Journal of Pharmaceutical and Biomedical Analysis 202 (2021) 114150

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

New perspective for the in-field analysis of cannabis samples using handheld near-infrared spectroscopy: A case study focusing on the determination of Δ^9-tetrahydrocannabinol

Riccardo Deidda [a,*,1], Florentin Coppey [b,1], Dhouha Damergi [c,d], Cédric Schelling [c,d], Laureen Coïc [a], Jean-Luc Veuthey [c,d], Pierre-Yves Sacré [a], Charlotte De Bleye [a], Philippe Hubert [a], Pierre Esseiva [b], Éric Ziemons [a]

[a] University of Liège (ULiège), CIRM, Vibra-Santé HUB, Laboratory of Pharmaceutical Analytical Chemistry, B36 Tower 4 Avenue Hippocrate 15, 4000, Liège, Belgium
[b] University of Lausanne, School of Criminal Justice, 1015, Lausanne, Switzerland
[c] School of Pharmaceutical Sciences, University of Geneva, CMU - Rue Michel Servet 1, 1211, Geneva 4, Switzerland
[d] Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, CMU – Rue Michel Servet 1, 1211, Geneva 4, Switzerland

ARTICLE INFO

Article history:
Received 15 March 2021
Received in revised form 6 May 2021
Accepted 10 May 2021
Available online 19 May 2021

Keywords:
Cannabinoids
THC
Cannabis
Handheld
Near infrared spectroscopy
NIR

ABSTRACT

The aim of the present study was to explore the feasibility of applying near-infrared (NIR) spectroscopy for the quantitative analysis of Δ^9-tetrahydrocannabinol (THC) in cannabis products using handheld devices. A preliminary study was conducted on different physical forms (entire, ground and sieved) of cannabis inflorescences in order to evaluate the impact of sample homogeneity on THC content predictions. Since entire cannabis inflorescences represent the most common types of samples found in both the pharmaceutical and illicit markets, they have been considered priority analytical targets. Two handheld NIR spectrophotometers (a low-cost device and a mid-cost device) were used to perform the analyses and their predictive performance was compared. Six partial least square (PLS) models based on reference data obtained by UHPLC-UV were built. The importance of the technical features of the spectrophotometer for quantitative applications was highlighted. The mid-cost system outperformed the low-cost system in terms of predictive performance, especially when analyzing entire cannabis inflorescences. In contrast, for the more homogeneous forms, the results were comparable.

The mid-cost system was selected as the best-suited spectrophotometer for this application. The number of cannabis inflorescence samples was augmented with new real samples, and a chemometric model based on machine learning ensemble algorithms was developed to predict the concentration of THC in those samples. Good predictive performance was obtained with a root mean squared error of prediction of 1.75 % (w/w). The Bland-Altman method was then used to compare the NIR predictions to the quantitative results obtained by UHPLC-UV and to evaluate the degree of accordance between the two analytical techniques. Each result fell within the established limits of agreement, demonstrating the feasibility of this chemometric model for analytical purposes.

Finally, resin samples were investigated by both NIR devices. Two PLS models were built by using a sample set of 45 samples. When the analytical performances were compared, the mid-cost spectrophotometer significantly outperformed the low-cost device for prediction accuracy and reproducibility.

1. Introduction

The analysis of cannabis samples mainly concerns two general areas: quality control laboratories (often for medicinal cannabis) and forensic laboratories (seized cannabis samples). The simplest medicinal cannabis samples available on the market consist of dried flower tips with the aim for use in various therapeutic indications (from multiple sclerosis to epilepsy) [1]. This product is

* Corresponding author.
E-mail address: riccardo.deidda@uliege.be (R. Deidda).
[1] These authors have equally contributed to this article.

https://doi.org/10.1016/j.jpba.2021.114150
0731-7085/

generally available in two main physical forms (entire and ground inflorescences) and with different concentrations of the two phytocannabinoids (CNBs) of interest for clinical practice: psychotropic Δ^9-tetrahydrocannabinol (THC) and nonpsychotropic cannabidiol (CBD). Their concentrations generally range from less than 1 to approximately 22 % (w/w) for THC and from less than 1 to approximately 9% (w/w) for CBD, depending on the cannabis strain [1–3]. Other pharmaceutical forms are available in Europe for medicinal purposes, is the case for products such as Sativex (sublingual spray), Marinol and Cesamet (capsule) [1].

Seized cannabis samples consist of potentially illicit cannabis products conceived for recreational use that need to be analyzed to determine whether they comply with the law of the reference country. In this context, there is no harmonization in Europe, which results in different tolerated levels that vary from country to country. In Switzerland, for instance, cannabis with a THC content exceeding 1% (w/w) is classified as an illicit drug [4]. Belgium, like many other European countries (such as Greece, France and Norway), sets this limit at 0.2 % (w/w) while Italy and Luxemburg at 0.6 and 0.3 % (w/w) respectively [1,5,6]. Several types of cannabis samples can be found in various and imaginative forms: herbal materials, resins, concentrates, edibles, vaped oils, and drinks are just some examples [7]. However, the most widespread types of illicit cannabis samples are herbal material (dried flowering tops) and resin [5]. The latter consists of a compressed solid fabricated from the resinous parts (glandular trichomes) of female flowers and is characterized by high THC contents, sometimes up to 36 % (w/w) [8]. It is worth noting that, according to the latest European drug report of the European Monitoring Centre for Drugs and Drug Addiction, cannabis is the most commonly used illicit drug in Europe [5].

As a consequence of the pharmacological and forensic interest in cannabis, there is currently wide interest in developing innovative analytical tools able to properly characterize cannabis samples and these tools face different challenges in terms of legislation compliance, sample complexity and technological limitations.

Among the most employed analytical techniques for cannabinoid analysis are chromatographic techniques, as recently reviewed by *Citti* et al. [9]. Gas chromatography (GC) has been widely applied to analyze cannabis samples, mostly for forensic purposes, and to assess the total THC content [10]. However, liquid chromatography (LC) has rapidly replaced GC when interest was not focused only on the THC content, and a more complete characterization of the cannabinoid profile was asked (in a pharmaceutical context, for instance) [9,11]. In addition to LC, supercritical fluid chromatography (SFC) is progressively gaining momentum in cannabis analysis due to its advantages in terms of lower analysis time and lesser ecological impact compared to LC [8,12,13]. However, as with chromatographic technique, when the original sample is in a solid form, a prior extraction procedure (generally solid-liquid extraction) is needed, which is time-consuming and achievable only with appropriate instrumentation in a laboratory [14]. Moreover, this represents a critical procedure that is at the basis of intra- and interlaboratory variabilities, as recently highlighted in the literature [15–17].

Regardless, the toolbox of an analytical chemist is rich in techniques, and the application of other techniques can be envisaged depending on the analytical purpose. In the context of applicable in-field techniques, near-infrared (NIR) spectroscopy is gaining increasing attention and relevance among the scientific community. The reason lies behind the many advantages offered by this spectroscopic technique. First, NIR analysis can avoid (or reduce) sample preparation and does not involve the use of large quantities of organic solvents, resulting in a very fast and green technique. Second, several portable/handheld NIR spectrophotometers are currently available on the market, allowing the movement of this analytical technique directly where analysis is needed. Further details about the application of handheld NIR spectrophotometers to pharmaceutical analysis can be found in a recent review [18].

Table 1
Some details about the two handheld NIR devices used for this study.

Parameters	*NIR-S-G1*	*MicroNIR*
Wavelength range (nm)	900 – 1700	950 – 1650
Light source	Two integrated vacuum tungsten halogen lamps	Two integrated vacuum tungsten lamps
Mode	Reflectance	Reflectance
Measurement time	8.32 s (32 scans)	0.8 s (100 scans)
Sample window	0.35 cm^2	7 cm^2
Dispersing element	Digital light processing (DLP)	Linear variable filter (LVF)
Detector	1 mm InGaAs uncooled	128 pixels InGaAs photodiode array
Signal to noise ratio	5000:1	25000:1
Integration time	1 s	10 ms
Weight	136 g	250 g
Number of data points	228	125
Cost	Low	Middle

Regarding the application of NIR spectroscopy to cannabis analysis, only a few works can be found in the literature, showing some interesting applications [19–22]. In a recent work, the possibility of differentiating between legal and illegal cannabis inflorescences by using a qualitative model implemented with a handheld system was explored. The chemometric model showed 100 % sensitivity (true positive) and selectivity (true negative) during validation studies [21].

On that basis, the aim of this study was to explore and showcase other possible applications of NIR spectroscopy in the context of quantitative cannabinoid analysis using handheld devices. Two handheld spectrophotometers (NIR-S-G1 and MicroNIR) were selected to simultaneously evaluate the performance of two devices differing in terms of technical features and price (low-cost vs. mid-cost). Concerning the analytical targets, focus was given to entire cannabis inflorescences and cannabis resins. Concerning inflorescences, the effect of sample homogeneity on THC predictions was studied on three different physical forms (entire, ground and sieved) by using two handheld NIR spectrophotometers. Afterwards, a chemometric model based on artificial intelligence was developed to predict the concentration of THC in entire cannabis inflorescences. The Bland-Altman method was applied to evaluate the degree of accordance between NIR predictions and UHPLC-UV results on THC content to demonstrate the interchangeability of both techniques. Finally, the application of NIR spectroscopy to cannabis resin analysis was investigated.

2. Materials and methods

2.1. Cannabis samples

The cannabis samples consisted of police seizures under the form of female plant inflorescences and resins that were provided by the School of Criminal Justice of the University of Lausanne, Switzerland.

Cannabis samples were first analyzed by handheld NIR spectroscopy and then underwent solid-liquid extraction to be analyzed by UHPLC-UV.

2.2. Near-infrared analysis

2.2.1. NIR spectrophotometers

Table 1 summarizes some technical aspects of the two handheld NIR spectrophotometers employed for this study and described hereafter.

2.2.1.1. MicroNIR onsite W 1700. This mid-cost handheld NIR spectrophotometer is commercialized by Viavi Solutions Inc. (Santa Rosa, California, USA) and operates in reflectance mode. It is equipped with a 128-pixel linear indium-gallium-arsenide (InGaAs) array detector with two tungsten light bulbs as the radiation source. The spectra were acquired in the range of 950–1650 nm (10,526 – 6060 cm^{-1}). The signal-to-noise ratio was 25000, and the integration time was 10 ms. The system is ultracompact, weighing only 250 g, and is powered by USB or built-in Li-ion batteries. It communicates with a mobile application using the Bluetooth protocol.

2.2.1.2. NIR-S-G1. This NIR system is a low-cost handheld NIR spectrophotometer distributed by InnoSpectra Corporation (Taiwan, China) operating in reflectance mode. It is equipped with a 1 mm indium-gallium-arsenide (InGaAs uncooled) detector with two integrated tungsten halogen lamps as the radiation source. The spectra were acquired in the range of 900–1700 nm (11111 – 5882 cm^{-1}). The signal-to-noise ratio was 5000:1 in a 1 s scan. The system is ultracompact, weighing only 136 g, and is powered by USB or built-in Li-ion batteries. The system offers the possibility to communicate with a mobile application using the Bluetooth protocol. In the present study, the device was piloted using the DLP NIRscan Nano GUI software version 2.1.0.

2.2.2. Sample presentation

Entire inflorescence and resin samples were analyzed by putting the sample in direct contact with the sampling window of the equipment. Ground and sieved inflorescence samples were analyzed through a PELD transparent plastic bag.

Each sample represented a part of a unique police seizure and no sub-sample was used for this study. As this type of sample is highly heterogeneous, four spectra were acquired during each NIR analysis and in different locations of the sample in order to increase the representativity of the spectra.

2.3. Reference technique: UHPLC-UV

UHPLC coupled to UV detection was chosen as the reference technique to generate the quantitative results used to calibrate and evaluate all the quantitative models. Such results were expressed as the total THC concentration, which was the sum of Δ^9-tetrahydrocannabinolic acid (THCA-A) and THC naturally present in the cannabis inflorescences. The analytical method and the protocol of analysis are accurately described in [8]. Hereafter, some details about this chromatographic method are reported.

2.3.1. Chemical and reagents

The phytocannabinoid standard solutions at 1 mg mL^{-1} in ethanol (for THC) and in 2-isopropanol (for THCA-A) used for this study were purchased from Lipomed AG (Arlesheim, Switzerland).

Ethanol (EtOH) and acetonitrile (ACN) of OPTIMA LC–MS grade, as well as water of UHPLC-MS grade, were purchased from Fischer Scientific (Loughborough, UK). Formic acid was obtained from Biosolve (Valkenswaald, Netherlands).

2.3.2. Standard solutions and cannabis samples

A unique standard stock solution was prepared at a concentration of 250 μg mL^{-1} for both THC and THCA-A. This standard stock solution was then diluted in water-ACN (3/7, v/v) to obtain final concentrations of 7.81, 15.62, 31.25, 62.50, and 125.0 μg mL^{-1} of both analytes.

Sample preparation for the cannabis samples consisted of solid liquid extraction using ethanol as the extraction solvent. To extract the CNBs from cannabis inflorescences and resins, 10 mL of ethanol was added to 500 mg of each cannabis sample in an Ika ultratube drive system for agitation and grinding for 4 min at 6000 rpm with two glass beads 6 mm diameter. The mixture was left at ambient temperature for 9 min, and centrifugation was performed for 3 min at 14000 rpm.

Cannabis extracts were diluted 50-fold before injection in the chromatographic system. A mixture of water and ACN (3:7, v/v) was used as the dilution solvent [8].

2.3.3. Apparatus and methodology

All experiments were performed on a Waters Acquity UPLC H-class system (Waters, Milford, MA, USA) equipped with a quaternary solvent manager, a sample manager with flow through needle (FTN) injector, and a column manager. A mix of ACN, EtOH and water (4:4:2 v/v) was used as the wash solvent, and a mix of ACN and water (7:3 v/v) was used as the purge solvent, with a postinjection wash of 6 s. The chromatographic system was connected to a Waters PDA detector for UV detection set at 214 nm.

Separation was performed at 30 °C on an InfinityLab Poroshell 120 EC-C18 column (150 × 2.1 mm, 2.7 μm) from Agilent (Santa Clara, USA) with an InfinityLab Poroshell 120 EC-C18 guard column (5 × 2.1 mm, 2.7 μm) from Agilent (Santa Clara, USA). Mobile phase A contained water with 0.1 % formic acid, and mobile phase B, ACN with 0.1 % formic acid. The gradient profile was as follows: 2.8 min in isocratic mode with 68 % B, followed by 68–73% B over 0.5 min, hold for 3.7 min, then to 95 % over 5.0 min and hold for 1.0 min. The percentage of B was finally brought to the initial conditions over 0.5 min and maintained for 4.5 min to reequilibrate the system. The flow rate was set at 0.5 mL min^{-1}, and the injection volume was 1 μL [8].

2.4. Software

The chromatographic system was equipped with Empower™ 3 software that was used to pilot the system and for data acquisition.

MATLAB R2018a software (The MathWorks, Massachusetts, USA) and PLS Toolbox® (version 8.6.2, Eigenvector Research, Washington, USA) were used for spectral data treatment and computation.

Python (version 3.7.7) was used with the scikit-learn library version 0.21.3 to train ensemble regression models.

2.5. Chemometrics

Various combinations of spectral preprocessing were evaluated in a previous study on similar substances [21]. The Savitzky-Golay 2nd derivative (polynomial order: 2, window size: 3) and standard normal variate (SNV) were the preprocessing with the best results to quantify substances with the MicroNIR. The same preprocessing methods were also used for the spectra obtained by NIR-S-G1, but setting this time a 7 window size for the Savitzky-Golay 2nd derivative. Replicate measures were grouped together during the cross-validation (leave-one-out approach) and not split between the calibration and validation sets.

Two types of regressors were trained using the preprocessed datasets:

- Genetic Algorithm – Partial least squares regressors were used to compare the predictions of the two devices (MicroNIR and NIR-S-G1) on the three forms of cannabis inflorescences and for

cannabis resins. Genetic algorithm was employed to select the variables correlated to the concentration of THC (population size, 64; window width, 30; % initial terms, 30; max generation 100; % at convergence, 50; mutation rate, 0.005; crossover, double). The complete list of variables selected for each GA-PLS model is reported in Figure S1 of the supplementary material.
- Ensemble regression models were trained with Python on the dataset of entire inflorescences and resins to optimize the predictions. Algorithms and parameters are reported in Figure S2 of the supplementary material.

3. Results and discussion

3.1. Preliminary study on the different physical forms of cannabis inflorescences

Although NIR spectroscopy has already been applied to the quantitative analysis of cannabinoids in ground cannabis inflorescences [20,23], the application of handheld systems to quantitatively analyze entire cannabis inflorescences has not yet been described in the literature. Entire cannabis inflorescences, due to their high degree of heterogeneity, represent a challenging analytical target for NIR analysis. In fact, NIR spectroscopy is sensitive not only to the chemical properties of matter but also to the physical properties [24]. As a consequence, parameters such as humidity, granulometry and appearance can significantly affect the NIR response. Therefore, it follows that a predictive model developed with NIR spectra obtained from entire inflorescences cannot be applied for predictions on ground inflorescences. Despite their complexity, entire inflorescences represent the most common type of sample found on both pharmaceutical and illicit markets and are a very coveted target for in-field NIR spectroscopy.

In order to correctly approach this sample, it was decided to conduct a preliminary study on different physical forms (entire, ground and sieved) of cannabis inflorescences. The aim was to evaluate the impact of sample homogeneity on the THC content predictions and investigate how the different spectrophotometers treated the obtained data. Two handheld NIR spectrophotometers were employed: a low-cost device, named NIR-S-G1, and a middle-cost device, named MicroNIR. Spectra were acquired on 26 samples that consisted of entire inflorescences confiscated by the police in the Canton of Vaud, Switzerland. Each sample was analyzed first by NIR on its original form (entire inflorescence), then after grinding (ground inflorescence) and finally after sieving (sieved inflorescence) with the aim of progressively increasing the grade of homogeneity. Figures S3a, S3b and S3c of the supplementary material show raw and preprocessed spectra for each physical form and NIR spectrophotometer. Marked spectral differences were observed between entire inflorescence spectra and those of ground and sieved inflorescences for both devices. In this case, NIR spectroscopy does not allow assigning a part of the spectrum to a unique expression of a chemical compound. The whole spectrum must be considered to find the combination of wavelength absorbances more correlated to the desired chemical information (THC content). In fact, the considered spectral region is associated to the 1st, 2nd and 3rd overtones of CH, CH_2, CH_3, 1st and 3rd overtones Ar−CH, 2nd and 3rd overtones of Ar−OH that are also present in many other chemical compounds that can be found in cannabis samples [20].

After NIR analysis, each sample was analyzed by UHPLC-UV in order to generate reference quantitative data about THC content [8]. The THC concentration range was relatively wide, ranging from 0.92 to 22.21 % (w/w). The acquired NIR spectra were preprocessed as follows. The Savitzy-Golay 2nd derivative was used to remove unimportant baseline signals and correct the so-called additive effects caused by noise and not captured photons. The standard normal variate (SNV) was used to correct multiplicative scaling effects generated by the lengthening of the optical path length. The Kennard-Stone algorithm (2/3:1/3) was used to split the database into calibration and validation sets [25]. Six different genetic algorithm-partial least square (GA-PLS) models (one for each physical form and NIR system) were generated. Fig. 1 shows the correlation between the NIR predictions and UHPLC data on THC concentration for each PLS model. Table 2 shows the results obtained for the various predictions, comprising the root mean square error of calibration (RMSEC), of cross-validation (RMSECV), and of prediction (RMSEP); cross-validation bias (CV bias); prediction bias; coefficient of determination in calibration (R^2_{cal}), in cross-validation (R^2_{CV}), and in prediction (R^2_{pred}) and ratio of prediction to deviation (RPD). As can be immediately noticed from these data, the performance of the two NIR spectrophotometers was quite different, particularly when analyzing inflorescences in their entire form. Looking at the RMSEP, this value was 2.07 % (w/w) for the chemometric model built with the NIR-S-G1 spectra and 1.04 % (w/w) for the MicroNIR spectra. The results become more comparable when the homogeneity of the sample increased, highlighting the critical impact of this parameter on the THC content predictions by NIR spectroscopy. This is particularly true for the NIR-S-G1 models but not for the MicroNIR models. In this case, model performance decreased for the ground and sieved inflorescences with respect to those obtained for the entire inflorescences.

In general, the MicroNIR spectrophotometer outperformed NIR-S-G1, particularly when approaching entire inflorescences. One explanation could be found in some different features of these spectrophotometers, which are detailed in 2.2.1. Moreover, the MicroNIR spectrophotometer presents a larger sample analysis window (7 cm^2 compared to 0.35 cm^2 for the NIR-S-G1), which allows a better representation of samples, resulting in the best THC concentration predictions. In addition, the theoretical signal-to-noise ratio is higher for the MicroNIR spectrophotometer compared to the NIR-S-G1 spectrophotometer.

On the basis of these results, it was stated that the use of a spectrophotometer with a larger sample analysis window is better fitted for highly heterogeneous samples. In contrast, when the degree of homogeneity increases (ground and sieved inflorescences), the use of smaller windows could be envisaged. In this case, a high number of detection spots from different parts of each sample is recommended in order to ensure a proper representation of the sample surface.

3.2. Feasibility study on entire cannabis inflorescence: application of ensemble regression models to THC content predictions

Based on the aforementioned preliminary results, the MicroNIR spectrophotometer was selected as the best suited for the analytical target. The aim here was to more deeply investigate the applicability of this spectrophotometer for THC quantitative analysis directly on an inflorescence that did not undergo a sample preparation process. This application could be very interesting for police road controls, for instance, because of the rapidity and ease of use of this technique. No laboratory material would be necessary, and the desired analytical response can be obtained in a few seconds. In fact, UHPLC results on THC content take several hours to become available.

For the creation of the chemometric model, the dataset was integrated with new real samples and another chemometric approach tested. Ensemble regression models were chosen as a promising chemometric solution to approach this arduous sample. Ensemble regression models evaluate a wide variety of machine learning algorithms and select the best combination of those associated with a better accuracy in cross-validation. Preprocessing algorithms are evaluated as well in order to choose the best fit. One such model is

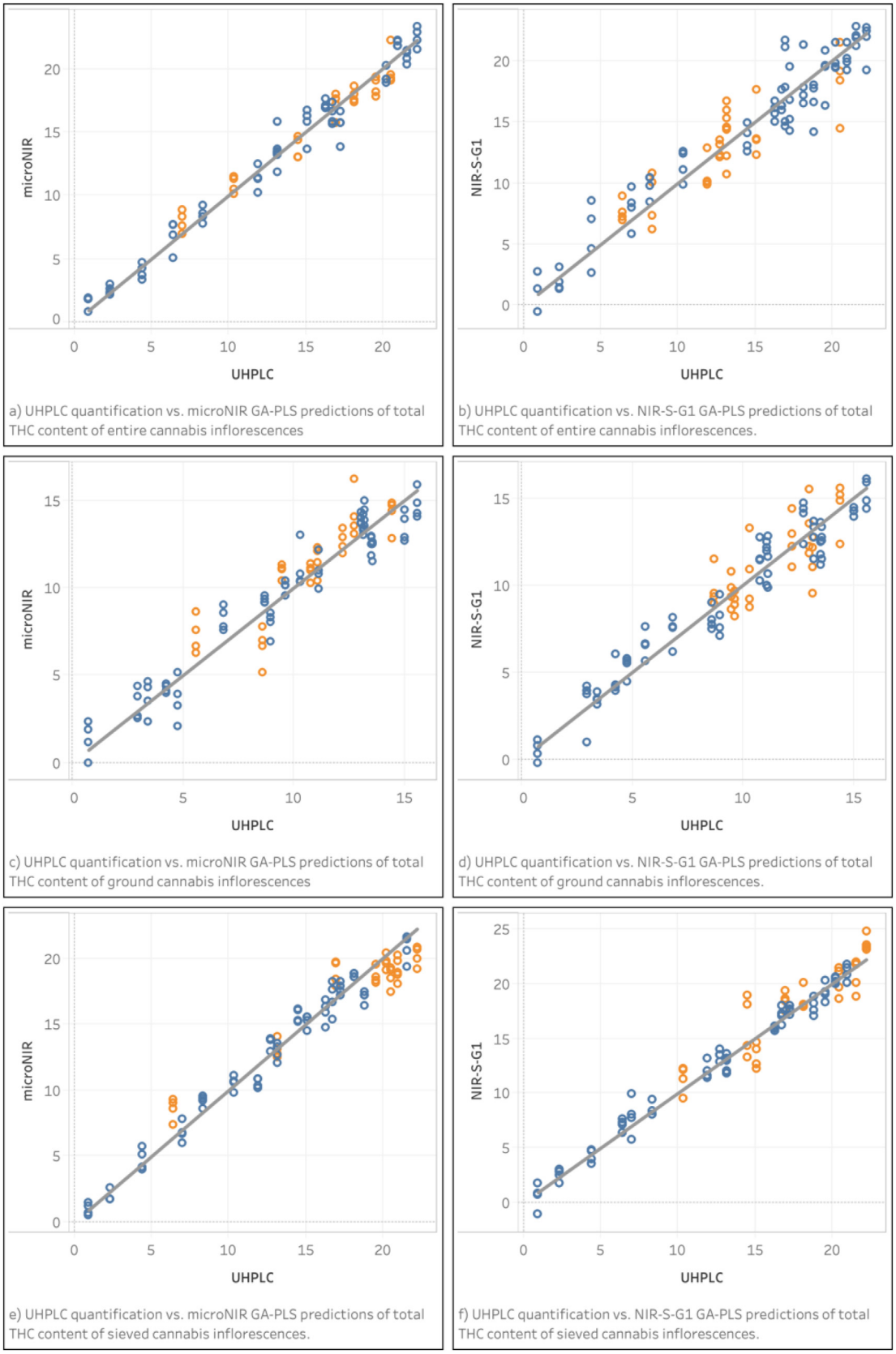

Fig. 1. GA-PLS models: UHPLC quantification vs. NIR predictions of total THC content on entire (a, b), ground (c, d) and sieved (e, f) cannabis inflorescences by MicroNIR (a, c, e) and NIR-S-G1 (b, d, f).

Table 2
Details about the GA-PLS models built to predict the THC content in three different physical forms of cannabis inflorescences by using a NIR-S-G1 and MicroNIR system.

	NIR-S-G1			MicroNIR		
PLS details	Entire inflorescences	Ground inflorescences	Sieved inflorescences	Entire inflorescences	Ground inflorescences	Sieved inflorescences
No LVs	6	7	7	6	4	4
RMSEC	1.74	0.88	0.86	0.74	1.02	0.96
RMSECV	2.79	1.62	1.99	0.98	1.28	1.08
RMSEP	2.07	1.99	1.55	1.04	1.78	1.75
Bias	0.01	−0.02	0.01	0.00	0.00	0.00
CV bias	−0.02	0.04	0.02	0.01	0.06	−0.00
Prediction bias	−0.03	0.43	0.22	0.23	0.92	1.34
R^2 calibration	0.93	0.96	0.98	0.98	0.94	0.98
R^2 CV	0.82	0.85	0.90	0.97	0.91	0.97
R^2 prediction	0.73	0.74	0.93	0.93	0.76	0.77
RPD	1.95	1.00	2.48	4.54	1.46	2.86

No LVs, number of latent variables; RMSEC, root mean squares error of calibration; RMSECV, root mean squares error of cross-validation; RMSEP, root mean squares error of prediction; CV bias, cross-validation bias; R^2 CV, determination coefficient in cross-validation; RPD, ratio of prediction to deviation.

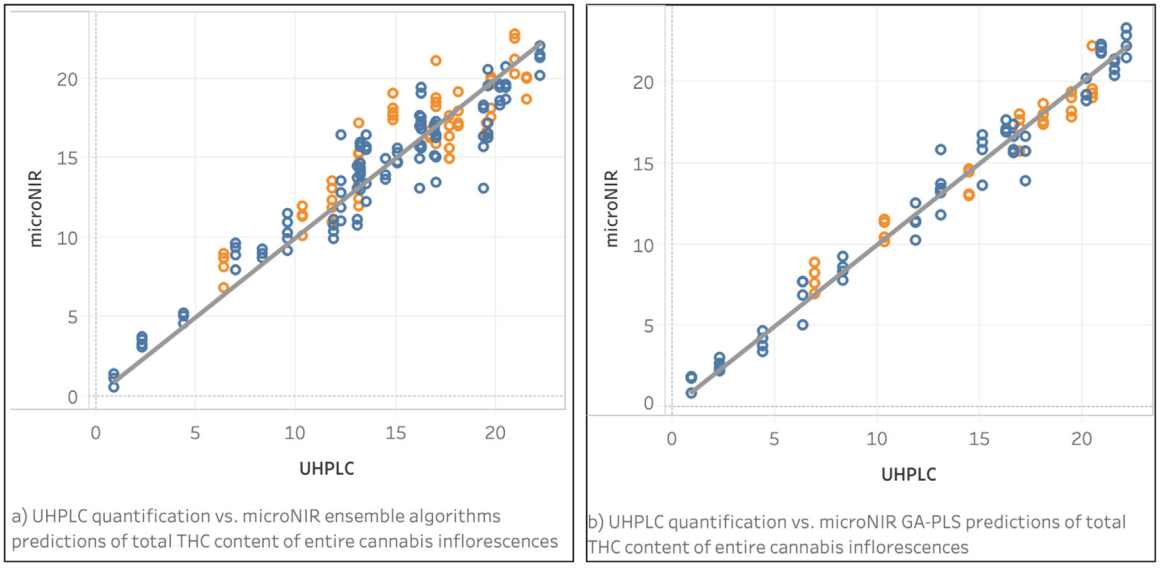

Fig. 2. UHPLC quantification vs. NIR predictions of total THC content on entire cannabis inflorescences by a) ensemble regression models and b) GA-PLS model built on MicroNIR.

referred to in [26] for more details about the implementation of this chemometric approach. For the development of this model, a total of 36 cannabis inflorescence samples were analyzed by means of the MicroNIR spectrophotometer. Samples were then analyzed by UHPLC-UV [8] in order to obtain reference data to calibrate and validate the chemometric model. The THC concentration range went from 0.92 to 22.21 % (w/w). The Kennard-Stone algorithm (2/3:1/3) was used to split spectra into calibration and validation sets.

Fig. 2 shows the correlation between the THC concentrations predicted by the NIR model and the reference values obtained via UHPLC-UV compared to those of GA-PLS on the entire inflorescences with the same device. Figure S2 of the supplementary material shows the parameters selected for this model. Table 3 summarizes the model performance in terms of RMSEC, RMSEP, R^2_{cal}, R^2_{val}, RPD and prediction bias. Good accuracy was obtained with an RMSEP of 1.75 % (w/w). However, it must be said that the advantages in terms of performance of this type of chemometric model are better exploited when large-scale datasets are involved, as demonstrated in a previous study [21]. In contrast, GA-PLS models could become less accurate when heterogeneous samples are integrated into the dataset due to the effect of other variable parameters, such as humidity, color and size.

In order to deeper evaluate interchangeability, the Bland-Altman statistical method was applied to evaluate the degree of agreement between NIR predictions and UHPLC reference values. This statistical method evaluates the model with a "difference plot", which consists of plotting the differences between quantitative measures on the vertical axis against the best estimate of the true value (the mean of the pair of measurements, for instance) on the horizontal axis [27]. Since the UHPLC-UV method was used as a reference method to calibrate the NIR model, its measures were used as the true values and plotted on the horizontal axis [28]. The relative difference percentages were plotted on the vertical axis, as this is convenient when methods show variability linked to increased magnitude. The limits of agreement (LoA) were built by calculating the mean and the standard deviation of the differences

Table 3
Details about the model built to predict the THC content in entire cannabis inflorescences by using a MicroNIR system.

Model details	
RMSEC	2.04
RMSEP	1.75
R^2 calibration	0.91
R^2 prediction	0.85
Prediction bias	0.21
RPD	2.52

RMSEC, root mean squares error of calibration; RMSEP, root mean squares error of prediction; R^2, determination coefficient; RPD, ratio of prediction to deviation.

Table 4
Quantitative data on THC content used to evaluate the agreement between the two analytical methods, UHPLC and NIR. The table includes also the true values, the relative difference percentages, the bias and the values of the limits of agreement used to obtain the Bland-Altman plot.

Sample No	Conc UHPLC (% w/w)	Conc NIR(% w/w)	True value (% w/w)	Relative difference percentage (%)
1	19.8	19.2	19.8	3.03
2	16.7	16.4	16.7	1.80
3	17.0	16.5	17.0	2.94
4	13.1	15.6	13.1	−19.08
5	14.5	14.0	14.5	3.45
6	10.4	11.6	10.4	−11.54
7	7.0	9.0	7.0	−28.57
8	16.2	15.7	16.2	3.09
9	20.2	19.5	20.2	3.47
10	16.3	17.0	16.3	−4.29
11	13.2	14.4	13.2	−9.09
12	19.5	15.1	19.5	22.56
13	17.7	16.1	17.7	9.04
14	6.4	7.8	6.4	−21.88
Bias (%)	−3.22			
Std. Dev. (%)	13.57			
uLoA (%)	23.37			
lLoA (%)	29.81			

Sample No, sample number; Conc UHPLC, THC concentration obtained by UHPLC; Conc NIR, THC concentration obtained by NIR; True value, THC concentration obtained by UHPLC; Relative difference, relative difference between Conc UHPLC and Conc NIR; Bias, average of relative difference percentages; Std. Dev., standard deviation of relative difference percentages; uLoA, upper limit of agreement; lLoA, lower limit of agreement.

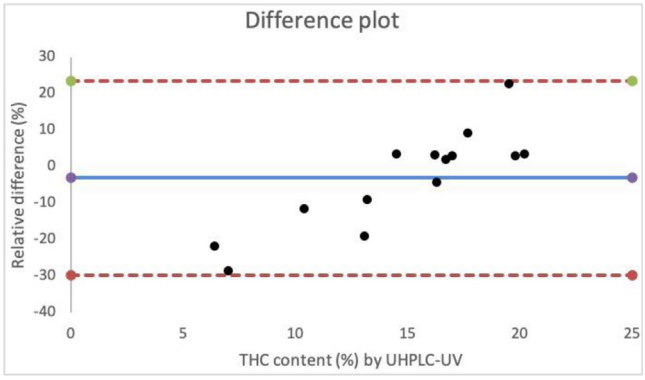

Fig. 3. Bland-Altman plot built on the basis of the THC quantitative results obtained by UHPLC and NIR.

between measurement pairs. These limits include both systematic (bias) and random error (precision) and allow comparison of the likely differences between individual results measured by two analytical methods [8,27,28]. Table 4 summarizes these data as well as the statistical data treatment, while Fig. 3 displays the Bland-Altman plot. In this plot, the dotted red lines represent the upper and lower limits of agreement (LoA = ± bias - 1.96 standard deviation), while the solid line represents the bias. The bias gives us important information about the agreement between the NIR predictions and the reference values. It is calculated by the mean of the differences, which should ideally be zero. When this value is not equal to zero, it means that on average, the method tested overestimates or underestimates the analyte concentration. In this specific case, the bias was equal to -3.22, meaning that the NIR model overestimates the THC content. A possible explanation of this trend could be found looking deeper at the nature of the sample under investigation. In fact, cannabinoids are synthesized by glandular trichomes, which are particularly expressed in the external part of cannabis female inflorescences [29]. Thus, higher concentrations of cannabinoids are expected on the superficial parts of these inflorescences than in flower support structures, such as stalks or leaves. In the NIR analysis, spectra were acquired on these superficial parts of the sample, while for UHPLC-UV analysis, the cannabinoids were extracted from the whole inflorescence before analysis.

However, as one can notice in the difference plot that all predictions presented relative difference percentage values within the established LoA. Only two analyses presented values that were close to these LoAs.

In conclusion, it can be stated that the quantitative data generated by applying the NIR predictive model strongly agree with those generated by UHPLC-UV. This aspect confirms the applicability of NIR spectroscopy to THC quantification in entire cannabis inflorescences.

However, it is worth noting that the creation of a definitive model to be applied to routine analysis would need a large number of samples covering the whole concentration range and able to represent all the intrinsic variability that characterizes these samples. This was not the aim of this project, which, in contrast, aimed to demonstrate only the feasibility of employing this innovative technique by using a relatively limited number of samples, which were available for this study. This is just an estimate of what NIR spectroscopy is able to do in terms of analytical performance.

3.3. *Preliminary study on cannabis resins*

When discussing cannabis samples, the first image that comes to mind is that of dry inflorescence buds. Effectively, this is the most common appearance of cannabis samples (and the reason for which it was decided to prioritize this type of sample in the first part of this work), but not the only one. This next part of the present work was dedicated to the exploration of the applicability of NIR spectroscopy to another type of cannabis sample: resins. Resins consist of brown/black bricks of a compressed solid fabricated from the resinous parts of the female inflorescences and are characterized by high THC contents. This is a very common sample in a forensic context, since it is largely used for recreational purposes. As cannabis inflorescences, this is a challenging sample to analyze by NIR spectroscopy. In fact, in addition to its heterogeneity, its dark brown/black color can interfere with the reflectance measured during NIR analysis. Moreover, its physicochemical properties are unpredictable and variable, which are aspects that should be taken into account during model calibration.

For this part of the study, 45 resin samples, consisting of those from police seizures, were analyzed to evaluate the applicability of NIR spectroscopy to predict their THC concentration. UHPLC-UV analysis was performed to obtain the reference values, which

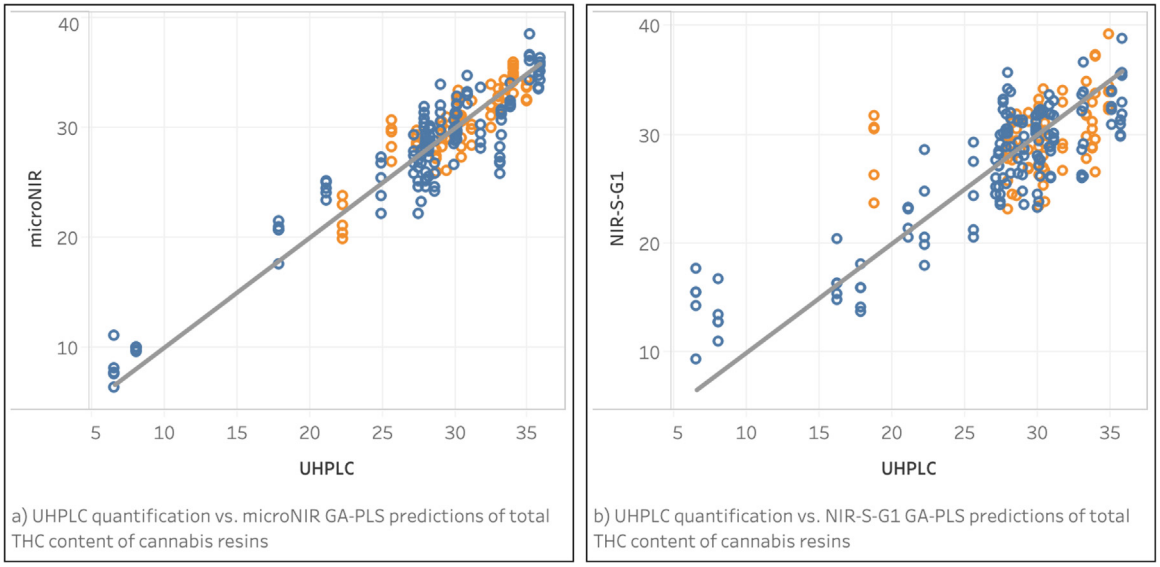

Fig. 4. GA-PLS models: UHPLC quantification vs. NIR predictions of total THC content on cannabis resins by MicroNIR and NIR-S-G1.

Table 5
Details about the GA-PLS models built to predict THC content on resin samples by using NIR-S-G1 and MicroNIR systems.

PLS details	*MicroNIR*	*NIR-S-G1*
No LVs	5	6
RMSEC	2.45	4.25
RMSECV	2.50	5.19
RMSEP	1.46	3.87
bias	0.02	0.03
CV bias	0.04	0.01
Prediction bias	−0.14	−1.06
R^2 calibration	0.87	0.72
R^2 CV	0.87	0.59
R^2 prediction	0.67	0.02
RPD	2.26	1.51

No LVs, number of latent variables; RMSEC, root mean squares error of calibration; RMSECV, root mean squares error of cross-validation; RMSEP, root mean squares error of prediction; CV bias, cross-validation bias; R^2 CV, determination coefficient in cross-validation; RPD, ratio of prediction to deviation.

confirmed the expected high THC concentration values of most samples. In fact, the THC concentration range was from 6.56 to 35.84 % (w/w). Figure S4 of the supplementary material shows an UHPLC-UV chromatogram obtained from the analysis of a resin sample, as an example. Unfortunately, the samples available for the study did not cover a large range of concentrations, and this sample shifted and was dense at high THC concentrations (only 11 % of samples presented THC content < 20 % (w/w)). This does not represent the best conditions for the calibration of a quantitative model since it does not allow a proper representation of low THC concentration samples.

Nevertheless, NIR analyses were performed with both spectrophotometers (MicroNIR and NIR-S-G1) to evaluate the performance of the spectrophotometers. NIR spectra samples were preprocessed by applying a Savitzky-Golay 2nd derivative and SNV normalization to reduce irrelevant information from spectra. Figure S5 shows the NIR spectra before and after preprocessing. The Kennard-Stone algorithm was used to split spectra into calibration and validation sets (2/3:1/3). A PLS model was then built for each spectrophotometer by using the spectra acquired during the analysis. Table 5 reports the analytical performance of these models. Analyzing the results, one can state that the MicroNIR spectrophotometer also outperformed the other equipment for this application. In fact, the MicroNIR device allowed obtaining a lower RMSEP (1.46 % versus 3.87 %) and higher coefficients of determination in calibration, cross-validation and prediction. As mentioned for cannabis inflorescences, the technical features of the spectrophotometer are fundamental to assure the success of its application. Fig. 4 shows the correlation between the NIR predictions and UHPLC results that were obtained for both models. The integration of this dataset with the spectra collected from low THC samples would help refine model accuracy.

4. Conclusions

The aim of this work was to demonstrate the feasibility of the quantitative applications of NIR spectroscopy to the in-field-based analysis of cannabis samples. The applicability of this technique for the quantitative analysis of THC of the two main types of cannabis samples has been demonstrated. To the best of our knowledge, this is the first work describing a quantitative application of NIR spectroscopy to entire cannabis inflorescences and resins and more specifically employing handheld devices allowing for in-field analyses.

However, when approaching cannabis sample analysis by NIR spectroscopy, it is essential to consider certain crucial aspects. First, the chemometric model should be specific for a defined physical form of sample. Second, instrumental features have to be considered as well, since several handheld devices are now available on the market, but not all necessarily fit with the complexity of the analytical target or with the intended purpose. Cannabis samples, due to their heterogeneity and continuous changes in their chemical profile, can be challenging to properly analyze. A large sample size is needed in order to correctly represent all the naturally occurring variabilities. Predictions should be done within the range used to calibrate the chemometric model. Moreover, before routine applications, the quantification range should be determined through proper validation studies. To avoid false predictive results in routine application, one shall first apply a one class classification (of the type of nearest neighbours for instance) so as to exclude predictive regression on spectra that are too different from the calibration set.

NIR models need to be implemented with new samples when the analytical target may present physicochemical variation. Concerning the pharmaceutical context, when relevant changes in the cultivation process occur, such as a different strain of cannabis or

changes in the cultivation parameters that can have an impact on the chemical profile of the product, a new calibration process would be necessary. Similarly, a new calibration process would be needed when the physical properties change (passing from entire to ground samples, etc.).

In a forensic context, implementation with new samples should be a continuous process since cannabis samples vary greatly over time and regions where they are illegally sold. For this reason, it is essential to continuously implement a library of NIR spectrophotometers with representative samples.

CRediT authorship contribution statement

Riccardo Deidda: Writing - original draft, Writing - review & editing, Conceptualization, Formal analysis, Visualization. **Florentin Coppey:** Writing - review & editing, Conceptualization, Formal analysis, Visualization, Software. **Dhouha Damergi:** Formal analysis. **Cédric Schelling:** Conceptualization, Formal analysis, Writing - review & editing. **Laureen Coïc:** Software. **Jean-Luc Veuthey:** Writing - review & editing, Conceptualization, Supervision, Funding acquisition. **Pierre-Yves Sacré:** Writing - review & editing. **Charlotte De Bleye:** Writing - review & editing, Conceptualization, Supervision. **Philippe Hubert:** Supervision, Funding acquisition, Conceptualization, Project administration. **Pierre Esseiva:** Supervision, Funding acquisition, Conceptualization, Project administration. **Éric Ziemons:** Writing - review & editing, Conceptualization, Supervision, Funding acquisition, Project administration.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

The Fondation Léon Fredericq is acknowledged for the “bourse de voyage” grant given to R. Deidda.

This project has been supported by the European funds of regional development (FEDER) and by Walloon Region of Belgium as part of the operational program “Walloon-2020.EU” (L. Coïc).

The financial support of this research by the Walloon Region of Belgium in the framework of the Vibra4Fake project (convention n°:7517) is gratefully acknowledged.

The canton police of Vaud department and the police of the city of Lausanne is acknowledged for the granting of the samples used for this study.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jpba.2021.114150.

References

[1] European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), Cannabis Legislation in Europe, 2020 https://www.emcdda.europa.eu/system/files/publications/4135/TD0217210ENN.pdf.
[2] R. Deidda, H.T. Avohou, R. Baronti, P.L. Davolio, B. Pasquini, M. Del Budda, C. Hubert, Ph. Hubert, S. Orlandini, S. Furlanetto, Analytical quality by design: development and control strategy for a LC method to evaluate the cannabinoids content in cannabis olive oil extracts, J. Pharm. Biomed. Anal. 166 (2019) 326–335.
[3] Bedrocan website, Products, 2021, Accessed on 4th March https://bedrocan.com/products-services/.
[4] Sensi Seeds website, Blog, Legal & Politics, 2021, Accessed on 13rd March https://sensiseeds.com/en/blog/category/legal-politics/page/9/.
[5] European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), European Drug Report, Trends and Developments, 2020, Accessed on 1st October 2020 https://www.emcdda.europa.eu/system/files/publications/13236/TDAT20001ENN_web.pdf.
[6] M. Deville, N. Dubois, R. Denooz, C. Charlier, Validation of an UHPLC/DAD method for the determination of cannabinoids in seized materials: analysis of 213 samples sold in Belgian CBD shops, For. Sci. Int. 310 (2020), 110234.
[7] S. Goodman, E. Wadsworth, C. Leos-Toro, D. Hammond, Prevalence and forms of cannabis use in legal vs. Illegal recreational cannabis markets, Int. J. Drug Policy 76 (2020), 102658.
[8] R. Deidda, C. Schelling, J.-M. Roussel, A. Dispas, C. De Bleye, E. Ziemons, P. Hubert, J.-L. Veuthey, The analysis of cannabinoids in cannabis samples by supercritical fluid chromatography and ultra-high-performance liquid chromatography: a comparison study, Anal. Sci. Adv. 2 (2021) 2–14.
[9] C. Citti, D. Braghiroli, M.A. Vandelli, G. Cannazza, Pharmaceutical and biomedical analysis of cannabinoids: a critical review, J. Pharm. Biomed. Anal. 147 (2018) 565–579.
[10] L. Nahar, M. Guo, S.D. Sarker, Gas chromatographic analysis of naturally occurring cannabinoids: a review of literature published during the past decade, Phytochem. Anal. 31 (2020) 135–146.
[11] L. Nahar, A. Onder, S.D. Sarker, A review on the recent advances in HPLC, UHPLC and UPLC analysis of naturally occurring cannabinoids (2010-2019), Phytochem. Anal. 31 (2020) 413–457.
[12] M.L. Miller, J.D. Pinkston, L.T. Taylor, Analysis of cannabis products by supercritical fluid chromatography, in: M.F. Vitha (Ed.), Modern Supercritical Fluid Chromatography Carbon Dioxide Containing Mobile Phases, John Wiley & Sons, Inc;, N.J. Hoboken, 2020, pp. 77–78.
[13] M. Wang, Y.-H. Wang, B. Avula, M.M. Radwan, A.S. Wanas, Z. Mehmedic, J. van Antwerp, M.A. ElSohly, I.A. Khan, Quantitative determination of cannabinoids in cannabis and cannabis products using ultra-high-performance supercritical fluid chromatography and diode array/mass spectrometric detection, J. Forensic Sci. 62 (2017) 602–611.
[14] L. Nahar, S.J. Uddin, A. Alam, S.D. Sarker, Extraction of naturally occurring cannabinoids: an update, Phytochem. Anal. (2020) 1–14.
[15] C. Citti, F. Russo, S. Sgrò, A. Gallo, A. Zanotto, F. Forni, M.A. Vandelli, A. Laganà, C.M. Montone, G. Gigli, G. Cannazza, Pitfalls in the analysis of phytocannabinoids in cannabis inflorescences, Anal. Bioanal. Chem. 412 (17) (2020) 4009–4022.
[16] N. Jikomes, M. Zoorob, The cannabinoid content of legal Cannabis in Washington State varies systematically across testing facilities and popular consumer products, Sci. Rep. 8 (1) (2018) 4519.
[17] B.C. Smith, How standardized sample preparation can solve the inter-laboratory variation problem, Cannabis science and technology 3 (2) (2020) 10–15.
[18] R. Deidda, P.-Y. Sacré, M. Clavaud, L. Coïc, H. Avohou, Ph. Hubert, E. Ziemons, Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers, Trends Analyt. Chem. 114 (2019) 251–259.
[19] B.T. Borille, M.C.A. Marcelo, R. Scorsatto Ortiz, K. de Cássia Mariotti, M.F. Ferrão, R. Pereira Limberger, Near infrared spectroscopy combined with chemometrics for growth stage classification of cannabis cultivated in a greenhouse from seized seeds, Spectrochim. Acta A. 173 (2017) 318–323.
[20] C. Sánchez-Carnerero Callado, N. Núñez- Sánchez, S. Casano, C. Ferreiro-Vera, The potential of near infrared spectroscopy to estimate the content of cannabinoids in Cannabis sativa L.: a comparative study, Talanta 190 (2018) 147–157.
[21] F. Coppey, A. Bécue, P.-Y. Sacré, E.M. Ziemons, P. Hubert, P. Esseiva, Providing illicit drug results in five seconds using ultra-portable NIR technology: an opportunity for forensic laboratories to cope with the trend toward the decentralization of forensic capabilities, Forensic Sci. Intern. 317 (2020), 110498.
[22] C. Duchateau, J.-M. Kauffmann, M. Canfyn, C. Stévigny, K. De Braekeleer, E. Deconinck, Discrimination of legal and illegal *Cannabis spp.* According to European legislation using near infrared spectroscopy and chemometrics, Drug Test. Anal. 12 (9) (2020) 1309–1319.
[23] D. Townsend, I. Eustis, M. Lewis, S. Rodgers, K. Smith, A. Bohman, The Determination of Total THC and CBD Content in Cannabis Flower by Fourier Transform Near Infrared Spectroscopy (application Note), 2021 https://www.perkinelmer.com/lab-solutions/resources/docs/app_determination_of_thc_and_cbd_cannabisflower.pdf.
[24] J.S. Shenk, J.J. Workman, M.O. Westerhaus, Application of NIR spectroscopy to agricultural products, in: D.A. Burns, E.W. Ciurczak (Eds.), Handbook of Near-Infrared Analysis, Taylor & Francis Group, LCC, Boca Raton, FL, 2008, pp. 347–382.
[25] R.W. Kennard, L.A. Stone, Computer aided design of experiments, Technometrics 11 (1969) 137–148.
[26] F. Pedregosa, G. Varoquaux, A. Gramfort, M. Vincent, B. Thirion, Scikit-learn: Machine Learning in Phython, J. Mach. Learn. Res. 12 (2011), 2925–2830.
[27] J.M. Bland, D.G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 1 (8476) (1986) 307–310.
[28] D. Giavarina, Understanding Bland altman analysis, Biochem. Med. (Zagreb) 25 (2) (2015) 141–151.
[29] O. Aizpurua-Olaizola, U. Soydaner, E. Öztürk, D. Schibano, Y. Simsir, P. Navarro, N. Etxebarria, A. Usobiaga, Evolution of the cannabinoid and terpene content growth of *Cannabis sativa* plants from different chemotypes, J. Nat. Prod. 79 (2016) 324–331.