

Construction and Validation of Quantification Methods for Determining the Cannabidiol Content in Liquid Pharma-Grade Formulations by Means of Near-Infrared Spectroscopy and Partial Least Squares Regression

Joan Espel Grekopoulos

AI LAB Swiss AG, Wald-Schönengrund, Switzerland

Keywords

Near-infrared spectroscopy · European Medicines Agency · Good Manufacturing Practices · Good Laboratory Practices · Partial least squares regression · Cannabidiol · Quality control

Abstract

There is an increasing interest in cannabinoids as they are being proved to effectively treat the symptoms of a variety of medical conditions. Commercialization of cannabinoid-based pharmaceutical products is expected to grow in the near future, favored by the recent changes in medical regulations in many developed countries. Hence, robust and reliable analytical methods for determining the content of the active pharmaceutical ingredient will be needed, as this is one of the most relevant parameters for the decision to release the final pharmaceutical product into the market. The aim of this work was to demonstrate that near-infrared (NIR) spectroscopy fulfills the needed requirements for this purpose, as well as to provide a methodology to be applied to other cannabinoid-based products. We present two validated methods for the quantification of different liquid pharma-grade cannabidiol (CBD) formulations based on NIR spectroscopy and partial

least squares regression modelling. The methods were constructed and validated with spectra belonging both to production samples and to laboratory samples specifically made for this purpose, and they fulfill European Medicines Agency and International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use guideline requirements. These methods allow determining the CBD content with results comparable to the usual method of choice while saving reagent- as well as time-related costs.

© 2019 The Author(s)

Published by S. Karger AG, Basel

Introduction

The recent discovery of the endocannabinoid system (ECS) as a regulator of cellular homeostasis has pushed cannabinoids into a central position from the medical and drug research point of view. This role of the ECS potentially opens many doors to preventing and treating many diseases related to the ECS, such as Alzheimer's disease [1], obesity [2, 3], multiple sclerosis [4, 5], and epilepsy [6], among many others.

KARGER

E-Mail karger@karger.com
www.karger.com/mca

© 2019 The Author(s)
Published by S. Karger AG, Basel

Karger
Open access

This article is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND) (<http://www.karger.com/Services/OpenAccessLicense>). Usage and distribution for commercial purposes as well as any distribution of modified material requires written permission.

Joan Espel Grekopoulos
AI LAB Swiss AG
Tüfi 450
CH-9105 Wald-Schönengrund (Switzerland)
E-Mail j.espel@ailabswiss.ch

Cannabidiol (CBD) is one of the major cannabinoids found in the *Cannabis sativa* L. plant, and, in contrast to tetrahydrocannabinol, it is not psychoactive. Nonetheless, many potential pharmacological activities of CBD are being investigated at the stage of clinical trials regarding their antipsychotic, antidepressant, anxiolytic, and anti-inflammatory [7] as well as their anticancer properties [8].

These features permit the use of CBD as a versatile active pharmaceutical ingredient (API), and, thus, strict control of the CBD content in CBD-based drugs is of the utmost importance for patients to ensure correct dosage of the drug. In addition, all steps needed for the purification and production of new pharmaceutical products based on this molecule need to be aligned with the guidelines used in the pharmaceutical industry.

In this sense, advances in near-infrared (NIR) spectroscopy instrumentation and data processing have reached a point where this technique is becoming a realistic alternative to traditional analytical options used in pharmaceutical quality control during the whole process, i.e., from the incoming raw materials and process monitoring to the outgoing final product.

The advantages of NIR spectroscopy lie in its non-destructive, noninvasive, and reagent-free nature; once a method has been developed and settled into routine, the analysis of a sample takes around 30 s to complete, and almost no preparation of the sample (e.g., ultrasound, dilutions, etc.) is required regardless of the drug's presentation or physical form.

NIR spectroscopy is based on the sample's absorption of light at a wavelength of 800–2,500 nm, i.e., wavenumbers range from 12,500 to 4,000 cm^{-1} . This interaction is a consequence of the vibrational features of the sample and results in overtones and combination bands that build up the resulting spectrum. The spectrum is thus sensitive to the physical and chemical properties of the sample, and this relationship can be mathematically modelled by multivariate regression methods.

Partial least squares (PLS) regression belongs to the so-called linear methods, which also include classical least squares, inverse least squares, and principal components regression. Despite the variety of mathematical methods, PLS regression is implemented and used most often in the context of NIR spectroscopy.

Although previous studies have reported the successful use of NIR spectroscopy and PLS regression to determine different cannabinoids in *Cannabis sativa* L. plant material [9, 10], to date no reports have been found describing the quantification of a cannabinoid in a final pharmaceutical product and its validation consistent with the Euro-

pean Medicines Agency (EMA) [11] and the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) [12] guidelines.

Materials and Methods

Scope of the Methods

In the present study, two quantification methods based on NIR spectroscopy were developed for 12 different CBD-based liquid formulations (Table 1). These pharmaceutical products are made and commercialized in Switzerland by a GMP (Good Manufacturing Practices)-certified company.

As can be seen in Table 1, a single NIR spectroscopy method can be used for different doses and flavors, given that the solvent, i.e., the major excipient (and the only one in the flavorless formulations), is the same.

All measurements related to the development and validation of the methods were carried out at the facilities of AI LAB Swiss AG, while the NIR spectroscopy system (including software, cuvettes, and a translector) was kindly provided by Büchi Labortechnik AG (Flawil, Switzerland) for the purposes of this study.

Preparation of Laboratory Samples

Our approach is based on that used by Blanco et al. [13], where laboratory samples at different API content levels are prepared with the aim of obtaining a wide enough concentration range. This practical and straightforward approach permits including variability due to concentration through laboratory samples, while variability in the production process is incorporated through the inclusion of production samples.

A total of 30 laboratory samples at 15 different concentration levels were prepared for each NIR spectroscopy method by making dilutions of a stock solution containing 25% of CBD in each solvent. Both the stock solutions and the dilutions were made by weighing on an analytical balance (Mettler Toledo, model XPE-205DR) and proper mixing and homogenization.

The concentration range of the laboratory samples spanned the whole nominal concentration range of the commercial products (i.e., 5–15% m/m) plus an additional 30% of the nominal concentration at both ends of the calibration curve, yielding a final range of 3.5–19.5%.

No flavoring (mint or orange) was added to the laboratory samples. The spectral variability due to these substances was included in the methods in a natural way through inclusion in both the calibration set (C-set) and the validation set (V-set) of production samples containing these flavors.

Reference Data for the Production Samples

The reference data for the production samples were obtained by solution of the CBD to be analyzed in methanol/methylene chloride. The CBD content was determined by high-performance liquid chromatography (HPLC; Agilent 1100 Series equipped with a Nucleosil 120-3 C8 column) separation and subsequent analysis by UV/Vis absorption spectroscopy.

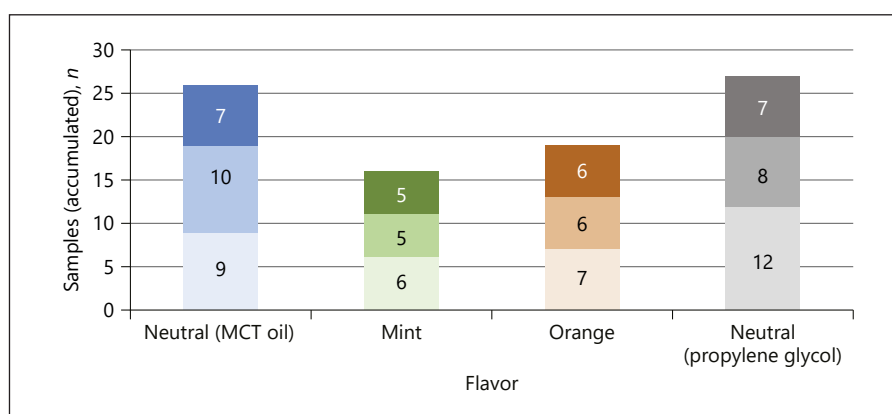
Quantification was carried out by three-point calibration using the CBD standard from Lipomed AG (Arlesheim, Switzerland), while identification was performed by comparison of diode array detector spectra. The reference method is properly validated according to the ICH guidelines.

Table 1. Main features of used products and associated NIR spectroscopy methods

| Product No. | Commercial name | CBD concentration, % m/m | Flavoring concentration, % m/m | Solvent | NIR spectroscopy method |
|-------------|--------------------------|--------------------------|--------------------------------|------------------|-------------------------|
| 1 | CBD Öl-Tropfen (Neutral) | 5 | – | MCT oil | 1 |
| 2 | | 10 | | | |
| 3 | | 15 | | | |
| 4 | CBD Öl-Tropfen (Orange) | 5 | Orange (0.9) | | |
| 5 | | 10 | | | |
| 6 | | 15 | | | |
| 7 | CBD Öl-Tropfen (Minze) | 5 | Mint (0.7) | | |
| 8 | | 10 | | | |
| 9 | | 15 | | | |
| 10 | CBD wasserlöslich | 5 | – | Propylene glycol | 2 |
| 11 | | 10 | | | |
| 12 | | 15 | | | |

NIR, near-infrared; CBD, cannabidiol; MCT, medium-chain triglyceride.

Fig. 1. Distribution of measured production samples. Color intensity indicates dose: for each flavor, the bottom part of the column displays the number of samples with a 5% CBD dose, the middle part those with a 10% CBD dose, and the upper part those with a 15% CBD dose. CBD, cannabidiol; MCT, medium-chain triglyceride.



In all, 61 production samples based on the medium-chain triglyceride (MCT) oil formulation and 27 production samples based on the propylene glycol formulation were analyzed by HPLC and stored under controlled conditions for their subsequent analysis by NIR spectroscopy. The distribution of the samples ensured that all doses and flavors were included (Fig. 1).

Spectrum Collection

A sufficient amount of sample was placed in a cylindrical optical glass cuvette (Hellma Analytics, model 692.091-OG). The use of a hard chrome plate steel transmittance cover (Büchi Labortechnik, Flawil, Switzerland) fitting the diameter of the cuvette ensured the same optical path length (0.3 mm) and thus results independent of the sample amount, given that the amount was sufficient.

Each sample (including both laboratory and production samples) was measured 3 consecutive times. Between each of these 3

consecutive measurements, the cell and the translector were moved with the aim of maximizing variability related to the preparation of the sample.

A measurement was an average of 16 scans with a resolution of 4 cm⁻¹. These measurements were performed on Büchi's NIRFlex N-500 Fourier Transform NIR Spectrophotometer along with the "Solids" measurement module. This module can be used to measure solids by diffuse reflectance, but the use of a translector, where infrared light is reflected after passing through the sample and crosses it again towards the sensor, de facto acts as a transmittance measurement. The measurement range of the instrument spans 4,000–10,000 cm⁻¹.

The NIR spectroscopy instrument was run by the "BUCHI Operator" application, included in the NIRWare software suite. This software periodically carried out the machine's system suitability tests, as well as internal and external reference tests, ensuring the quality of the samples' spectra.

Data Processing

Once the spectra of all samples were collected, each reference value was assigned to each corresponding spectrum by means of the NIRWare Software Package.

As the number of available samples was large enough, a cross-validation scheme was not necessary, and thus the spectra were split into two sets for each model: (1) a C-set, used to construct the model, and (2) a V-set, used to test and validate the model. This separation was made such that each set contained approximately the same number of samples. Besides, the selection was made in an alternating way, thereby ensuring that the whole range of variability was included in both sets: flavor, dose, and laboratory as well as production origin.

Since each sample had three associated spectra, the split was made into groups of three spectra (i.e., blockwise) to guarantee that all three spectra belonging to the same sample were either in the C-set or in the V-set. This guaranteed that batches included in one set (i.e., the C-set) were different from those included in the other set (i.e., the V-set).

Construction of the PLS Regression Models

The NIRCal software, included in Büchi's aforementioned software package, was used to construct two different PLS regression models: one for the MCT oil-based formulations and another for the propylene glycol-based ones. Since no outliers were detected, no spectra were deleted.

Regarding the mathematical model that predicts the concentration from a given spectrum, PLS regression was the method of choice. The PLS algorithm decomposes the spectral matrix into PLS factors or components with the objective of finding the best correlation between the spectra and the reference values. The number of PLS factors is therefore an important parameter to be set, as it determines the characterization of the data set. In addition, an optimal spectral range (i.e., the considered wavenumbers) and the pretreatment of the spectra (in order to increase the signal-to-noise ratio) have to be chosen based on obtaining the best predictive capacity of the model.

To ensure that the models quantified the analyte (CBD) and not the absence of placebo, the first loading for each model was plotted (Fig. 2, 3). It can be seen that the wavelengths that retained most of the information coincided among them as well as with the active regions of CBD (see reflectance of CBD in Fig. 6, 7).

The main quality parameters used to evaluate the model were the coefficient of determination, R^2 , and the standard error of estimation, SEE (referred to as SEC when calculated for the C-set and SEP for the V-set):

$$SEE = \sqrt{\frac{1}{m} - \sum_{i=1}^m (y_{p,i} - y_{r,i} - BIAS)^2},$$

where $y_{p,i}$ and $y_{r,i}$ refer to the predicted and the reference value of spectrum i , respectively, m is the total number of data points, and $BIAS$ is the bias of the plot concerning predicted versus reference values, and

$$R^2 = \frac{\sum_{i=1}^m (y_{p,i} - \bar{y}_r)^2}{\sum_{i=1}^m (y_{r,i} - \bar{y}_r)^2},$$

where $y_{p,i}$ and $y_{r,i}$ refer to the predicted and reference value of spectrum i , respectively, and \bar{y}_r is the mean value of the reference values.

R^2 and SEE apply to both the C-set and the V-set and are non-redundant parameters. Their value in both sets determines the optimal parameters (wavenumber range, pretreatment, and the number of PLS factors) for each model.

Additionally, NIRWare allows a qualitative identity check of the measured sample based on comparing the residual of the measured spectrum with the average spectrum of all spectra composing the PLS quantification model. This identity check prevents the analyst from quantifying CBD in a drug that looks similar to the product to be measured.

Validation of the Methods

To properly validate the methods according to the aforementioned guidelines, the following parameters need to be assessed: specificity, linearity, accuracy, precision, and robustness. Also, the standard error of laboratory (SEL) of the reference method needs to be determined and compared to the SEP .

The selectivity of the methods was assessed by measurement of out-of-scope samples such as placebo and other liquid formulations. The software is expected to reject these samples, and thus the value of CBD content is not valid.

Linearity was evaluated by performing a Passing and Bablok [14, 15] regression between the predicted values and the reference values and testing that the slope and the intercept were statistically equivalent to 1 and 0, respectively, for a 95% level of significance.

Although this approach is not the usual one, compared to others such as lack of fit [16], it permits us to compare two methods. Since in this case the other method was HPLC, which is validated and has been shown to be linear, this allowed testing for NIR spectrum linearity, as well as stating that the NIR spectroscopy method is free from systematic errors and matrix effects.

Accuracy evaluations were performed for each method both at a given nominal concentration and along the whole concentration range. The predicted values were compared to the reference ones by a paired t test at a 95% confidence level.

Regarding precision, both repeatability and intermediate precision had to be assessed for each method. Repeatability was evaluated by having the same operator measure the same sample 6 times and calculating the relative standard deviation. Evaluation of intermediate precision was carried out by evaluating the impact of different preparations and different operators. This calculation was done by means of two-way analysis of variance at the 95% confidence level.

With respect to robustness, its evaluation over time by means of paired t tests is not possible yet, since the method has recently been developed and there is a lack of reference and NIR spectroscopy data over a long period of time. Thus, accuracy values have to be provisionally assumed to assess this parameter.

The SEL of the HPLC method can be determined by performing two different measures on different production samples:

$$SEL = \sqrt{\frac{1}{m} \sum_{i=1}^m (Differ)_i^2},$$

where m is the number of samples and $Differ$ is the difference between the two measurements.

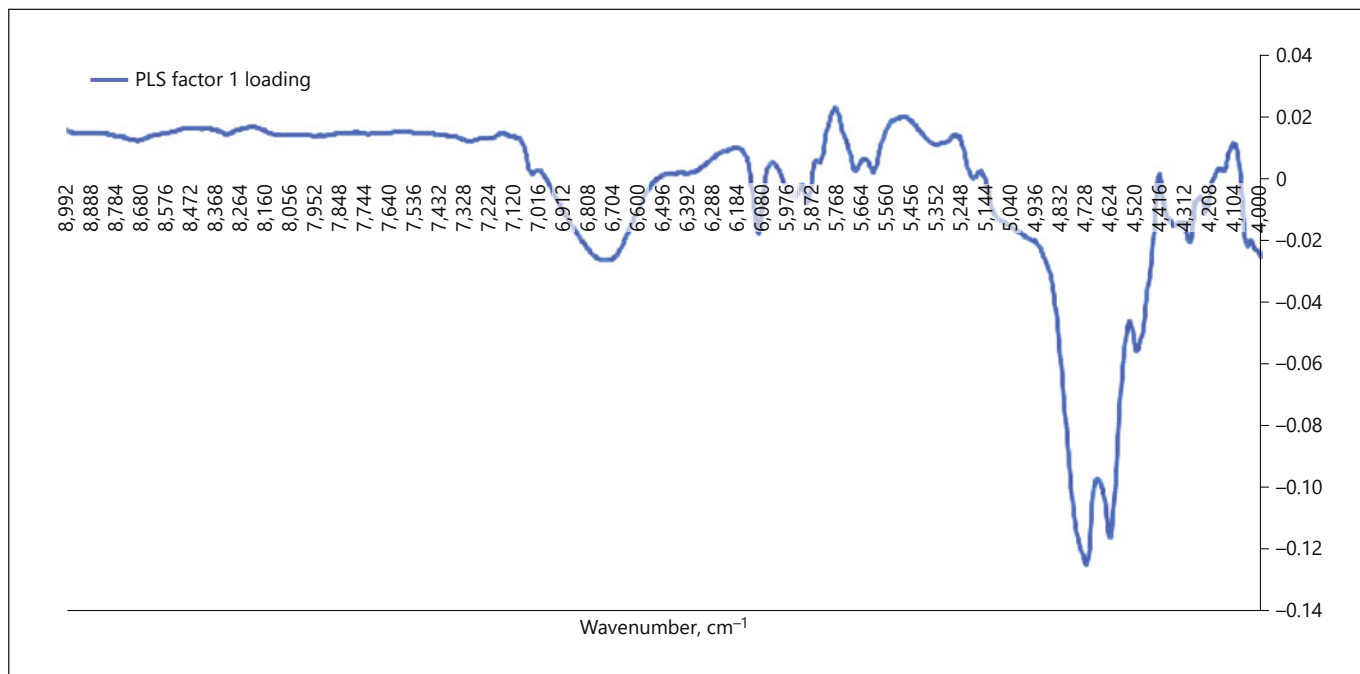


Fig. 2. Loading corresponding to the first PLS factor for the MCT oil-based formulation model. PLS, partial least squares; MCT, medium-chain triglyceride.

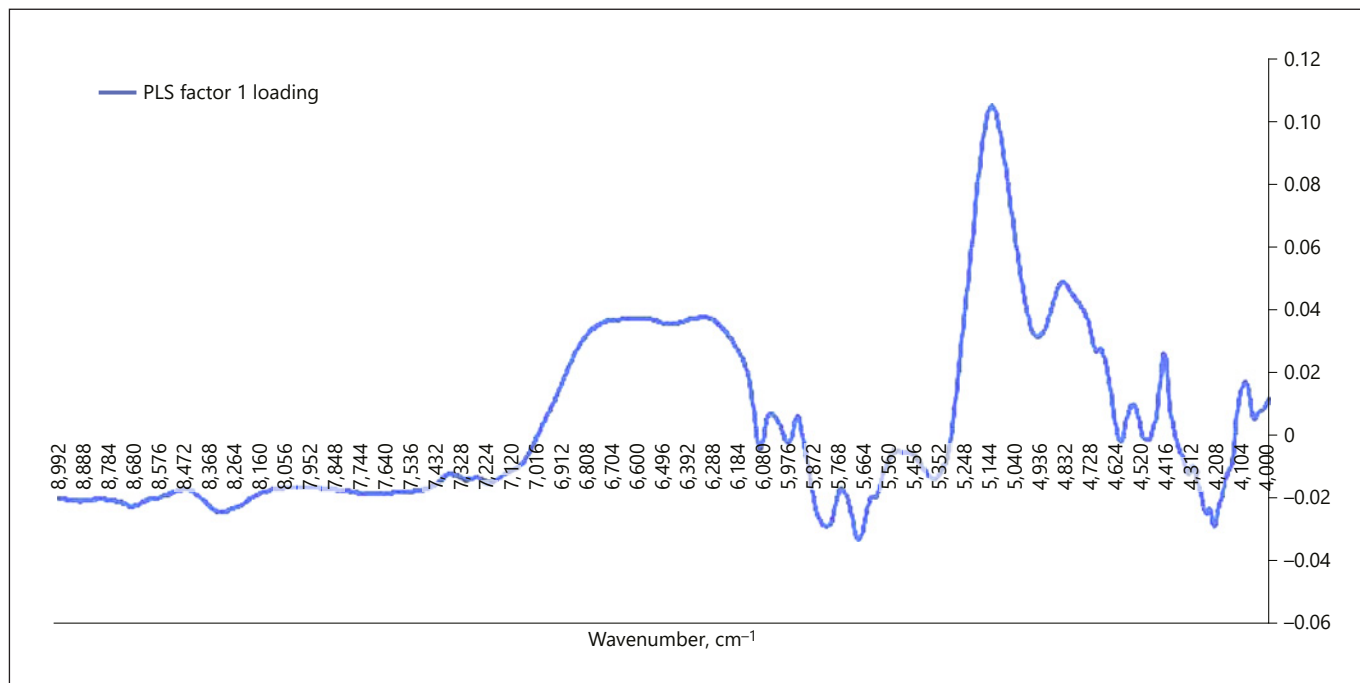


Fig. 3. Loading corresponding to the first PLS factor for the propylene glycol-based formulation model. PLS, partial least squares.

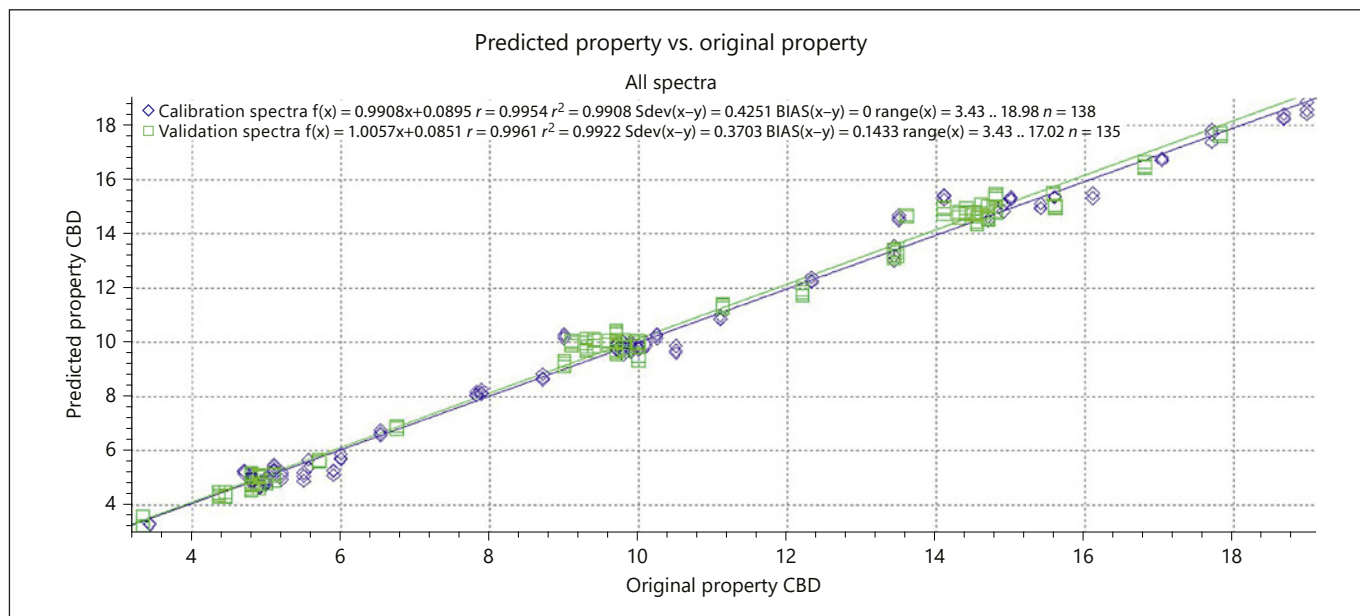


Fig. 4. Predicted versus reference values for spectra included in the MCT oil-based formulation PLS regression model. Blue: calibration set; green: validation set. CBD, cannabidiol; PLS, partial least squares; MCT, medium-chain triglyceride.

Table 2. Main parameters and properties of the constructed PLS regression models

| Property | MCT oil-based formulations | Propylene glycol-based formulations |
|--------------------------------------|----------------------------|-------------------------------------|
| Spectrum pretreatment | SNV transformation | MSC |
| Wavenumber range, cm^{-1} | 4,000–9,020 | 4,000–9,000 |
| Number of PCs | 4 | 6 |
| C-set concentration range, % m/m | 3.4–19.0 | 3.4–19.4 |
| V-set concentration range, % m/m | 3.3–17.8 | 3.6–19.5 |
| C-set number of samples ^a | 46 (15+31) | 30 (16+14) |
| C-set number of spectra | 138 | 90 |
| V-set number of samples ^a | 45 (15+30) | 27 (14+13) |
| V-set number of spectra | 135 | 81 |
| SEC, % m/m | 0.43 | 0.30 |
| SEP, % m/m | 0.37 | 0.32 |

PLS, partial least squares; MCT, medium-chain triglyceride; SNV, standard normal variate; MSC, multiplicative scatter correction; PCs, principal components; C-set, calibration set; V-set, validation set; SEC/SEP, standard error of estimation for the calibration/validation set. ^a Expressed as $x(y+z)$, where x = total number of samples, y = number of laboratory samples, and z = number of commercial samples.

Results

Calibration of the Models

The values of R^2 and SEE , as parameters related to the predictive capacity of the models, were taken as main criteria for constructing the models and choosing an appro-

priate pretreatment of the spectra and number of PLS factors. The models have been found to predict better based on the following adjustments (Table 2).

The calibration and validation plots for the two models can be seen in Figures 4 and 5.

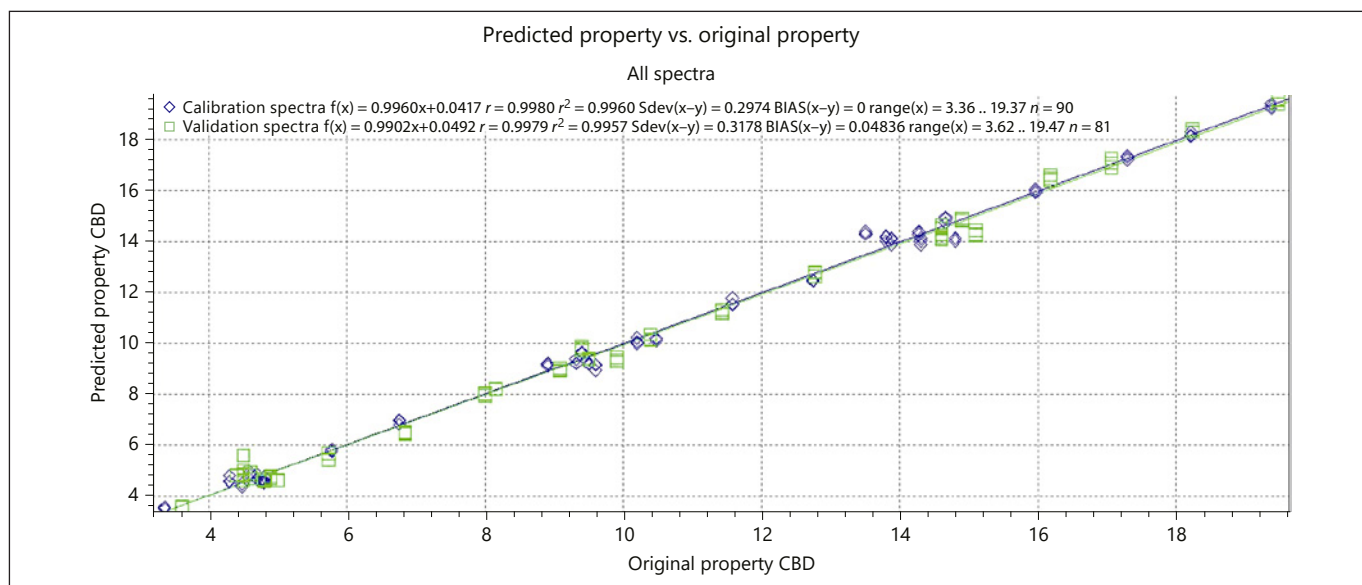


Fig. 5. Predicted versus reference values for spectra included in the propylene glycol-based formulation PLS regression model. Blue: calibration set; green: validation set. CBD, cannabidiol; PLS, partial least squares.

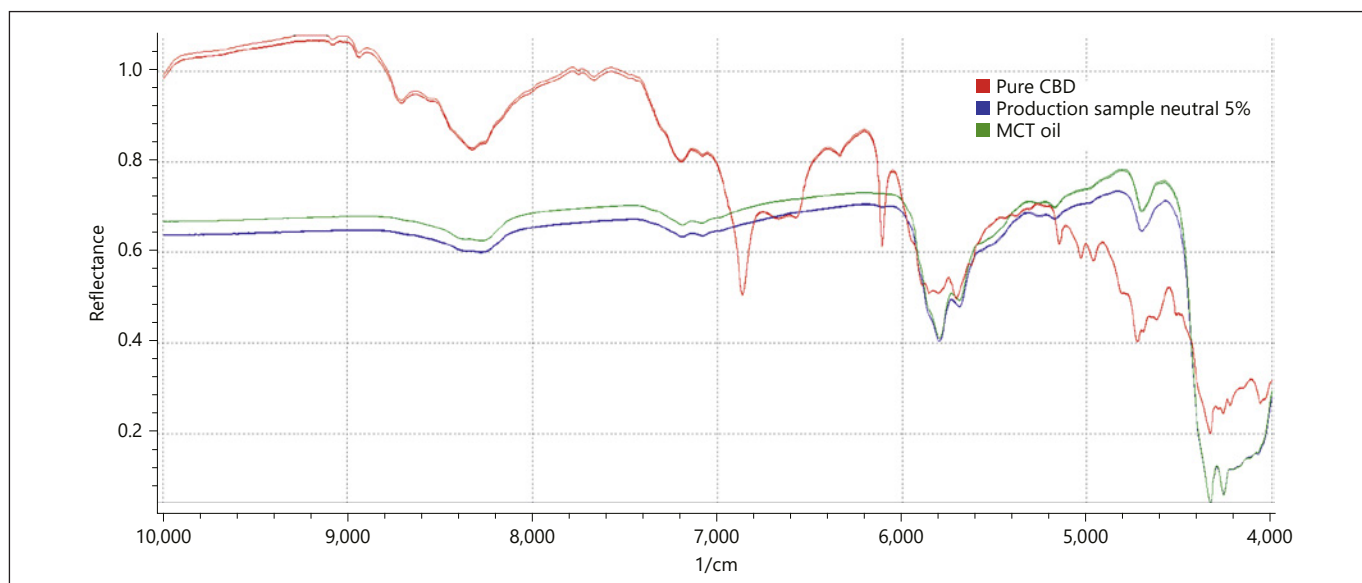


Fig. 6. Full-range, untreated spectra corresponding to the active pharmaceutical ingredient (pure CBD), placebo (MCT oil), and a 5% neutral production sample. CBD, cannabidiol; MCT, medium-chain triglyceride.

Spectral Range

Figures 6 and 7 show comparisons of the pure CBD spectra with placebo (MCT oil or propylene glycol) and a production sample at 5% nominal concentration. Note that each sample was measured in triplicate, and thus there are three spectra per sample.

As can be seen, no spectral features were available from 9,000 to 10,000 cm^{-1} , so this range was discarded right at the beginning. For the rest of the spectral range (i.e., 4,000–9,000 cm^{-1}), there are two possible approaches depending on the absorption features of the compound to be quantified: either selecting fragmented ranges around the most

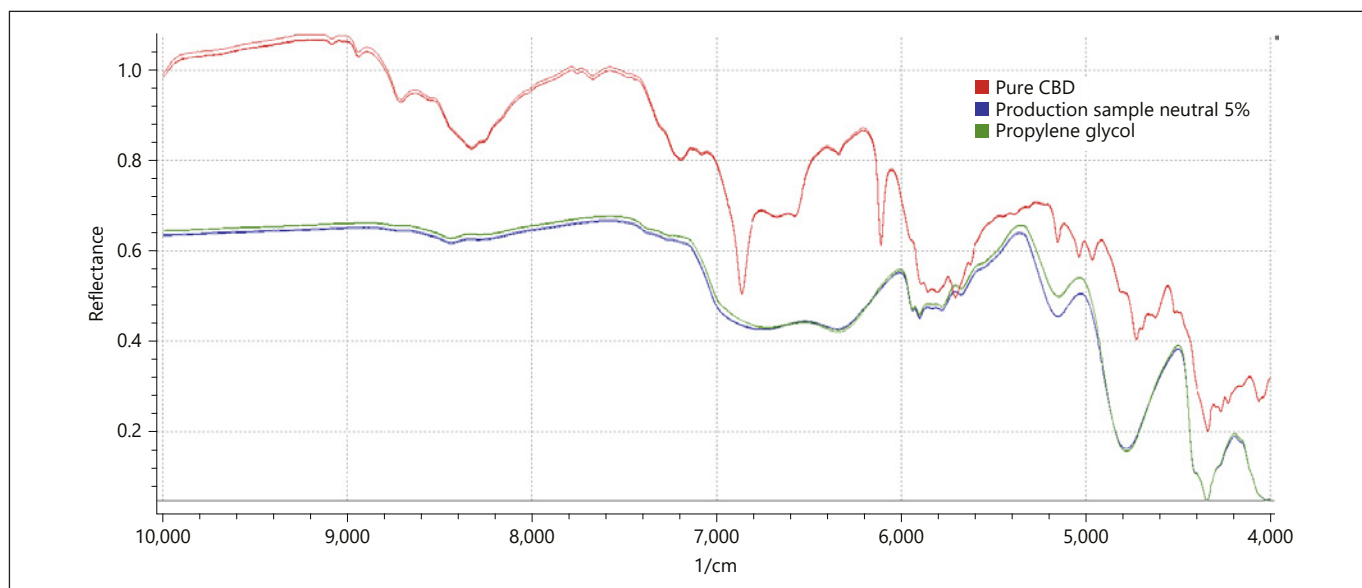


Fig. 7. Full-range, untreated spectra corresponding to the active pharmaceutical ingredient (pure CBD), placebo (propylene glycol), and a 5% neutral production sample. CBD, cannabidiol.

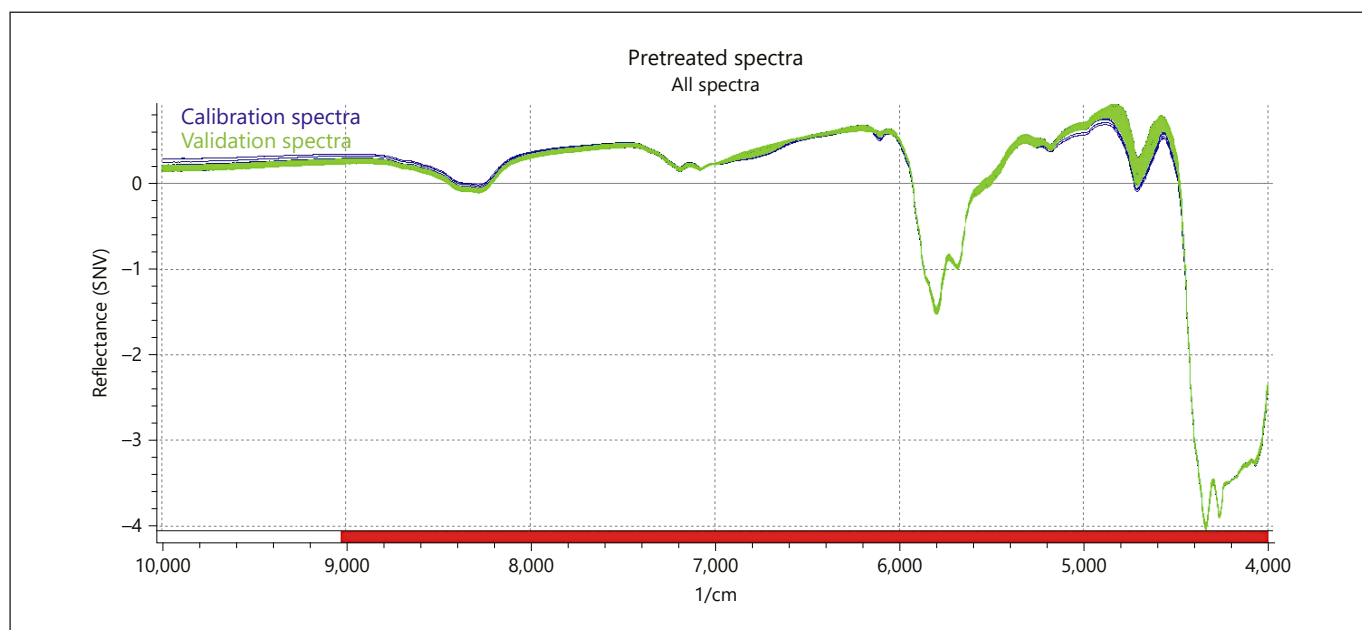


Fig. 8. SNV transformation-pretreated spectra (blue: calibration set; green: validation set) for the MCT oil-based formulation PLS regression model. Red: PLS regression model working spectral range. SNV, standard normal variate; MCT, medium-chain triglyceride; PLS, partial least squares.

intense absorption bands of the API or selecting the whole spectral range. Since CBD displays many absorption bands along all wavenumbers from 4,000 to 9,000 cm^{-1} , this range has been selected for constructing the PLS regression mod-

els. The results were better than those obtained by selecting smaller ranges around intense absorption bands of CBD.

In addition, focusing on a wide spectral range provides better selectivity for the qualitative identification of sam-

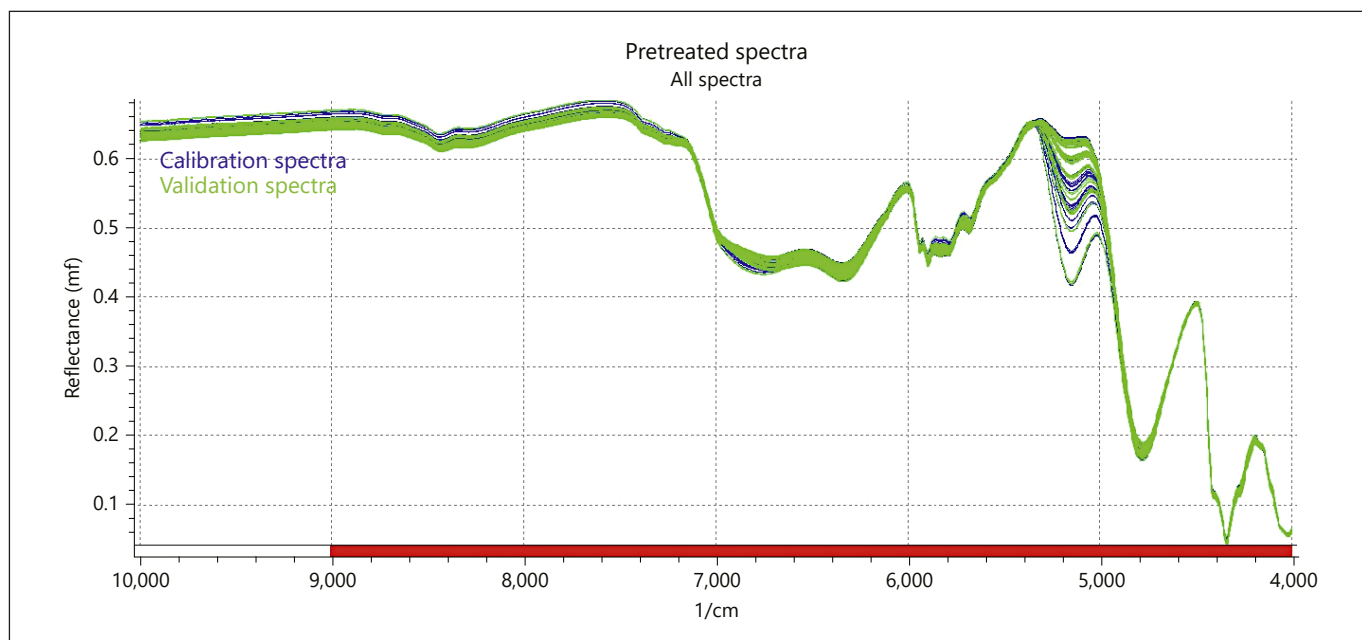


Fig. 9. Multiplicative scatter correction-pretreated spectra (blue: calibration set; green: validation set) for the propylene glycol-based formulation PLS regression model. Red: PLS regression model working spectral range. PLS, partial least squares.

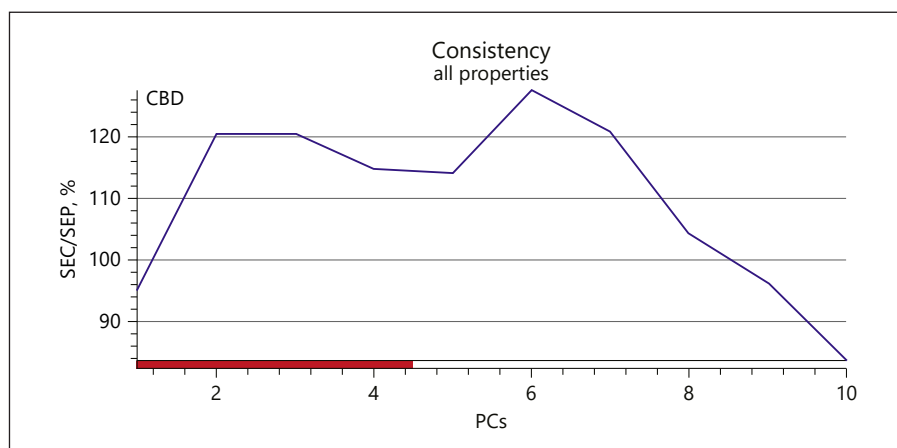


Fig. 10. Consistency versus number of PLS factors for the MCT oil-based formulation PLS regression model. Red: chosen number of PLS factors. PLS, partial least squares; CBD, cannabidiol; SEC/SEP, standard error of estimation for the calibration/validation set; PCs, principal components.

ples, since spectral features that may not be related to the API are also included in the model.

Pretreatment

Compared to other pretreatments or combinations of pretreatments, standard normal variate transformation and multiplicative scatter correction provided the best results when pretreating the spectra of the MCT oil-based and the propylene glycol-based formulations, respectively.

These methods are linearly correlated [17], but some authors have shown that they may also lead to different results [18], as happened in the case at hand.

The pretreated spectra can be seen in Figures 8 and 9. As these figures display, there is a dispersion of the spectra around $5,000\text{ cm}^{-1}$. Since baseline shifts should have been eliminated due to normalization of the pretreatment, this feature may be attributed to the increasing CBD concentrations along with a low absorbance of the solvent in that particular range.

Table 3. Specificity of the calibrated PLS regression models

| PLS regression model | Sample | Expected result | Obtained result | Failed test |
|-------------------------------------|-------------------------|-----------------|-----------------|--------------------|
| MCT oil-based formulations | Placebo | NO OK | NO OK | Range |
| | E-liquid 0.5% CBD | NO OK | NO OK | Range and residual |
| | MCT oil 10% CBD | NO OK | NO OK | Range and residual |
| Propylene glycol-based formulations | Placebo | NO OK | NO OK | Range |
| | E-liquid 0.5% CBD | NO OK | NO OK | Range and residual |
| | Propylene glycol 5% CBD | NO OK | NO OK | Range and residual |

If one of the tests (residual- or range-based) fails, NO OK is reported. PLS, partial least squares; MCT, medium-chain triglyceride; CBD, cannabidiol.

Table 4. Linearity parameter fulfillment of the calibrated PLS regression models

| PLS regression model | Property | Obtained result (value ± confidence interval) | Fulfills requirement? |
|-------------------------------------|-----------------|---|-----------------------|
| MCT oil-based formulations | C-set slope | 0.991±0.016 | Yes |
| | C-set intercept | 0.090±0.173 | Yes |
| | C-set R^2 | 0.991 | – |
| | V-set slope | 1.006±0.015 | Yes |
| | V-set intercept | 0.085±0.170 | Yes |
| | V-set R^2 | 0.992 | – |
| Propylene glycol-based formulations | C-set slope | 0.996±0.013 | Yes |
| | C-set intercept | 0.042±0.152 | Yes |
| | C-set R^2 | 0.996 | – |
| | V-set slope | 0.990±0.015 | Yes |
| | V-set intercept | 0.049±0.160 | Yes |
| | V-set R^2 | 0.996 | – |

PLS, partial least squares; MCT, medium-chain triglyceride; C-set, calibration set; V-set, validation set.

Fig. 11. Consistency versus number of PLS factors for propylene glycol based formulations PLS regression model. Red: chosen number of PLS factors. PLS, partial least squares; CBD, cannabidiol; SEC/SEP, standard error of estimation for the calibration/validation set; PCs, principal components.

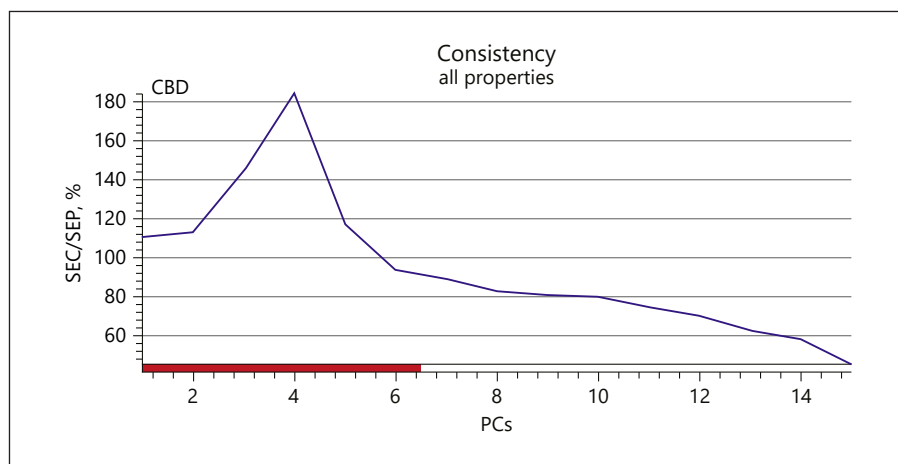


Table 5. Accuracy evaluation of the calibrated PLS regression models

| PLS regression model | Property | Nominal concentration or concentration range | Obtained <i>t</i> value | Critical <i>t</i> value | Fulfills requirement? |
|-------------------------------------|-------------------------|--|-------------------------|-------------------------|-----------------------|
| MCT oil-based formulations | Target level accuracy | 10% | 2.101 | 2.262 | Yes |
| | Over-the-range accuracy | 4.8–15.6% | 0.629 | 2.571 | Yes |
| Propylene glycol-based formulations | Target level accuracy | 10% | 0.639 | 2.306 | Yes |
| | Over-the-range accuracy | 4.4–14.6% | 0.135 | 2.571 | Yes |

PLS, partial least squares; MCT, medium-chain triglyceride.

Table 6. Repeatability evaluation of the calibrated PLS regression models

| PLS regression model | Obtained RSD, % | Threshold, % | Fulfills requirement? |
|-------------------------------------|-----------------|--------------|-----------------------|
| MCT oil-based formulations | 0.4 | 2 | Yes |
| Propylene glycol-based formulations | 1.1 | 2 | Yes |

PLS, partial least squares; RSD, relative standard deviation; MCT, medium-chain triglyceride.

Number of PLS Factors

The appropriate number of PLS factors was chosen based on an equilibrium between minimizing the values of *SEC* and *SEP* and maximizing R^2 . A reasonably low number of PLS factors (1–9) should be chosen to avoid overfitting the PLS regression models.

The selection of the model's number of PLS factors is also supported by consistency, defined as the ratio between the *SEC* and the *SEP*. This parameter should be around 100% (Fig. 10, 11) to ensure an optimal balance between the C-set and the V-set.

An optimal compromise for the validation parameters, without overfitting the model, has been found to be 4 and 6 factors for the MCT oil-based and the propylene glycol-based formulations, respectively.

Validation of the Models

The following sections display the validation results following the approach described in the EMA guidelines [11].

Specificity

The specificity of the method was assessed by demonstrating that samples outside the method's scope were rejected (e.g., labeled as NO OK by the software).

For this purpose, the software features two different tools: an identity test based on comparing the residual of

the sample with the average spectrum of the spectra used to calibrate the model, and a test to discard samples whose predicted values are outside the model's range. For a sample's CBD to be quantified, both tests must be OK.

In order to test the selectivity of the calibrated models, samples consisting of placebo and different formulations produced at AI LAB Swiss AG were measured (Table 3). As can be seen, the method was selective regarding samples out of range.

Linearity and Range

Table 4 shows that the calibrated models presented slope and intercept values – both for the C-set and the V-set – that were statistically equal to 1 and 0, respectively, for a 95% confidence level. Also, R^2 values were close to 1.

Accuracy

Accuracy was assessed by comparing the reference and predicted values of a representative set of commercial samples taken from the V-set. Evaluation was done by means of a paired *t* test for a 95% confidence level. The test was performed at one of the product's nominal concentrations on the one hand (target level accuracy, assessed with 10 spectra) and by comparing reference and predicted values' spread along the validation line on the other hand (accuracy over the range, assessed with 6 spectra) (Table 5).

Table 7. Evaluation of the intermediate precision of the calibrated PLS regression models

| PLS regression model | Variation | Obtained <i>F</i> value | Critical <i>F</i> value | Fulfills requirement? |
|-------------------------------------|--------------------|-------------------------|-------------------------|-----------------------|
| MCT oil-based formulations | Sample preparation | 0.33 | 19.00 | Yes |
| | Analyst | 1.33 | 18.51 | Yes |
| Propylene glycol-based formulations | Sample preparation | 0.33 | 19.00 | Yes |
| | Analyst | 7.41 | 18.51 | Yes |

PLS, partial least squares; MCT, medium-chain triglyceride.

Precision: Repeatability

Repeatability was assessed by having the same operator repeat the same measurement of a production sample 6 times in a row (Table 6) and calculating the relative standard deviation.

Precision: Intermediate Precision

The methods were demonstrated to provide statistically equivalent results with different sample preparations and analysts. For assessment of intermediate precision, 3 measurements stemming from the same sample (production sample) were performed by 2 different operators. The obtained values were evaluated by means of two-way analysis of variance at the 95% level of significance (Table 7).

Robustness

Accuracy results are assumed to show the robustness of a method. As new samples will be analyzed in the future, data on these will be used to assess the robustness of the method over time.

Standard Error of Laboratory

The EMA guidelines do not provide any critical value or threshold for the *SEL*, but they recommend its determination in order to put the *SEP* into perspective. The *SEL* of the reference method was determined from the HPLC results of 3 different production samples measured twice, with a new sample preparation per measurement. A value of 0.128 was obtained.

The *SEP/SEL* quotient equals 2.9 for the MCT oil formulation model and 2.5 for the propylene glycol model. Values of *SEP* are usually twice the *SEL* [19]. The slightly higher values obtained for the calibrated models may be due to inclusion of a larger amount of variability (concentration and flavor) in a single method, while most of the methods described in the literature are oriented towards a single final product at a single nominal concentration.

Discussion and Conclusions

Two methods based on NIR spectroscopy and subsequent PLS regression modelling have been developed and validated for the release of pharma-grade CBD-based liquid formulations. These methods allow direct measurement (no sample dissolution, dilution, or any volumetric operations are needed) and quick quantification of the API.

The methods have been shown to be selective and to cover a range of different concentrations, each corresponding to different doses, i.e., different products. In addition, the inclusion of samples with different flavors in the C-set and the V-set has allowed further extending the range of products to which the method may be applied.

The validation has shown that the NIR spectroscopy predictions are comparable to the reference method and allow for the market release of liquid CBD preparations using MCT oil as well as propylene glycol containing 5, 10, or 15% CBD.

All in all, it has been shown that NIR spectroscopy is a great candidate for the quality control of cannabis-based pharmaceutical products because of its versatility and efficiency.

Acknowledgements

The author would like to kindly thank: AI LAB Swiss AG and its team for granting access to the facilities and laboratory equipment and for the scientific support during the research; Büchi Labortechnik AG and the designated persons of contact for providing the NIR spectroscopy-related instrumentation and material as well as technical and scientific support; GHM Genetic Development SL for enhancing the collaboration with AI LAB Swiss AG and financially supporting the project.

Statement of Ethics

The author has no ethical conflicts to disclose.

Disclosure Statement

The author has no conflicts of interest to declare. The author worked for GHM Genetic Development SL at the time of this project. GHM Genetic Development SL is a company committed to the study of the therapeutic potential of cannabinoids.

Funding Sources

AI LAB Swiss AG assumed laboratory-derived expenses: reagents and analytical equipment; Büchi Labortechnik AG provided the spectrophotometer, transflector, cuvettes, and software; GHM Genetic Development SL took over the transport and accommodation costs of the author.

References

- 1 Aso E, Ferrer I. Cannabinoids for treatment of Alzheimer's disease: moving toward the clinic. *Front Pharmacol*. 2014 Mar;5:37.
- 2 Bisogno T, Mahadevan A, Coccorello R, Chang JW, Allarà M, Chen Y, et al. A novel fluorophosphonate inhibitor of the biosynthesis of the endocannabinoid 2-arachidonoylglycerol with potential anti-obesity effects. *Br J Pharmacol*. 2013 Jun;169(4):784–93.
- 3 Engeli S, Lehmann AC, Kaminski J, Haas V, Janke J, Zoerner AA, et al. Influence of dietary fat intake on the endocannabinoid system in lean and obese subjects. *Obesity (Silver Spring)*. 2014 May;22(5):E70–6.
- 4 Feliú A, Moreno-Martet M, Mecha M, Carrillo-Salinas FJ, de Lago E, Fernández-Ruiz J, et al. A Sativex[®]-like combination of phytocannabinoids as a disease-modifying therapy in a viral model of multiple sclerosis. *Br J Pharmacol*. 2015 Jul;172(14):3579–95.
- 5 Kozela E, Juknat A, Gao F, Kaushansky N, Coppola G, Vogel Z. Pathways and gene networks mediating the regulatory effects of cannabidiol, a nonpsychoactive cannabinoid, in autoimmune T cells. *J Neuroinflammation*. 2016 Jun;13(1):136.
- 6 Reddy DS. The utility of cannabidiol in the treatment of refractory epilepsy. *Clin Pharmacol Ther*. 2017 Feb;101(2):182–4.
- 7 Rong C, Lee Y, Carmona NE, Cha DS, Ragguett RM, Rosenblat JD, et al. Cannabidiol in medical marijuana: research vistas and potential opportunities. *Pharmacol Res*. 2017 Jul;121:213–8.
- 8 Pisanti S, Malfitano AM, Ciaglia E, Lamberti A, Ranieri R, Cuomo G, et al. Cannabidiol: state of the art and new challenges for therapeutic applications. *Pharmacol Ther*. 2017 Jul;175:133–50.
- 9 Borille BT, Marcelo MC, Ortiz RS, Mariotti KC, Ferrão MF, Limberger RP. Near infrared spectroscopy combined with chemometrics for growth stage classification of cannabis cultivated in a greenhouse from seized seeds. *Spectrochim Acta A Mol Biomol Spectrosc*. 2017 Feb;173:318–23.
- 10 Sánchez-Carnerero Callado C, Núñez-Sánchez N, Casano S, Ferreiro-Vera C. The potential of near infrared spectroscopy to estimate the content of cannabinoids in *Cannabis sativa* L.: a comparative study. *Talanta*. 2018 Dec;190:147–57.
- 11 European Medicines Agency. [Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations](#). EMEA/CHMP/CVMP/QWP/17760/2009 Rev1. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-use-near-infrared-spectroscopy-pharmaceutical-industry-data-requirements-new-submissions_en.pdf.
- 12 ICH. Q2 (R1), "Validation of analytical procedures: text and methodology", ICH Harmonised Tripartite Guideline. Chicago, USA: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2005. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1_Guideline.pdf.
- 13 Blanco M, Coello J, Iturriaga H, Maspoch S, de la Pezuela C. Strategies for constructing the calibration set in the determination of active principles in pharmaceuticals by near infrared diffuse reflectance spectrometry. *Analyst (Lond)*. 1997;122(8):761–5.
- 14 Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem*. 1983 Nov;21(11):709–20.
- 15 Passing H, Bablok W. Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part II. *J Clin Chem Clin Biochem*. 1984 Jun;22(6):431–45.
- 16 Araujo P. Key aspects of analytical method validation and linearity evaluation. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2009 Aug;877(23):2224–34.
- 17 Dhanoa MS, Lister SJ, Sanderson R, Barnes RJ. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. *J Near Infrared Spectrosc*. 1994;2(1):43–7.
- 18 Fearn T, Riccioli C, Garrido-Varo A, Guerrero-Ginel JE; FEARN. On the geometry of SNV and MSC. *Chemom Intell Lab Syst*. 2009;96(1):22–6.
- 19 Míka V, Pozdíšek J, Tillmann P, Nerušil P. Development of NIR calibration valid for two different grass sample collections. *Czech J Anim Sci*. 2003;48(10):419–24.