classification rate of cross validation of 100%, which indicates that all samples of the training set were classified correctly. No sample was misclassified or unclassified. The external validation showed a correct classification rate of 85.7%, which is due to the misclassification of only one sample (out of 7) of the test set, i.e., a class 2 sample classified as class 3. This misclassification might be elucidated by its total content of ICH class III solvents of 73 ppm, which is quite close to the set boundary of 100 ppm between classes 2 and 3. Both the internal and external validation exhibited a 100% correct classification rate for the discrimination between genuine and counterfeit samples.

3.5.2. Cialis® like samples

Since SIMCA is a technique which models every class separately, it is required that every class is represented by a minimum number of samples. Furthermore it is not possible to model classes represented by only one sample by usage of SIMCA. Therefore this analysis was applied to the second Cialis® data set, taking only classes 1, 2 and 3 into account. The number of principal components was selected using leave-one-out cross validation, resulting in the retention of two PCs for class 1 (genuines), 12 PCs for class 2 and 10 PCs for class 3. The internal cross validation gave a correct classification rate of 100%. This demonstrates that all samples of the training set were classified correctly. The external validation resulted in a correct classification rate of 85.7%. This is due to the misclassification of one (out of 7) test set samples; a class 3 sample which was classified as class 2. This particular sample was also misclassified in the CART tree. It contains 1679.4 ppm ICH class III solvents and therefore the misclassification cannot be explained. A 100% correct classification rate for the discrimination between

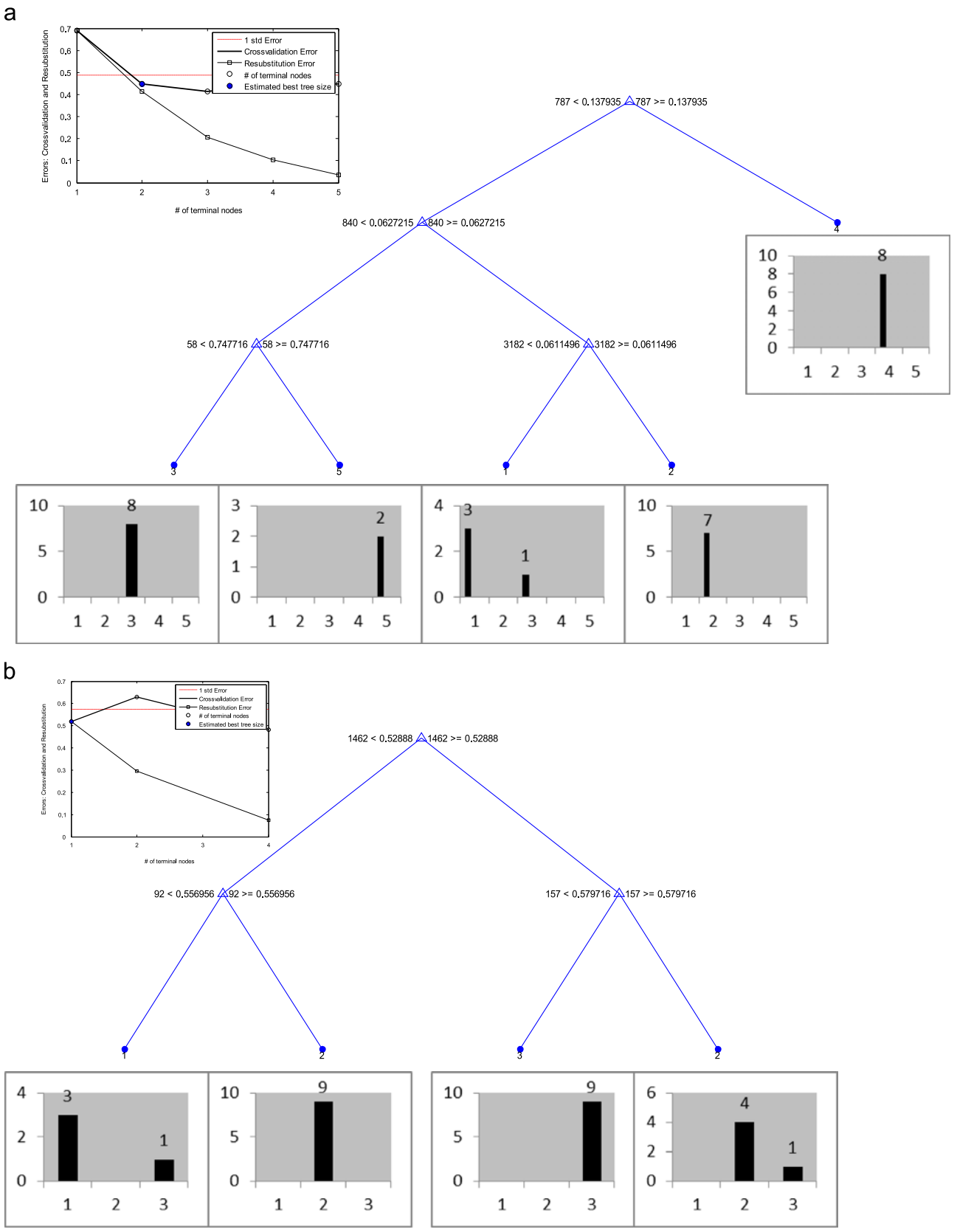

Fig. 5. (a) Classification tree obtained for the Viagra® data set using the Gini index as split criterion. Each split is described by the selected time point and its split value for the abundance (MS). Each leaf is defined by the number of the class which is highest represented in the respective leaf. Each graph represents the number of training set samples in each leaf in function of the classes they belong to. (b) Classification tree obtained for the Cialis® data set using the Gini index as split criterion. Each split is described by the selected time point and its split value for the abundance (MS). 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.

genuine and counterfeit samples was obtained for both the internal and external validation.

4. Discussion and conclusions

Two data sets, one Viagra® and one Cialis® set, were analyzed using chemometric techniques. Two exploratory techniques, i.e., Principal Component Analysis and Projection Pursuit, were applied in an attempt to reveal the structure in both data sets. This exploratory analysis focused on differences between genuine and counterfeit samples and on differences among the counterfeits.

For the Viagra® data set, a clear distinction between genuine and counterfeit samples is obtained by applying PCA. This differentiation between genuine and counterfeit could also be observed with PP for both tested projection indices, i.e., entropy and yenyukov, yet the distinction obtained with PCA is superior. Unfortunately, none of the applied exploratory techniques reveals any structure in the Viagra® data set. All counterfeit samples are clustered together and no pattern can be observed.

PCA yields the best results for the Cialis® data set as well. A clear differentiation between genuine and counterfeit samples is acquired. Moreover, PCA shows slight clustering among the counterfeit samples, which unfortunately turns out to be of no use for the creation of predictive models. However the exploratory analysis shows differences in the fingerprints between genuine and counterfeit samples that could be modelled. The results obtained with PP are inferior to PCA.

Since no patterns are found in both data sets, an arbitrary classification system is set up based on residual solvents content. The resulting classification system comprises 5 classes (Table 4). Once the classes are defined, a test and training set were selected for both data sets using either the Kennard and Stone or the Duplexx algorithm. Subsequently, 2 modelling techniques, i.e., Classification and Regression Trees and Soft Independent Modelling of Class Analogy, were applied and tested for their predictive properties for distinguishing between genuine and counterfeit samples and for the classification in the different defined classes.

The CART model, obtained for the Viagra® data set, showed 3 misclassified samples in total. Unfortunately one misclassification concerns a class 3 sample which is classified as a genuine. Therefore this model is not ideal to describe the Viagra® data set.

The predictive model obtained by SIMCA is characterized by a 100% correct classification rate for cross validation, indicating that all training set samples are classified correctly. The external cross validation of the obtained model shows a correct classification rate of 85.7%, due to the misclassification of one test set sample. Both the internal and external cross validation result in a 100% correct classification rate for the discrimination between genuine and counterfeit samples. This indicates that all genuine samples are classified correctly and no counterfeit sample is classified as being genuine. Consequently, it can be concluded that SIMCA is a suitable model to describe the Viagra® data set. The obtained CART model is inferior to the SIMCA model and SIMCA also has better predictive properties compared to CART.

For the Cialis® data set it was chosen to eliminate the samples of classes 4 and 5 since these classes were too small to model correctly. A CART tree showed that all these samples were misclassified. This disability to model is probably due to the fact that these classes are not represented by an appropriate amount of samples.

The CART model, obtained for the Cialis® data set, showed 4 misclassified samples in total. Similar to the Viagra® data set, one misclassification concerns a class 3 sample which is classified as a genuine. Therefore this model is not suited for the description of the Cialis® data set.

The model obtained by SIMCA shows a 100% correct classification rate for cross validation, indicating that all training set samples are classified correctly. For the external validation a correct classification rate of 85.7% is obtained since one test set sample is misclassified. Both the internal cross validation and the external validation show a 100% correct classification rate for the discrimination between genuine and counterfeit samples, meaning that only the genuine samples were classified as genuine. This clearly shows that SIMCA is superior to CART. SIMCA also has better predictive properties compared to CART.

This study investigated whether differences in GC-fingerprints could be useful to discriminate between genuine and counterfeit medicines and to distinguish between different types of counterfeit medicines according to the risk they pose to public health. Up to now, the risk evaluation of counterfeit medicines comprises mainly the identification and quantification of the active substances present, while potential toxic secondary components, such as impurities, are not explored. This implies that, if the identity and dosage of the active ingredients are correct, a counterfeit medicine can be evaluated as relatively safe, while in actual fact the presence of secondary components might cause serious health problems. Residual solvents are often present in genuine and counterfeit drugs and some of them might be harmful. Therefore it is important that the content of residual solvents is limited. We believe that investigating the residual solvent content of counterfeit medicines provides a valuable approach to evaluate these pharmaceuticals for the risks they pose to public health and to attain a more complete risk evaluation.

This study shows that chemometric analysis of GC impurity fingerprints and analysis of residual solvents are valuable tools to discriminate between genuine and counterfeit medicines. Furthermore, this approach gives a prime notion of the health risks these products constitute. Based on the obtained results it can be stated that for both data sets PCA yields the best discrimination between genuine and counterfeit drugs and SIMCA generates the best predictive models. However, it should be noticed that the proposed methods are only valid for the PDE-5 inhibitors. For other groups of counterfeit drugs, such as slimming products, 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.

References

[1] International Medical Products Anti-Counterfeiting Taskforce (IMPACT), Counterfeit drugs kill, May 2008. ⟨http://www.who.int/impact/FinalBrochureWHA2008a.pdf⟩ (accessed 27.09.13).
[2] World Health Organization (WHO), General information on counterfeit medicines. ⟨http://www.who.int/medicines/services/counterfeit/overview/en⟩ (accessed 27.09.13).
[3] 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.
[4] G. Jackson, S. Patel, S. Khan, 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/dqqt3sge9hwssgcgcos440g40/455_EAASM_counterfeiting%20report_020608(1).pdf⟩ (accessed 27.09.13).
[6] E. Deconinck, M. Canfyn, P.Y. Sacré, P. Courselle, J.O. De Beer, J. Pharm. Biomed. Anal. 81–82 (2013) 80–88.
[7] E. Deconinck, M. Canfyn, P.Y. Sacré, S. Baudewyns, P. Courselle, J.O. De Beer, J. Pharm. Biomed. Anal. 70 (2012) 64–70.
[8] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Q3C: Impurities: Guideline for Residual Solvents, Step 5, February 2009.
[9] Y. Liu, C.Q. Hu, J. Chromatogr. A 1175 (2007) 259–266.
[10] K. Grodowska, A. Parczewski, Acta Pol. Pharm. 67 (2010) 3–12.
[11] European Pharmacopoeia, 7.8, Council of Europe, Strasbourg, France, 2013.
[12] M.J. Rocheleau, M. Titley, J. Bolduc, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 805 (2004) 77–86.
[13] A. Brault, V. Agasse, P. Cardinael, J.C. Combret, J. Sep. Sci. 28 (2005) 380–386.
[14] F. David, R. Szücs, J. Makwana, P. Sandra, J. Sep. Sci. 29 (2006) 695–698.

[15] S. Legrand, J. Dugay, J. Vial, J. Chromatogr. A 999 (2003) 195–201.
[16] C. B'Hymer, Pharm. Res. 20 (2003) 337–344.
[17] K. Grodowska, A. Parczewski, Acta Pol. Pharm. 67 (2010) 13–26.
[18] E. Deconinck, C. De Leersnijder, D. Custers, P. Courselle, J.O. De Beer, Anal. Bioanal. Chem. 405 (2013) 2341–2352.
[19] E. Deconinck, P.Y. Sacré, P. Courselle, J.O. De Beer, Talanta 100 (2012) 123–133.
[20] I. Stanimirova, A. Kubik, B. Walczak, J.W. Einax, Anal. Bioanal. Chem. 390 (2008) 1273–1282.
[21] E. Deconinck, T. Hancock, D. Coomans, D.L. Massart, Y.V. Heyden, J. Pharm. Biomed. Anal. 39 (2005) 91–103.
[22] E. Deconinck, M.H. Zhang, D. Coomans, Y. Vander Heyden, J. Chem. Inf. Model. 46 (2006) 1410–1419.