Development and validation of a method for separation of pregabalin and gabapentin capsules using Near Infrared hyperspectral imaging

Emelie Persson

Examiner: Maria Lundqvist
Supervisors: Louise Elmlund and Simon Dunne
Development and validation of a method for separation of pregabalin and gabapentin capsules using Near Infrared hyperspectral imaging

Seizures containing large numbers of units of narcotics, goods dangerous to health and doping are often sent to the Swedish National Forensic Centre (NFC). Only a fraction of these capsules or tablets can be analyzed, therefore the samples need to represent the whole seizure. If the samples show content variations, Near Infrared (NIR) spectroscopy in combination with hyperspectral imaging has been shown to be a promising tool to gauge the homogeneity in the seizures based on chemical content. The objective of this thesis was to further develop and then validate a method for the separation of pregabalin and gabapentin capsules using NIR hyperspectral imaging and Principal Component Analysis (PCA). Capsules containing different amounts of pregabalin and gabapentin were prepared and analyzed. Additionally, authentic seizures were analyzed to confirm that the method fulfilled its purpose.

The result of this study showed that use of hyperspectral data in the wavelength range 1650-1750 nm gave the best differentiation between pregabalin and gabapentin capsules. Capsules containing the ratio 70-30 % gabapentin and pregabalin could be separated distinctively from capsules containing pure gabapentin. Multiple authentic seizures could be separated into groups correctly depending on the capsules or tablets content.

Nyckelord
NIR, hyperspectral imaging, PCA, homogeneity, chemometrics, pregabalin, gabapentin
Abstract
Seizures containing large numbers of units of narcotics, goods dangerous to health and doping are often sent to the Swedish National Forensic Centre (NFC). Only a fraction of these capsules or tablets can be analyzed, therefore the samples need to represent the whole seizure. If the samples show content variations Near Infrared (NIR) spectroscopy in combination with hyperspectral imaging has been shown to be a promising tool to gauge the homogeneity in the seizures based on the chemical content. The objective of this thesis was to further develop and then validate a method for the separation of pregabalin and gabapentin capsules using NIR hyperspectral imaging and Principal Component Analysis (PCA). Capsules containing different amounts of pregabalin and gabapentin were prepared and analyzed. Additionally, authentic seizures were analyzed to confirm that the method fulfilled its purpose.

The result of this study showed that use of hyperspectral data in the wavelength range 1650-1750 nm gave the best differentiation between pregabalin and gabapentin capsules. Capsules containing the ratio 70-30 % gabapentin and pregabalin could be separated distinctively from capsules containing pure gabapentin. Multiple authentic seizures could be separated into groups correctly depending on the capsules or tablets content.
5.8 Method comparison ........................................................................................................... 23
5.9 Authentic seizures containing other substances ............................................................... 25
5.10 Combined wavelength ranges ......................................................................................... 25
5.11 Process analysis ............................................................................................................... 27
5.12 Impact and ethical implications in a broad sense ............................................................. 27

6. Conclusions ....................................................................................................................... 29

Acknowledgments .................................................................................................................. 30

References ............................................................................................................................. 31

Appendix .................................................................................................................................. 33
  Appendix 1. Software- Spectrum Extraction ........................................................................ 33
  Appendix 2. Repeatability ..................................................................................................... 36
  Appendix 3. Intermediate precision/ruggedness .................................................................. 38
  Appendix 4. Selectivity/robustness ....................................................................................... 39
  Appendix 5. Method comparison ......................................................................................... 41
1. Introduction

1.1 Field of application for Near Infrared (NIR) hyperspectral imaging

Near Infrared (NIR) spectroscopy has been utilized in industrial applications since the 1950s. [1] The method was first used only as an addition to other optical techniques but soon became available as single-unit systems. These systems were later further developed focusing on chemical analysis. Since then, NIR has become a useful technique in many different areas such as pharmaceutical, biomedical, food and agricultural quality control and environmental studies.

In one previous study [2], differences between three strains of three different species of fungi associated with maize were studied using NIR hyperspectral imaging and different multivariate image analysis techniques. A short wave infrared camera within the spectral range 1000-2498 nm with a wavelength interval of 6.3 nm was used. The images were analyzed using the Evince hyperspectral image analysis software package and MATLAB. Different preprocessing methods including, standard normal variate (SNV), multiplicative scatter correction (MSC) and Savitzky-Golay derivatives, were used combined with Principal Component Analysis (PCA) to analyze the results. Score plots including all strains could be differentiated using PCA.

Another study [3] used near infrared chemical imaging (NIR-CI) to determine homogeneity within oral pharmaceutical solid dosage forms. In CI the term homogeneity is used when the concentration of the active pharmaceutical ingredient (API) is nearly the same in every pixel. The hyperspectral images of the tablets were obtained by diffuse reflectance mode using NIR-CI spectrometer with a HgCdTe detector, covering the spectral range 1000-2400 nm. The homogeneity indices were computed using MATLAB and two different spectral pretreatments were used; standard normal variate (SNV) and Savitzky-Golay smoothing. For calculation of the API concentrations, multivariate curve resolution alternating least squares (MCR-ALS) and partial least squares (PLS) models were used. The conclusion of the study was that hyperspectral images provide highly useful qualitative information and can therefore be used to assess homogeneity based on the API distribution in a sample.

In addition, NIR spectroscopy, together with multipoint measurements, has been used to control both the distribution and the content of API within final drug products. [4] This was done by installing a NIR spectrometer equipped with a specially designed multipoint measurement probe on a conveyor belt to dynamically measure pharmaceutical tablets at the end of the manufacturing process.

1.2 Swedish National Forensic Centre

The Swedish National Forensic Centre (NFC) is an independent national expert organization within the Swedish Police Authority. [5] In addition to the central forensic laboratory located in Linköping there are three additional regional forensic laboratories located in Stockholm, Gothenburg and Malmö. NFC has the overall responsibility for the forensic process. Their main task is to conduct forensic investigations and analyzes on behalf of the judicial authorities. They also have an obligation to conduct research and development within the forensic field.
The Drug Analysis Section, located in Linköping, consists of 73 employees (December 2018) who together completed 50,400 cases during 2018. [6] Their principal role is to perform analyses, both qualitatively and quantitatively, to identify restricted substances, such as narcotics, goods dangerous to health and doping agents in seizures.

1.3 NFC’s current routine analysis regarding capsules and tablets
When a seizure consisting of capsules or tablets arrives at NFC a sample is selected for identity analysis. It would be too time and resource consuming to analyze all units in a seizure. Therefore a representative sample must be selected from the seizure as a whole.

The identification of capsules and tablets begins with a detailed visual assessment of color, size and imprinting. Capsules are more difficult to assess since they can be easily opened and modified and then reclosed. The visual assessment is followed by a chemical analysis of the content. For identity investigations a sample of maximum 20 capsules or tablets is analyzed. Seizures of counterfeit pharmaceuticals or doping preparations can often display significant variation in their content, despite identical packaging and appearance. Therefore a method that can screen large seizures for a rapid separation of dissimilar items would be a valuable tool to verify the validity of the sample taken to represent the entire seizure.

In 2012 NFC purchased the UmBio Inspector, a NIR hyperspectral imaging tool, for studying chemical homogeneity in large seizures. Previous studies of NIR hyperspectral imaging at NFC [7] have shown promising results regarding the separation of capsules or tablets containing some particular substances.

1.4 Background and problem formulation
Lately, NFC has discovered a problem regarding capsules from clandestine laboratories that imitate the prescription drug LYRICA. Deviations are often found among the capsules in these seizures where the narcotic substance pregabalin is often partly or fully substituted with gabapentin. However, this deviation is not found in all seizures.

In June 2018 S-pregabalin was legislated as a narcotic substance in Sweden, whereas gabapentin is still legal in the country, which makes the situation above problematic. Therefore, it is essential to separate capsules containing pregabalin from capsules containing gabapentin. The current identifying strategy cannot analyze all capsules, only a few samples. If the samples show variations in chemical content it is desirable that the UmBio Inspector can separate the capsules into different groups. Since this is not an identifying method it must be used as a complement to the current identity analysis. The advantage of using UmBio Inspector is that it can analyze the whole seizure within minutes and can be used on any material which absorbs in the NIR region. In addition, it is a non-destructive technique. This method is an improvement on the current method which involves forensic assessment of the entire seizure based on statistical calculations.

1.5 Thesis objective
The objective of this thesis was to further develop followed by validation of a method for separation of pregabalin and gabapentin capsules using NIR hyperspectral imaging. The deliverable endpoint was to give NFC a basis for determining acceptance criteria as well as for implementing the method into the routine analysis process. It was also desirable to provide focal points for further development for other applications.
1.6 Limitations
The major limitation regarding this project was its timeframe (only ten weeks). This project was a part of a larger project that NFC is conducting with the goal of including the UmBio Inspector method into the routine analysis process. When using NIR hyperspectral imaging it provides data representing the entire contents of the objects. Therefore, the NIR intensity of each substance present in the object influences the analysis, not only the substance of interest. Furthermore, it is used to analyze the homogeneity of the seizures, not the identity.
2. System and Process

The project started with an introduction at NFC with the supervisors. During the first week the UmBio Inspector was introduced as well as the associated software and the capsule preparation equipment. Literature studies concerning NIR, PCA, homogeneity and chemometrics were performed.

During the project, capsules were prepared with different contents in varying concentrations. Analyses were performed on prepared capsules as well as tested on authentic seizures using NIR hyperspectral imaging. All analyses were evaluated before continuing with another. In parallel, the thesis was written continuously. The last couple of weeks were reserved for finishing the thesis and preparing the presentation of the project.

The project was carried out over the period 25th of March to the 14th of June. A GANTT-chart (Figure 1) was designed summarizing the projects various activities with their estimated time frames.

<table>
<thead>
<tr>
<th>GANTT-chart</th>
<th>Development and validation of a method for separation of pregabalin and gabapentin capsules using Near Infrared hyperspectral imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>Hours</td>
</tr>
<tr>
<td>Pre-study</td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>6</td>
</tr>
<tr>
<td>Introduction to NIR-method</td>
<td>3</td>
</tr>
<tr>
<td>Introduction to capsule production</td>
<td>2</td>
</tr>
<tr>
<td>Literature studies</td>
<td>30</td>
</tr>
<tr>
<td>Write project plan</td>
<td>8</td>
</tr>
<tr>
<td>Laboratory</td>
<td></td>
</tr>
<tr>
<td>Produce capsules</td>
<td></td>
</tr>
<tr>
<td>50 % pregabalin or gabapentin</td>
<td>5</td>
</tr>
<tr>
<td>Different bulk</td>
<td>6</td>
</tr>
<tr>
<td>10-30 % pregabalin + 90-70 % gabapentin</td>
<td>8</td>
</tr>
<tr>
<td>Analyze capsules</td>
<td></td>
</tr>
<tr>
<td>Repeatability</td>
<td>2</td>
</tr>
<tr>
<td>Intermediate precision/ ruggedness</td>
<td>10</td>
</tr>
<tr>
<td>Selectivity/ robustness</td>
<td>6</td>
</tr>
<tr>
<td>Matrix effects</td>
<td>8</td>
</tr>
<tr>
<td>Limit of detection</td>
<td>6</td>
</tr>
<tr>
<td>Optimization of wavelength range</td>
<td>4</td>
</tr>
<tr>
<td>Method comparison</td>
<td>8</td>
</tr>
<tr>
<td>Analyze data</td>
<td>30</td>
</tr>
<tr>
<td>Develop method</td>
<td>20</td>
</tr>
<tr>
<td>Final report</td>
<td></td>
</tr>
<tr>
<td>Write introduction</td>
<td>10</td>
</tr>
<tr>
<td>Write theory</td>
<td>30</td>
</tr>
<tr>
<td>Write material and method</td>
<td>30</td>
</tr>
<tr>
<td>Write results and discussion</td>
<td>40</td>
</tr>
<tr>
<td>Write abstract, conclusion and acknowledgement</td>
<td>10</td>
</tr>
<tr>
<td>Submit to examiner</td>
<td></td>
</tr>
<tr>
<td>Submit to opponent</td>
<td></td>
</tr>
<tr>
<td>Correct thesis</td>
<td>30</td>
</tr>
<tr>
<td>Presentation</td>
<td></td>
</tr>
<tr>
<td>Decide material for presentation</td>
<td>8</td>
</tr>
<tr>
<td>Make a PowerPoint</td>
<td>6</td>
</tr>
<tr>
<td>Practice for the presentation</td>
<td>10</td>
</tr>
<tr>
<td>Oppose</td>
<td>4</td>
</tr>
<tr>
<td>Present</td>
<td>5</td>
</tr>
<tr>
<td>Evaluate the project</td>
<td>5</td>
</tr>
<tr>
<td>Other</td>
<td>12</td>
</tr>
<tr>
<td>Meetings with examiner</td>
<td>3</td>
</tr>
<tr>
<td>Total time</td>
<td>361</td>
</tr>
</tbody>
</table>

Figure 1. A GANTT-chart displaying the estimated time frames for the various activities during the project.
3. Theory

3.1 LYRICA, pregabalin and gabapentin
LYRICA (Figure 2) is a Swedish registered drug produced by Pfizer, an American multinational pharmaceutical company. [8] The capsules consist of 25 to 300 mg of pregabalin. In Sweden it is a prescription drug and since 2018 pregabalin is classified as a narcotic substance (classification label V). LYRICA is used, in addition to other medicines, for the treatment of one form of epilepsy. It is also used for the treatment of peripheral and central neuropathical pain; long term pain caused by nerve damage, as well as generalized anxiety disorder in adults.

![Figure 2. LYRICA capsule. [8].](image-url)

Seizures mimicking LYRICA often contain pregabalin, gabapentin or a mixture of the two. Gabapentin is also an active substance in drugs for the treatment of epilepsy and peripheral neuropathical pain and has similar medicinal effects to pregabalin. [9] However, gabapentin is not classified as a narcotic substance in Sweden. The structural formulae of pregabalin and gabapentin are shown in Figure 3. [9, 10] Since gabapentin has a ring structure it will absorb NIR light at different wavelengths compared to pregabalin.

![Figure 3. Structural formulae of pregabalin (left) and gabapentin (right).](image-url)

3.2 NIR hyperspectral imaging
NIR spectroscopy covers the wavelength range 780-2500 nm (Figure 4) corresponding to the wavenumbers 12800-4000 cm\(^{-1}\). Absorptions in the NIR region are generated from fundamental vibrations (stretching and bending) by two distinctive processes; overtones and combinations.
Figure 4. An illustration of the different wavelength ranges in the electromagnetic spectrum.

Overtones can be thought of as harmonics. Every fundamental vibration will produce a series of absorptions at multiples of the fundamental level frequency. Frequency is the reciprocal of the wavelength (equal to the wavenumber expressed in cm\(^{-1}\)). The number of detected overtones from a group of fundamental absorptions in a molecule is limited to a few.

NIR absorptions are at a higher excitation than fundamental absorption so they require more energy. If two or more fundamental absorptions match this NIR energy combination bands can be observed. The most common and energetic combination bands arise from stretch and bend combinations in the same group. Therefore, the combination of O-H stretch with O-H bend and C-H stretch with C-H bend are the absorptions that are often observed. These occur in different positions in the spectrum. A very large number of combinations are possible. Many NIR spectra consist of only a few broad peaks and appear relatively featureless, as a consequence of all these combined absorptions. However, NIR spectra are much more complex than they appear to be.

NIR analysis is a widely accepted analytical tool in, for example, pharmaceutical analysis. One of the reasons why NIR analysis is so useful is that one can use reflected energy. Reflected energy is complex. It consists of two components; specular (mirror-like) and diffuse. Looking at this with a NIR spectroscopy perspective the specular component does not give any information. Unlike the diffuse component which depends on the physical nature of the sample, particle size being particularly important. Changes in the spectrum can be observed caused by variation in physical parameters of a sample, leading to the consequence that the observed spectrum is a mixture of chemical and physical information. Due to the measurement of reflected energy NIR analysis can be performed with little or no sample preparation. [11]

3.3 Pre-processing of data
3.3.1 Why pre-processing?
During data collection variations in conditions as well as instrumental variations can occur. These variations can cause differences in global trend, total energy, high-frequency noise and/or local background. [12] For solid samples, undesired systematic variations are primarily caused by light scattering. The main types of these scatterings, that do not include energy transfer with the sample, are Rayleigh and Lorentz-Mie. They are processes in which the electromagnetic radiation is scattered by for instance small particles, bubbles, droplets or surface roughness. Another major cause of undesired systematic variations is differences in effective path length. [13]
The total variation in a sample set often consists primarily of undesired variations. This can be observed as shifts in baseline and through another phenomenon called non-linearities. For NIR reflectance measurements the specular reflections, which do not contain any chemical information as previously mentioned, are normally minimized by instrument design and sampling geometry. On the other hand, the diffusely reflected light is the primary source of information in the NIR spectra. Unfortunately it will not only contain information on chemical composition of the sample, but also the micro-structure. [13]

Based on the chemical content, it is necessary to eliminate undesired variations to be able to perform an accurate separation or classification of a data set. This can be done by using one or more of the pre-processing techniques available for NIR spectra.

3.3.2 Normalization
Spectra often require normalization for an effective comparison between different samples; to be able to find similarities and differences based on the chemical content. Normalization is supposed to do what cannot be done physically during data collection, numerically recover exact replicates when no chemical differences exist. [12]

Data from highly sensitive instruments are often contaminated by noise due to subtle changes in instrumental settings or conditions. Noise is more accurately described as non-discriminatory sample-specific signal ranging from broadband background to high-frequency jitter. By normalization of data, the graphs integral is set to one. This means that the sample is divided into groups according to contents, not concentrations. [12]

3.3.3 Quadratic detrend
Quadratic detrending is used for data sets to remove non-linear trends generated from spectroscopic measurement. The detrending function is separated from all other operations and is applied to each individual spectrum. It is done by calculating the baseline function as a least square to the polynomial determined by the samples spectrum. There are different orders of polynomial that the detrending function can use. For the zero-degree polynomial the dislocation is removed, for the first-degree polynomial the gradient is removed as well and for the second-degree the curvature is also removed. [14]

3.3.4 Mean centering
Subtract average is used when the differences between a data point and the mean of the whole data set wants to be determined. A mean spectrum, based on all the data, is calculated and the differences for each spectrum compared to the mean spectrum is shown.

3.4 PCA
There can be difficulties detecting patterns and relationships when analyzing multivariate data because of the sheer volume of information. Moreover, the information can be partly hidden in the complex data which makes interpretation difficult. Exploratory Data Analysis (EDA) provides visual approaches to find patterns in data and one of the most powerful multivariate EDA is called PCA. [14] PCA is a useful technique for reducing the dimensionality of data if there is some correlation between the variables. [15] It quantifies the amount of useful information in the data, which is equivalent to variables that exhibit large systematic variation, and disregards the noise, variables that exhibit small variations. [14]
In a multidimensional space each object in a data table represents a point. The location of that point is determined by its coordinates, which are equal to the cell values of their corresponding row in the table. Each variable composes a coordinate axis. In Figure 5 the wavelengths represent different variables and the numbers represent different objects. The picture to the right in Figure 5 illustrates the data point for the marked cells in the table in a multidimensional space.

Figure 5. To the left is a table from The Unscrambler and the picture to the right illustrates the data point for the marked cells in a multidimensional space.

If two objects should be considered as similar most of their variables should have values close to each other (see Figure 6). Their data points are then close to each other in space. The opposite is if two samples have values that significantly differ for at least some of their variables. Their data points are then occupying different areas in the given multidimensional space.

Figure 6. Different objects in a data set separated into two groups.

PCA is based on the idea of finding latent variables called principal components; Z₁, Z₂, etc., which are linear combinations of the original variables describing each specimen; X₁, X₂, etc. [15] This is performed by finding the directions in space along which the dispersion of the data points is the largest (see Figure 7) [14]. The principal components are chosen so that the first principal component (PC1), Z₁, accounts for most of the irrespective variation in the data set and each subsequent principal component accounts for less variation than the previous one. The number of principal components computed depends on the number of variables or
samples in the data set. The principal components generate new dimensions and form a new coordinate system. [15] The principal components should not have a covariance or a correlation and they should be irrespective of each other. These new dimensions have two advantages over the original set of axes; the principal components are orthogonal to each other and they are ranked depending on how much information they carry. In this way the interpretation can be prioritized.

The principal components are obtained from a covariance matrix or a correlation matrix of the data. The latter is used when, for example, one variable has a much larger variance than the others and therefore dominates the first principal component. The correlation matrix standardizes each variable to zero mean and gives them the variance of 1, making all variables carry equal weight. Unfortunately, if the original variables have very different variances the standardization can have a considerable effect on the PCA scores plots. Covariance is a measure of the joint variance of two variables.

The principal components are the eigenvectors of the chosen matrix and the technique for finding these eigenvectors is called eigenanalysis. By doing an eigenanalysis on the chosen matrix the coefficients of the principal components are obtained. The values that the principal components take for each of the objects can be calculated by substituting the relevant values of \(X_1, X_2,\) etc., into the formula for \(Z_1, Z_2,\) etc. This value is sometimes referred to as a “score” and each object has a score on each principal component. The score defines an objects location along a principal component and it represents the coordinate of the object on the principal component. [14] The principal components relative importance is expressed in terms of how much variance of the original data set it describes. By plotting the scores object differences or similarities can be observed.

### 3.5 UmBio Inspector

The UmBio Inspector (Figure 8) consists of a conveyor (1), lamps with NIR light (2) and a NIR camera (3) and is operated through an operator screen (4). The sample is placed on a tray (5) which is transported under the camera while being irradiated with NIR light while the NIR camera registers the measurement. Before each analysis a white and a black reference are imaged. The camera generates a NIR spectrum for every pixel in the picture (320x320 pixels). [16] After the analysis the data is transferred to the software (Spectrum Extraction) for further data processing to receive mean spectra of all the pixels within each object.
3.6 Validation
A validation is associated with method development and is used to prove that a method fulfils its intentions. Moreover, the validation brings out knowledge that can be used to assure that a method fulfils its intentions while using it. The validation includes different supervisions to assure that a method is adequate to its aim. These supervisions need to represent different aspects of the methods ability to achieve results. [17] The validation in this study focused on supervision of repeatability, intermediate precision, selectivity, robustness, matrix effects, limit of detection and method comparison.
4. Materials and Methods

4.1 Instrument
The UmBio Inspector (UmBio AB) consists of a conveyor and a NIR camera operated by the software Breeze. For the NIR analysis, the spectral range 1100-2400 nm with a wavelength interval of 6.34 nm was used. For processing and analysis of the hyperspectral data the software developed in-house by the Image analysis group (NFC), Spectrum Extraction, was used (see Appendix 1). After selection of spectral range, the data was preprocessed by normalization, quadratic detrending and mean centering, resulting in a two principal component PCA scores plot.

4.2 Chemicals
The chemicals that have been used in this study are pregabalin (seized material), gabapentin (Sigma-Aldrich, Lot # LRAB7794), cellulose (Sigma-Aldrich, Lot # MKCG8490), talc (Sigma-Aldrich, Lot # MKCG6924), lactose (AnalaR, BDH Laboratory Supplies, Lot K23086687), starch (Merck, KGaA), D-(+)-glucose (KEBOLab) and magnesium oxide (Sigma-Aldrich, Batch # 0000036528).

4.3 Preparation of capsules
Feton capsules filling kit was used for preparation of capsules, allowing the preparation of 100 capsules/batch. The content of all capsules were weighed using a balance with an accuracy of three decimals.

Capsules containing a mixture of 50 % pregabalin (P50) or 50 % gabapentin (G50) in lactose were prepared. The concentration of pregabalin in the prepared capsules is equivalent to the amount of pregabalin in LYRICA (50-60 % pregabalin/capsule). Lactose is a common filling component in capsules and is present in LYRICA. Moreover, since seizures of false LYRICA often contain high concentrations of pregabalin, gabapentin or a mixture of both, capsules containing mixtures of gabapentin and pregabalin were prepared in different ratios; 90-10, 80-20 and 70-30 %. Finally, capsules containing pure cellulose, talc, lactose, starch, glucose and magnesium oxide were prepared. To study the precision in capsule weight for the prepared capsules, selected items were weighed using an analytical balance with an accuracy of 0.1 mg.

Table 1 shows a summary of the content in the prepared capsules containing either pregabalin, gabapentin or a mixture of them both.

Table 1. Summary of the content in the prepared capsules and quantities prepared.

<table>
<thead>
<tr>
<th>Capsule label</th>
<th>% pregabalin</th>
<th>% gabapentin</th>
<th>% lactose</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>P100</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>G100</td>
<td>-</td>
<td>100</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>P50</td>
<td>50</td>
<td>-</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>G50</td>
<td>-</td>
<td>50</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>P10G90</td>
<td>10</td>
<td>90</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>P20G80</td>
<td>20</td>
<td>80</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>P30G70</td>
<td>30</td>
<td>70</td>
<td>-</td>
<td>10</td>
</tr>
</tbody>
</table>
4.4 Validation

4.4.1 Optimization of wavelength range
To determine the optimal wavelength range for analysis in this application, capsules P50 and G50 were analyzed (see Table 1 for capsule definitions). Firstly, an image analysis was performed in the full wavelength range (1100-2400 nm). The spectrum was then divided into shorter segments to determine the wavelength ranges where the most evident separation was visualized in the preprocessed spectra and their corresponding PCA score plot.

4.4.2 Limit of detection
The limit of detection was studied by analyzing capsules with different ratios of gabapentin and pregabalin. Ten capsules each of P10G90, P20G80 and P30G70 (see Table 1) were analyzed on separate trays. To minimize variations due to positioning on the tray, all the capsules were placed in the same column.

4.4.3 Repeatability
The repeatability was examined in two different modes using capsules with different contents. To minimize variations due to positioning on the tray, ten consecutive analyses were performed on the same day with the capsules placed on the same position on the tray. The study was performed, in parallel, with one tray containing two items of G100 and two items of P30G70 and another tray containing ten items of G50 and ten items of P50.

4.4.4 Intermediate precision/ruggedness
Intermediate precision/ruggedness was investigated by analyzing a tray consisting of 80 items of P50 and ten items of G50, each day for ten days straight (excluding the weekends). The capsules were initially placed in random positions on the tray and to minimize variations due to positioning on the tray they were placed in the same position for each measurement.

4.4.5 Selectivity/robustness
Selectivity/robustness was examined by analyzing a tray with 10 items of G50 and 80 items of P50 ten times on the same day. Each time the capsules were placed in different random positions on the tray. A second study was made using two items of G100 and ten items of P30G70.

4.4.6 Matrix effects
To study the NIR intensity of different filling agents, ten of each capsules containing pure cellulose, talc, lactose, starch, glucose and magnesium oxide were analyzed together with ten empty capsules. In addition, capsules with pure pregabalin and pure gabapentin respectively were analyzed.

4.4.7 Method comparison
To confirm the results from the study, five authentic seizures were analyzed according to the same parameters as the prepared capsules. The seizures were separated into different groups after analysis with the UmBio Inspector, followed by analysis with the current identifying method at NFC to confirm the groupings.
4.5 Authentic seizures containing other substances
Alongside the validation in this project, the UmBio Inspector has been used to separate capsules or tablets in seizures containing substances other than gabapentin and pregabalin. All seizures that have shown deviation in content after the initial identity analysis have been analyzed using NIR hyperspectral imaging. Since different chemical substances are NIR active in different spectral ranges, the evaluation of all analyses was first performed using the full spectral range, 1100-2400 nm. If necessary, the spectral range was divided into shorter segments, just like for pregabalin and gabapentin, to find the optimal wavelength range for those specific substances.
5. Results and discussion

5.1 Preparation of capsules

A summary of all the weighed capsules, their calculated means and standard deviations is presented in Table 2. The pre-processing (normalization) of the NIR data before evaluation should remove the differences based on the amount of filling in capsules. However, at least for further studies it is good to have a knowledge regarding the differences in weight between the different contents, especially if other pre-processing methods will be used.

Table 2. A summary of the weighed capsules calculated means and standard deviations.

<table>
<thead>
<tr>
<th>Content</th>
<th>Weight</th>
<th>Mean (g)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empty capsules</td>
<td></td>
<td>0.05893</td>
<td>0.00159</td>
</tr>
<tr>
<td>Cellulose</td>
<td></td>
<td>0.19102</td>
<td>0.00361</td>
</tr>
<tr>
<td>Talc</td>
<td></td>
<td>0.17117</td>
<td>0.00565</td>
</tr>
<tr>
<td>Glucose</td>
<td></td>
<td>0.27705</td>
<td>0.01491</td>
</tr>
<tr>
<td>Magnesium oxide</td>
<td></td>
<td>0.13324</td>
<td>0.00318</td>
</tr>
<tr>
<td>Lactose</td>
<td></td>
<td>0.27929</td>
<td>0.01253</td>
</tr>
<tr>
<td>Starch</td>
<td></td>
<td>0.32423</td>
<td>0.00783</td>
</tr>
<tr>
<td>50 % pregabalin</td>
<td></td>
<td>0.30564</td>
<td>0.00570</td>
</tr>
<tr>
<td>50 % gabapentin</td>
<td></td>
<td>0.25868</td>
<td>0.02035</td>
</tr>
</tbody>
</table>

5.2 Optimization of wavelength range

To increase the ability of separation the spectral data was divided into shorter segments to determine the wavelength range which is most suitable for data evaluation in that particular case. The aim was to amplify the separation between capsules by only studying wavelength ranges of relevance. Spectra from capsules containing 50 % pregabalin (P50) and 50 % gabapentin (G50), respectively, were evaluated. The separation between the two substances was most evident in the wavelength range 1650-1750 nm.

In Figure 9, the wavelength ranges 1100-2400 nm and 1650-1750 nm are compared. The spectrum A1 and PCA scores plot A2 illustrates the separation between capsules P50 (pink) and G50 (green) in the wavelength range 1100-2400 nm. B1 and B2 illustrates the spectra and PCA scores plot for 1650-1750 nm. The separation is much more evident in the PCA scores plot when the spectral data from only 1650-1750 nm is processed.

To further confirm that the spectral range 1650-1750 nm is most evident for separation of gabapentin and pregabalin, the substances were analyzed together with empty capsules. Figure 10 presents the spectrum in the wavelength range 1650-1750 nm and the PCA scores plot after analysis of empty capsules (green) and capsules containing pure pregabalin (P100, pink) and gabapentin (G100, grey), respectively. The spectrum illustrates that pregabalin and gabapentin are NIR intense in different wavelength ranges, which makes it possible to separate the substances within this wavelength range.
Figure 9. The spectrum A1 has the wavelength range 1100-2400 nm and the PCA scores plot A2 show the separation between capsules containing 50 % pregabalin (pink) and 50 % gabapentin (green). B1 and B2 illustrates the separation between the same capsules utilizing data only in the wavelength range 1650-1750 nm for the PCA analysis.
Figure 10. The preprocessed spectra and PCA scores plot after analysis of empty capsules (green) and capsules containing pure pregabalin (pink) and gabapentin (grey).

During the evaluation of optimized spectral range also data at 1150-1300 nm displayed some separation of the capsules, although not as evident as at 1650-1750 nm. The wavelength ranges 1400-1600 nm and 1800-2000 nm showed no separation. At 2000-2400 nm the PCA scores plot showed grouping based on capsule position (rows) rather than chemical content. In Figure 11 this is demonstrated using two examples: the left plot illustrates the PCA scores plot from one seizure containing pregabalin and the right plot illustrates the analysis of 80 capsules containing 50 % pregabalin (P50) and 10 capsules containing 50 % gabapentin (G50). The number of the red circles illustrates the row numbers of the tray.

Given the results, all further analyses of pregabalin and gabapentin capsules were executed with the wavelength range 1650-1750 nm.

Figure 11. Left PCA scores plot of pregabalin capsules (seizure) and right PCA scores plot of 50 % pregabalin and 50 % gabapentin. The wavelength range is 2000-2400 nm. The number of the red circles illustrates the row numbers of the tray.

5.3 Limit of detection

Limit of detection in this application can be interpreted as the smallest deviation between capsules that can be elucidated in the scores plot of a PCA. To determine the limit of detection in this project the capsules containing mixtures, in different ratios, of gabapentin and pregabalin were analyzed and evaluated at 1650-1750 nm. Since pregabalin is classified as a narcotic substance it was important to determine the minimum concentration of pregabalin that could be separated from gabapentin in capsules.
Figure 12 presents the preprocessed spectra A1 and the PCA scores plot A2 after analysis of capsules containing pure gabapentin (G100, left grey), pure pregabalin (P100, right grey) and capsules containing 30 % pregabalin in gabapentin (P30G70, green). In the spectra (Figure 12, A1) there is a distinct difference between the capsules. However, the PCA scores plot (right) does not show an evident separation between G100 and P30G70. By removing data from one cluster (P100) in the analysis the separation between the two remaining groups (G100 and P30G70) are more distinct in the PCA scores plot (see Figure 12 B1 and B2). These results illustrates the possibility to increase separation by reducing the number of variations.

Figure 12. The preprocessed spectra A1 and PCA scores plot A2 present the results after analysis of capsules containing pure gabapentin (left grey) and pregabalin (right grey), respectively, and capsules containing the ratios 70-30 % gabapentin and pregabalin (green). In B1 and B2 the pregabalin capsules have been removed from the analysis.
Figure 13 presents the preprocessed spectra A1 and the PCA scores plot A2 after analysis of capsules containing pure gabapentin (G100, left grey), pure pregabalin (P100, right grey) and capsules containing 20 % pregabalin (P20G80, green). In the spectra (Figure 13 A1) there is still a difference between the capsules. However, the PCA scores plot (A2) does not show an evident separation between G100 and P20G80. By removing data from one cluster (P100) in the analysis the separation between the two remaining groups (G100 and P20G80) are more distinct in the PCA scores plot (see Figure 13 B1 and B2). However, it is not as trustworthy as the separation between G100 and P30G70. As to the capsules containing 10 % pregabalin, in gabapentin, no separation from G100 in the PCA scores plot was noticeable.

Figure 13. The preprocessed spectra A1 and PCA scores plot A2 present the results after analysis of capsules containing pure gabapentin (left grey) and pregabalin (right grey), and capsules containing the ratios 80-20 % gabapentin and pregabalin (green). In B1 and B2 the pregabalin capsules have been removed from the analysis.
5.4 Repeatability
Repeatability was studied in order to get an estimation of the spread in results between ten replicates, using the same method by the same person on the same day. Capsules containing 50 % pregabalin (P50) and 50 % gabapentin (G50), were analyzed and evaluated ten times on the same day at 1650-1750 nm. The evaluation of the PCA scores plots showed that the capsules can be confidently separated in all ten cases. Figure 14 shows the PCA scores plots for four of the ten repeatability analyses (all analyses are attached in Appendix 2). In each analysis two groups can be visualized separated due to their content, pregabalin (green) or gabapentin (pink). The different groups were equally distributed and the distance between the groups were comparable.

![PCA scores plots](image)

Figure 14. The PCA scores plot from the study of repeatability of capsules containing 50 % pregabalin (green) and 50 % gabapentin (pink). The pictures represent four of the ten analyses.

Likewise, capsules containing pure gabapentin (G100) and capsules containing the ratio 70-30 % gabapentin and pregabalin (P30G70) were analyzed and their PCA scores plots were evaluated. Figure 15 represents the PCA scores plots from four of the ten repeatability analyses (see all ten analyses in Appendix 3). Two groups can be visualized, G100 in grey and P30G70 in green.
5.5 Intermediate precision/ruggedness
The intermediate precision was studied to determine the variation between replicates being analyzed on different days. The intermediate precision was studied by analyzing 80 capsules containing 50% pregabalin (P50) and 10 capsules containing 50% gabapentin (G50) for ten days straight followed by data evaluation in the spectral range 1650-1750 nm. The PCA scores plots from four of the ten days are presented in Figure 16 (all analyses are attached in Appendix 3). It can be concluded that P50 (green) and G50 (pink) can be separated each day for ten days straight since the PCA shows two separate groups each day. In addition, the clusters in all analyses showed equal distribution.
Figure 16. PCA scores plot from the study of intermediate precision using 80 capsules containing 50% pregabalin (green) and 10 capsules containing 50% gabapentin (pink). The pictures represent four of the ten days.

5.6 Selectivity/robustness

Robustness describes how well an analytical method can perform unaffected by small, deliberate changes in operating parameters. In this project the selectivity and robustness were studied to determine the methods ability to separate capsules irrespective of their position on the tray. The selectivity study was performed by analyzing 80 capsules containing 50% pregabalin (P50) and 10 capsules containing 50% gabapentin (G50) ten times on the same day with the capsules in different positions on the tray each time. The spectral range for data evaluation was set to 1650-1750 nm. Figure 17 represents the PCA scores plot and their corresponding hyperspectral images of two of the ten analyses (the results from all ten analyses are attached in Appendix 4). The result shows that the capsules can be separated each time no matter position on the tray. Figure 17 shows two separate groups which are clearly separated, P50 in pink and G50 in green.

In addition, the selectivity was studied, under the same manners as above, using capsules of pure gabapentin and capsules containing 30% pregabalin in gabapentin (P30G70). Also in this set up, the separation was evident in the PCA scores plot, data shown in Appendix 4. These results prove that the method is reliable regardless of the capsules positions on the tray and that it provides a correct separation in the studied wavelength range.
Figure 17. The study of the selectivity with 80 capsules containing 50% pregabalin (pink) and 10 capsules containing 50% gabapentin (green). The pictures represent the PCA scores plot and the hyperspectral images of two of the ten. The wavelength range was 1650-1750 nm.

5.7 Matrix effects
Matrix effects were studied to gain knowledge regarding the possible interference by different filling agents in the measurements. By analyzing capsules containing pure filling agents followed by evaluation together with data from empty capsules, the NIR intensity of the filling agents, at different wavelengths, could be examined. Figure 18 illustrates the preprocessed and mean-centered NIR spectra for cellulose (1), talc (2), lactose (3), starch (4), glucose (5) and magnesium oxide (6). The bulk agents are in green and the empty capsules in pink.

Cellulose, lactose, starch and glucose are all sugars and they are therefore NIR active in similar wavelength ranges (1200-1400 nm and 1650-2100 nm), thus their spectra look alike. Talc and magnesium oxide have the lowest NIR activity in the wavelength range 1650-1750 nm. As previously mentioned, in Section 5.2, pregabalin and gabapentin showed most evident NIR intensity at 1650-1750 nm.

It is important to understand that, since the bulk agents are NIR active in different spectral ranges, capsules containing different bulk agents can be separated based on these and not necessarily on the active substances. However, an identity analysis of one sample from different clusters will show if the capsules contain the same active substances or not.

5.8 Method comparison

Five authentic seizures were analyzed and evaluated (spectral range 1650-1750 nm) to see that the method fulfilled its aim to separate pregabalin and gabapentin capsules. Figure 19 illustrates the preprocessed spectra and PCA scores plot for one of the seizures. Identity analysis of samples later confirmed that the green group represents capsules containing pregabalin and the grey gabapentin. Data from additional three analyzed seizures are attached in Appendix 5.

Figure 19. The preprocessed spectra in the wavelength range 1650-1750 nm and the PCA scores plot of an authentic seizure. The green group represents capsules containing pregabalin and the grey gabapentin.
The method was also tested on an authentic seizure containing 243 false LYRICA capsules. The initial identity analysis (20 capsules) showed variations in content; some capsules contained gabapentin and some contained pregabalin. By using the UmBio Inspector, the capsules could be separated into two groups (see Figure 20).

![Figure 20](image)

**Figure 20.** A seizure separated into two groups; the pink consisting of gabapentin and the green pregabalin.

Initially, a few capsules were analyzed by Fourier Transform Infrared Spectroscopy (FT-IR) for confirmation of the groupings and later by Liquid Chromatography-tandem Mass Spectrometry (LC-MS/MS) which confirmed the content. However, the LC-MS/MS analysis showed that the pregabalin-group (shown in green) contained minor amounts of gabapentin (shown in pink) in different ratios. Therefore, all capsules containing pregabalin and the same quantity of gabapentin capsules, placed in the same positions on different trays, were analyzed to locate capsules in the outskirts of the pregabalin group. Four capsules, marked with red circles in Figure 21, were analyzed with regard to concentration using Nuclear Magnetic Resonance (NMR). Capsule number one contained 36.6 % pregabalin, number two 97.4 %, number three 86.2 % and number four 55.0 %. The remaining material in these capsules consisted of gabapentin. This study proves that the capsules are separated due to concentration differences, which can explain why some clusters are wide.

![Figure 21](image)

**Figure 21.** A PCA scores plot showing the capsules from which concentrations were determined using NMR.
5.9 Authentic seizures containing other substances

The initial identity analysis of a seizure consisting of green, round tablets, showed that the tablets contained different concentrations of closely related substances (anabolic steroids). By analyzing the tablets with the UmBio Inspector, at 1100-2400 nm, the tablets could be separated into two groups, one pink and one green (see Figure 22). The following identity analysis confirmed the tablets contents, the tablets in the pink group contained mesterolon and the tablets in the green group contained oxymetholone.

![Figure 22](image)

Figure 22. The preprocessed spectra and the PCA scores plot after analysis of a seizure consisting of green, round tablets containing mesterolon (pink) and oxymetholone (green). The spectral range was 1100-2400 nm.

Another seizure consisting of almost 1000 green tablets was sent to NFC by the customs service for separation using the UmBio Inspector. According to previous identity analyses the tablets contained either etizolam or etizolam together with cyproheptadine and amantadine. After analyzing the tablets using the UmBio Inspector no clear separation could be obtained (see Figure 23). The broad group in the scores plot could be the result of variation in concentration of the various components. The marked green group in Figure 23 was thought to be separated due to content, but was actually separated due to positioning on tray.

![Figure 23](image)

Figure 23. The preprocessed spectra and PCA scores plot of one tray of a seizure consisting of green tablets. The wavelength range was 1100-2400 nm.

5.10 Combined wavelength ranges

During the project a new version of the software Spectrum Extraction was developed by the image analysis group at NFC. This new version made it possible to select two or more wavelength ranges at the same time. The upper spectral traces and PCA scores plot shown in
Figure 24 shows the separation of capsules containing the ratios 80-20 % gabapentin and pregabalin (P20G80, pink) and capsules containing pure gabapentin (G100, green) analyzed in the wavelength range 1650-1750 nm. The circled dot should belong to the green group. The lower spectral traces and PCA scores plot show the separation of the same capsules but data from both 1150-1300 nm and 1650-1750 nm were combined in the evaluation. Here the circled dot lies in closer proximity to the correct group. These results indicates the possibility to improve the separation by using two or more spectral ranges in combination.

Figure 24. Preprocessed and mean-centered NIR spectra and PCA scores plots after analysis of capsules containing the ratio 80-20 % gabapentin and pregabalin (P20G80, pink) and capsules containing pure gabapentin (G100, green). The upper pictures are in the wavelength range 1650-1750 nm and the lower are complemented with the wavelengths 1130-1300 nm.
In addition, studying capsules containing the ratio 70-30 % gabapentin and pregabalin (P30G70, grey) and capsules containing pure gabapentin (G100, green) with combinational spectral ranges displayed more evident separation along the PC1-axis. However, the separation is more diffuse. The PCA scores plot to the left in Figure 25 is based on the wavelengths 1650-1750 nm and the one to the right is complemented with the wavelengths 1150-1300 nm.

Another study was performed where the wavelength range was reduced to 1680-1720 nm, since that area has displayed the biggest differences between gabapentin and pregabalin. Nevertheless, no difference regarding the separation was observed.

5.1 Process analysis
The activities in the GANTT-chart (Figure 1) were all performed as planned. However, due to lack of gabapentin, because of the expense, the capsules containing either 70, 80 or 90 % gabapentin were produced after each other using the same material. First, the capsules containing 90 % gabapentin were produced and then analyzed with the UmBio Inspector, before the capsules were emptied and the content diluted with pregabalin to get 80 % gabapentin and so on.

After getting the knowledge that the limit of detection was 30 % pregabalin in gabapentin, one repeatability study and one selectivity study were performed using these capsules. They were compared to capsules containing pure gabapentin and the studies were performed to confirm that the method fulfilled its purpose when analyzing capsules containing the smallest detectable deviation.

After the validation was performed some problems regarding the license for the software occurred. Therefore, no further seizures could be analyzed during the final weeks of the project.

5.12 Impact and ethical implications in a broad sense
The purpose of the implementation of the developed method for separation of pregabalin and gabapentin capsules was to get a more precise method for determining the homogeneity in large seizures. Since pregabalin is classified as a narcotic substance the separation of capsules containing only non-classified gabapentin is of great importance for increased legal certainty.
To prevent that narcotic substances are found and thereafter used or end up in the environment, all materials containing these substances are being destroyed.

To protect anonymity and confidentiality no registration numbers of seizures have been mentioned in this report. Besides this no other ethical considerations need to be discussed since no people or biological materials have been used in this study.
6. Conclusions
Capsules from clandestine laboratories can often display variations in chemical content, despite identical packaging and appearance. The variability places serious demands on the sampling strategy used to generate a representative sample for analysis.

The aim of this study was to develop and validate a method for separation of capsules based on their chemical content; gabapentin or pregabalin. For this application, the most evident separation was displayed within the wavelength range 1650-1750 nm. Since pregabalin is a narcotic substance it was important to study the limit of detection, i.e. the lowest concentration of pregabalin that can be separated from pure gabapentin. The results showed that capsules containing 30 % pregabalin mixed with gabapentin can, with certainty, be separated from capsules containing pure gabapentin. Capsules containing 20 % pregabalin, in gabapentin, showed in some cases separation. 10 % pregabalin could not be separated from pure gabapentin. For the capsules mixed with lactose, 50 % pregabalin could be separated from 50 % gabapentin. The study of repeatability, intermediate precision and selectivity showed separation, as expected, each time and proved that the method is robust. Given the results, the method is validated and ready to implement into the daily work at NFC.

In this study authentic seizures containing other substances were also studied. Initial identity analyses showed variations within the seizures due to differing chemical content. The seizure containing tablets of either oxymetholone or mesterolone could be separated successfully but the seizure containing tablets of etizolam or a combination of etizolam, cyproheptadine and amantadine could not be separated. It is worth noting that these results are based on full spectral data, 1100-2400 nm. No information regarding the concentrations of each substance or filling agent was available.

The use of NIR hyperspectral imaging enables the possibility to separate capsules based on their chemical content. The outcome of the separation depends on numerous factors; the NIR intensity of the substance of interest as well as the filling agents, the concentrations, the nature of the variation and the ratio between the components. Therefore it is important in future studies to verify the optimal wavelength range for each application. To increase the separation ability, a combination of several shorter wavelength intervals should be further investigated.
Acknowledgments
I have had the indulgence to do my Bachelor Thesis project at NFC. This has given me valuable experiences and knowledge for the future. I have also gotten an inspiring insight into the Drug Analysis Sections work.

First and foremost I would like to thank my supervisors Louise Elmlund and Simon Dunne at NFC for all their help and encouragement during my project. I am grateful for all their knowledge that they have shared with me and they have taught me a lot. They are the reason that I could carry out this project in such a good way. Furthermore, I would like to thank all the coworkers who have willingly answered my questions when needed.

Lastly, I would like to thank Maria Lundqvist who has been my examiner during this project.
References


[14] The Unscrambler X 10.3 (64-bit).


Appendix

Appendix 1. Software- Spectrum Extraction

In Figures 26-31 the windows for the software developed in-house by the Image analysis group (NFC), Spectrum Extraction, are displayed. Figure 26 shows the window where the datasets were added and removed. In the hyperspectral image each object is marked with an identity number.

Figure 26. Software window showing the tray and the objects marked with numbers.

The objects on a tray are selected in the window displayed in Figure 27. Here all the background is removed.

Figure 27. Software window showing the selected objects.
The software has a feature for manually removing objects from the analysis of data (see Figure 28).

Figure 28. Software window showing a tool for manual removal of foreground regions.

Figure 29 shows a preprocessed spectra in the wavelength range 1100-2400 nm. The different colors represents the different objects.

Figure 29. Software window showing the preprocessed spectra.
During the project a new version of the software Spectrum Extraction was developed by the image analysis group at NFC. This new version made it possible to select two or more wavelength ranges at the same time. Figure 30 illustrates the new window for preprocessed spectra.

Figure 30. Software window from the new version, showing the preprocessed spectra.

Figure 31 displays the PCA scores plot which was used to see if there was a significant difference based on the content between capsules or tablets.

Figure 31. Software window showing the PCA scores plot.
Appendix 2. Repeatability
Capsules containing 50 % pregabalin (P50) and 50 % gabapentin (G50), were analyzed ten times on the same day at 1650-1750 nm. Figure 32 shows the PCA scores plots for the ten repeatability analyses, pregabalin shown in green and gabapentin in pink.

Figure 32. PCA scores plot for repeatability analysis of capsules containing 50 % pregabalin (pink) and capsules containing 50 % gabapentin (green).
Capsules containing pure gabapentin (G100) and capsules containing the ratio 70-30 % gabapentin and pregabalin (P30G70) were analyzed and their PCA scores plots evaluated. Figure 33 represents the PCA scores plots from the ten repeatability analyses. Two groups can be visualized, G100 in grey and P30G70 in green.

Figure 33. PCA scores plot for repeatability analysis of capsules containing pure gabapentin (green) and capsules containing the ratio 70-30 % gabapentin and pregabalin (grey).
Appendix 3. Intermediate precision/ruggedness
The intermediate precision was studied by analyzing 80 capsules containing 50 % pregabalin (P50) and 10 capsules containing 50 % gabapentin (G50) for ten days straight with the wavelength range 1650-1750 nm. The PCA scores plots from the ten days are presented in Figure 34, P50 shown in green and G50 in pink.

Figure 34. PCA scores plot for intermediate precision analysis of capsules containing 50 % pregabalin (green) and capsules containing 50 % gabapentin (pink).
Appendix 4. Selectivity/robustness

The selectivity study was performed by analyzing 80 capsules containing 50% pregabalin (P50) and 10 capsules containing 50% gabapentin (G50) ten times on the same day with the capsules in different positions on the tray each time. The wavelength range for analysis was set to 1650-1750 nm. Figure 35 represents the PCA scores plot of the ten analyses, P50 shown in pink and G50 in green.

Figure 35. PCA scores plot for selectivity analysis of capsules containing 50% pregabalin (pink) and capsules containing 50% gabapentin (green).
The selectivity study was also performed by analyzing two capsules containing pure gabapentin (G100, pink) and ten capsules containing the ratio 70-30 % gabapentin and pregabalin (P30G70, green). See Figure 36.

Figure 36. PCA scores plot for selectivity analysis of capsules containing pure gabapentin (pink) and capsules containing the ratio 70-30 % gabapentin and pregabalin (green).
Appendix 5. Method comparison

Figure 37 illustrates the preprocessed spectra and PCA scores plot for three of the seizures. In A1 and A2 the green group represents gabapentin and the grey pregabalin, in B1 and B2 the green group represents pregabalin and the grey gabapentin and in C1 and C2 the green group represents one row on the tray since the capsules were homogeneous, all containing pregabalin.

Figure 37. Preprocessed spectra and PCA scores plot after analysis in the wavelength range 1650-1750 nm of four seizures; A1 and A2 (gabapentin green), B1 and B2 (pregabalin green), C1 and C2 (homogeneous seizure, one row green).