Talanta Open 8 (2023) 100270

Contents lists available at ScienceDirect

Talanta Open

journal homepage: www.sciencedirect.com/journal/talanta-open

Verification of the active pharmaceutical ingredient in tablets using a low-cost near-infrared spectrometer

Julia Gabel [a], Gesa Gnegel [a], Waltraud Kessler [b], Pierre-Yves Sacré [c], Lutz Heide [a,*]

[a] Pharmaceutical Institute, Eberhard Karls University Tübingen, 72070 Tübingen, Germany
[b] Faculty of Life Sciences, Reutlingen University, 72762 Reutlingen, Germany
[c] Department of Pharmacy, University of Liège (ULiège), CIRM, Research Support Unit in Chemometrics, Liège, B-4000 Liège, Belgium

ARTICLE INFO

Keywords:
Near-infrared spectroscopy
Pharmaceuticals
Medicine quality
Substandard and falsified medicines
DD-SIMCA
Authentication

ABSTRACT

The present study investigated the possibilities and limitations of using a low-cost NIR spectrometer for the verification of the presence of the declared active pharmaceutical ingredients (APIs) in tablet formulations, especially for medicine screening studies in low-resource settings. Spectra from 950 to 1650 nm were recorded for 170 pharmaceutical products representing 41 different APIs, API combinations or placebos. Most of the products, including 20 falsified medicines, had been collected in medicine quality studies in African countries. After exploratory principal component analysis, models were built using data-driven soft independent modelling of class analogy (DD-SIMCA), a one-class classifier algorithm, for tablet products of penicillin V, sulfamethoxazole/trimethoprim, ciprofloxacin, furosemide, metronidazole, metformin, hydrochlorothiazide, and doxycycline. Spectra of amoxicillin and amoxicillin/clavulanic acid tablets were combined into a single model. Models were tested using Procrustes cross-validation and by projection of spectra of tablets containing the same or different APIs. Tablets containing no or different APIs could be identified with 100 % specificity in all models. A separation of the spectra of amoxicillin and amoxicillin/clavulanic acid tablets was achieved by partial least squares discriminant analysis. 15 out of 19 external validation products (79 %) representing different brands of the same APIs were correctly identified as members of the target class; three of the four rejected samples showed an API mass percentage of the total tablet weight that was out of the range covered in the respective calibration set. Therefore, in future investigations larger and more representative spectral libraries are required for model building. Falsified medicines containing no API, incorrect APIs, or grossly incorrect amounts of the declared APIs could be readily identified.

Variation between different NIR-S-G1 spectroscopic devices led to a loss of accuracy if spectra recorded with different devices were pooled. Therefore, piecewise direct standardization was applied for calibration transfer.

The investigated method is a promising tool for medicine screening studies in low-resource settings.

1. Introduction

Substandard and falsified (SF) medicines expose patients worldwide to the risk of prolonged illness, economic loss or even death, due to ineffective therapy or toxic effects. With an estimated prevalence of 10.5 %, low- and middle-income countries (LMICs) are especially affected by SF medicines [1,2].

In LMICs, the national medicine regulatory authorities and other stakeholders in medicine quality assurance are often underfunded, and they lack both personnel and access to fully equipped medicine quality control laboratories. This allows medical products to enter the market without sufficient quality control, or even through illicit channels [3,4]. Rapid, low-cost medicine quality screening technologies have therefore been recommended in the fight against SF medicines in LMICs [5], and several countries have started implementing such programmes [6–8]. Easy-to-use screening technologies may also be used by personnel specialized in other disciplines such as supply chain staff and health care workers. We recently reported on the use of the Global Pharma Health Fund (GPHF)-Minilab® by faith-based medicine supply organizations in Africa [9]. The GPHF-Minilab® employs thin-layer chromatography for the identification of active pharmaceutical ingredients (APIs), and is currently the most widely used medicine quality screening tool

* Corresponding author.
E-mail address: heide@uni-tuebingen.de (L. Heide).

https://doi.org/10.1016/j.talo.2023.100270
Received 29 August 2023; Received in revised form 30 October 2023; Accepted 6 November 2023
Available online 7 November 2023

worldwide [10,11].

In recent years, portable near-infrared (NIR) spectrometers have received attention as medicine quality screening technologies [10, 12-16]. They require neither consumables nor sample preparation/destruction, the time required for analysis is very short, and only minimal training is necessary for their operation [10,17,18]. However, interpretation of the complex NIR spectra usually requires the use of multivariate data analysis [19–24].

NIR spectra of pharmaceutical products depend both on the chemical composition, i.e. on the APIs as well as on the excipients, and on physicochemical characteristics such as crystallinity, particle size, moisture etc. [19]. Extensive research has been carried out on the use of NIR spectroscopy for the discrimination of authentic, usually branded products of a given manufacturer from falsified products, even in cases when both have a similar composition [25–32]. Obviously, this approach requires a complete and up-to-date library of the NIR spectra of all authentic products of interest, and the creation of such libraries requires considerable effort [19,33]. For many stakeholders involved in medicine procurement in LMICs, such as the faith-based medicine supply organizations mentioned above [9,34], complete and up-to-date libraries of the NIR spectra of all relevant authentic products are impossible to obtain, since these organizations procure generic medicines at low prices from a large and ever-changing number of manufacturers located in India, China and many other countries [9,35]. Unfortunately, the occurrence of falsified medicines which contain no APIs, grossly insufficient amounts of APIs or even incorrect APIs remains a constant problem in such settings [1,36].

In the present study, we investigated whether a low-cost NIR spectrometer combined with the chemometric tools of principal component analysis (PCA) [21] and DD-SIMCA [37,38] may represent a useful technology for the verification of the APIs in pharmaceutical tablets, irrespective of specific brands, in low-resource settings. While some previous studies have addressed API verification by NIR spectroscopy, they used more expensive equipment or included only small numbers of products [39–41]. We now investigated 170 pharmaceutical products. 110 of these had been collected in African countries, and all of these had been investigated in our laboratory for the identity, quantity and dissolution of the APIs using compendial high-performance liquid chromatography (HPLC) methods [9,35,42]. The present study also included 20 products which had been identified as falsified. We employed the low-cost handheld NIR-S-G1 spectrometer (InnoSpectra, Taiwan; 1200–1600 USD per unit [13]) which has been evaluated in previous studies for its use in low-resource settings [30-32,43-46].

2. Methods

2.1 Investigated pharmaceutical products

Table 1 shows the types of medicines investigated in the present study. All of these are listed in the 2021 WHO Model List of Essential Medicines [47], and all were formulated as tablets. Of these, 105 products had been collected in Cameroon, the Democratic Republic (DR) of Congo and Chad during recently published medicine quality studies of our group; the procedures for their collection are described in the respective publications [9,35,42]. Additionally, five falsified medicines had been purchased in an ongoing study in Nigeria, from different commercial suppliers in Anambra and Enugu states. Most of the samples collected in African countries had been sold without a package leaflet, and no information was available on the excipients they contained.

Furthermore, 60 products were purchased from the pharmacy of Tübingen University Hospital and from a licensed retail pharmacy in Tübingen, and represented medicines licensed and marketed in Germany. If more than one brand of a certain type of medicine was purchased in Germany (Table 1), brands were selected which differed from each other in their excipient composition.

As shown in Table S1 (Supplementary Material), 143 of the products were white tablets, while the remaining 27 products were green (1), yellow (12), blue (2), brown (2), beige (3), grey (1), pink (4) or orange

Table 1
Overview of the pharmaceutical products investigated in this study.

Stated API(s)	Stated strength (mg)	Number of products:					Collected in:		Total number	Mass% API(s) [$]
		Calibration set	External validation sets				Africa*	Germany		
			V1 different brand	V2 different batch	V3 different strength	V5-V8 falsified medicines				
Medicine types included in calibration and validation sets:										
Amoxicillin	500	6	2	3	2	0	11	2	13	52–82
Amoxicillin/ clavulanic acid	500/125	5	2	0	1	2	8	2	10	45–55/ 8–13
Ciprofloxacin	500	6	2	1	2	0	7	4	11	59–73
Doxycycline	100	6	2	3	0	0	10	1	11	32–68
Furosemide	40	6	2	2	0	4	12	2	14	21–38
Hydrochlorothiazide	50	3	1	1	5	1	7	4	11	12–50
Metformin	500	6	2	2	4	0	12	2	14	77–94
Metronidazole	250	6	2	3	3	2	13	3	16	34–85
Penicillin V	250	6	2	2	3	1	12	2	14	71–89
Sulfamethoxazole/ trimethoprim	400/80	6	2	1	1	8	16	2	18	58–77/ 12–15
Placebo	–	3	1	0	0	0	0	4	4	
Total		**59**	**20**	**18**	**21**	**18**	**108**	**28**	**136**	
Further medicine types:										
Falsified "chloroquine" [§]						2	2	0	2	
30 additional APIs [#]						0	0	32	32	32–92
Grand total						**20**	**110**	**60**	**170**	

All products were formulated as tablets. * Cameroon, Democratic Republic of Congo, Chad, and Nigeria [9,35,42].

[$] Proportion of active pharmaceutical ingredient(s) in total tablet weight, expressed as percent. The stated mass% range was calculated from tablets of calibration and validation sets and the tablets of different batches and strength but excluding the falsified products.

[§] Two products labelled as chloroquine tablets, containing no chloroquine but 124 mg or 14 mg of metronidazole, respectively [42].

[#] 32 products containing 30 additional active pharmaceutical ingredients from the WHO Essential Medicines List 2021 [47], purchased in Germany and listed in Table S1. Table S1 (Supplementary Materials) gives detailed information on each of the 170 products.

(2). In twelve of these products, the coating had a different colour than the tablet core. A total of 45 products showed smooth top and bottom sides, 83 had a score line or an embossing on one side, and 40 on both sides. For two falsified products, tablets from the same container showed different kinds of embossing and score lines.

In total, samples of 170 different pharmaceutical products were included into this study. Table S1 gives detailed information about each of these medicines, including brand names and stated manufacturers.

2.2 HPLC analysis

Medicines collected in African countries were shipped to Germany by commercial courier services and were stored at Tübingen University at 21 °C until analysis. Compendial analysis was carried out according to the monographs of the United States Pharmacopeia (USP 41) for the respective medicines. This comprised identification and quantification of the API as well as dissolution testing. For the 105 products collected in Cameroon, the DR Congo and Chad, the procedures and results of this testing have been published [9,35,42].

Additionally, five falsified products labelled as sulfamethoxazole/trimethoprim tablets 400/80 mg had been collected in Nigeria (see above). They were investigated by HPLC-UV analysis for identity and quantity of the APIs, using an Agilent 1260 Infinity II HPLC system (Agilent Technologies, Santa Clara, California, USA), with columns and solvents as specified in the USP 41. Certified pharmaceutical secondary reference standards from Sigma-Aldrich (St. Louis, Missouri, USA) were used. The results of this analysis are listed in Table S1.

2.3 Creation of calibration and external validation sets

For each of the ten APIs or fixed API combinations listed in Table 1, it was attempted to obtain eight products representing different brands containing the same API(s) in the same strength. Using the RAND function of MS Excel®, two of these products were selected randomly for inclusion into the external validation set V1, and the other six products were included into the calibration set.

As summarized in Table 2, further external validation sets were created from tablets belonging to different batches of the brands comprised in the calibration set, or from tablets of different strength than those comprised in the calibration set. Furthermore, an external validation set was created from tablet preparations of 30 additional APIs which were contained in the 2021 WHO Model List of Essential Medicines [47] and were commercially available in Germany. Finally, external validation sets were created from tablets which had been identified as falsified in the course of the aforementioned medicine quality studies [9,35,42] or in the present study.

All products collected in African countries and included into the calibration set or into the validation sets V1, V2 and V3 (Table 2) had been confirmed to comply with the USP specifications for identity, assay (= quantity) and dissolution of the API(s) by analysis in our laboratory.

Table 2
External validation sets for testing the chemometric models developed in this study.

Set	Products
V1	Products representing brands not comprised in the calibration set, but containing the same APIs in the same strength as the products in the calibration set
V2	Products representing different batches of brands comprised in the calibration set
V3	Products of different strength than those comprised in the calibration set
V4	Tablet preparations of 30 additional APIs not contained in the calibration set
V5	Falsified medicines containing no API at all
V6	Falsified medicines containing a different API instead of the declared one
V7	Falsified medicines containing the correct amount of the declared API but carrying a label misrepresenting the source or the expiry date of the product
V8	Falsified medicines containing the declared API in an incorrect amount

2.4 NIR spectrometers

Five NIR-S-G1 spectrometers (devices A-E), produced by InnoSpectra (Hsinchu, Taiwan), were purchased; however, device A had to be excluded due to a technical defect. As described by Crocombe [48] the NIR-S-G1 uses the digital light processing (DLP) technology of Texas Instruments (Dallas, Texas, USA), i.e. it contains a digital micromirror device and a single 1 mm uncooled InGaAs detector [49]. It performs measurements in diffuse reflection mode, in the wavelength range of 900–1700 nm. A performance evaluation of NIR-S-G1 devices, including wavelength and photometric accuracy and repeatability, as well as spectroscopic noise, has been published recently [50].

2.5 Software

ISC WinForms SDK GUI (v3.7.2, InnoSpectra, Hsinchu, Taiwan) was used to control the NIR-S-G1 devices. The Aspen Unscrambler® (V12.2, Aspen Technology Inc., Bedford, MA, USA) was utilized for the initial principal component analyses. PLS Toolbox (v.9.21, Eigenvector Research Inc., Manson, WA, USA) in a MATLAB environment (v.R2023a, The Mathworks Inc., Natick, MA, USA) was used for partial least squares discriminant analysis and piecewise direct standardization. The DD-SIMCA toolbox (v 1.2) [38] was used to compute the one-class classification models, and the MATLAB code for the Procrustes cross-validation was downloaded from the GitHub repository indicated by Kucheryavskiy et al. [51].

2.6 Spectra acquisition

For spectra acquisition, the NIR-S-G1 device was fixed in an upright position, either by using a 3D printed device holder (Fig. S1, Supplementary Materials) or by attaching self-adhesive plastic feet to the rounded bottom of the device. The device was connected to a computer using a USB cable. Tablets were removed from blisters and placed directly onto the sapphire scan window for spectra acquisition. Any vibrations or movements were avoided during the measurements.

The spectrometer was switched on one hour before the first measurement. Ten initial blind scans were then performed to ensure a system temperature between 30 and 40 °C for all measurements. The employed NIR-S-G1 instrument settings were: Hadamard mode; spectral range 900–1700 nm; pattern width 7.03 nm; digital resolution 583; exposure time 1.27 ms; 16 repeats. Spectra were saved as comma separated values (CSV) files. Four spectra were acquired of each tablet, two each from the bottom and the top side, with a vertical flip in between the two measurements. Of each product, two tablets were investigated.

Calibration set products were first measured by investigator BM using the devices C and D. After two weeks, measurements were repeated by investigator GG using devices C and B. Thereby, 32 spectra were acquired from each calibration set product. Another two weeks later, products of validation set V1 were measured by investigator JG using devices C and E. Thereby, 16 spectra were acquired from each product of this validation set. In this way, procedures were validated by measurements carried out on different days, by different persons and using different pieces of equipment, following the recommendations of the International Council for Harmonisation (ICH) [52] and the USP [53].

Products of the validation sets V2 – V8 (Table 2) were measured by investigator BM using device C. Eight spectra were acquired from each product. However, for two products representing falsified medicines, only one tablet was left after chemical analysis, and therefore only four spectra were recorded. In total, 2928 spectra were recorded and included into the data analysis.

Between measurements, tablets were stored at 21 °C in a dark place, in low-density polyethylene zip lock bags which were placed in a high-density polyethylene box together with desiccant silica gel.

2.7 Spectral data pretreatment

Each NIR-S-G1 device records data points at slightly different wavelengths, spaced in unequal steps. Therefore, the raw NIR spectra were interpolated in 1 nm intervals using natural cubic splines with the Aspen Unscrambler® software. According to the NIR-S-G1 manufacturer's information [49], the photosensitivity of the device is low at wavelengths <950 nm, and it changes with temperature at wavelengths >1650 nm, resulting in noise in these ranges. Furthermore, the common excipient talcum exhibits a prominent absorption peak at 1391 nm (Fig. S2) which interfered with the desired classification of spectra according to the APIs (rather than according to excipients). Therefore, only the spectral ranges 950–1375 nm and 1405–1650 nm were used for chemometric analyses.

Different pre-processing methods were tested, i.e. a Savitzky-Golay (SG) algorithm [54] using either no derivative or 1st or 2nd derivative; using 2nd, 3rd, 4th or 5th order polynomial; and using a window size of 3, 7, 13, 15 or 21. Either no standard normal variate (SNV) transformation [55] was applied, or SNV before or after SG. The best separation of different APIs in PCA and DD-SIMCA was obtained using SG with 2nd or 1st derivative, window size 13 or 15, and SNV performed after SG. Therefore, for most chemometric analyses SG (2, 2, 13) followed by SNV was used. However, for the spectral data of amoxicillin, amoxicillin/clavulanic acid and ciprofloxacin, SG with the 1st derivative and a window size of 15 was chosen, since using SG with the 2nd derivative resulted in the observed numbers of extremes falling outside of the tolerance limits for many values of α in the extreme plots (see Section 2.9). Spectra were mean centred before analysis.

2.8 Principal component analysis (PCA)

PCA models were computed from the spectra of APIs in the calibration set using pretreated spectral data. The singular value decomposition (SVD) algorithm was applied with cross validation. The results were visualized in the respective PC-1 vs. PC-2 scores plots.

2.9 DD-SIMCA analysis

In DD-SIMCA analysis [37,38], the first step is a PCA of the calibration samples for the investigated target class (in the present case, of the spectra of the calibration set products containing a specific API). Subsequently, for each spectrum a score distance h_i and an orthogonal distance v_i are calculated from the PCA results. These two distances are scaled and combined in a total distance that follows the scaled Chi-squared distribution. Based on the calibration set, the parameters of the Chi-squared distribution (degrees of freedom and scaling factor) are computed, and an acceptance area is calculated using a chosen type I error α. The calculated model is represented by an acceptance plot, usually with the x- and y-axis in logarithmic scale, i.e. $x = \log(1 + h_i/h_0)$ and $y = \log(1 + v_i/v_0)$. New samples are projected onto the acceptance plot and can be classified as members or non-members of the target class according to their projection falling inside or outside of the acceptance area, respectively.

Furthermore, extreme plots can be generated from the calibration set data by plotting the theoretically expected and the observed numbers of extreme samples for different values of the type I error α. The performance of the model is satisfactory if, for all values of α, the observed numbers of extreme samples are inside of the tolerance limits calculated using a binomial distribution [37,38,56].

In the present study, models were built applying the rigorous approach [56,57], with the probability of type I error set at $\alpha = 0.01$. The γ value for definition of the outlier area [37] was set at 0.01. Outlying spectra were identified with DD-SIMCA using the "robust" method [37,38], and were removed from the datasets. Subsequent analyses were carried out using the "classical" method [37,38]. The optimal number of principal components was chosen based on both the extreme plots (E-plots) and the sensitivity plots, using the calibration and the Procrustes cross-validation datasets [51,58]. The model with the E-plots showing the best extreme pattern (as close as possible to the tolerance corridor) and the sensitivity closest to the 100 × (1-α)% value were selected [56,59].

Subsequently, spectra of products of the external validation set V1 were projected onto the respective model. The models were further tested for sensitivity using products of external validation sets V2 and V3 (see Table 2) containing the same API, and for specificity using spectra of all products containing other APIs in the calibration set and in external validation sets V1, V2, V3 and V4. The results were visualized in acceptance plots.

2.10 Partial least squares discriminant analysis (PLS-DA)

A PLS-DA model was computed to separate the classes of amoxicillin and amoxicillin/clavulanic acid. The optimal PLS dimension (four latent variables) was determined through a Venetian blind cross-validation process and confirmed by predictions of the Procrustes cross-validation dataset.

2.11 Piecewise direct standardization (PDS)

PDS was used to compute models for calibration transfer between the employed spectroscopic devices. Device C was chosen as "master", and devices B and D as "slaves". Using the spectra of selected samples (see Section 3.10), a transform matrix was computed and employed to transfer spectra measured on the "slave" devices. The optimal window size, i.e. the "slave" spectrum's spectral range on which the transform matrix was computed, was selected as follows: a Procrustes cross-validation dataset (Xpv_{slave}) was computed for the slave transfer dataset. Subsequently, PDS models with window sizes between 1 and 201 were computed and applied on the Xpv_{slave} dataset. The transferred datasets were projected onto the DD-SIMCA model of the "master" device, and sensitivity was determined. The window size resulting in the highest sensitivity was selected.

2.12 Definitions

In this study, the current definitions of substandard and falsified medicines by the World Health Organization (WHO) were used [1,2]. Where necessary, additionally the criteria suggested by Hauk et al. [60] and by Ozawa et al. [36] were applied.

Calibration set spectra were classified as regular objects, extreme objects or outliers as described by Pomerantsev and Rodionova [37,56]. For the DD-SIMCA models, the *a posteriori* sensitivity was calculated as follows [38]:

$$\text{Sensitivity } (\%) = 100\ \% \times (n_{\text{spectra}} - n_{\text{extreme objects}})/n_{\text{spectra}}$$

Rigorous DD-SIMCA models [57] are based on the target dataset alone, therefore no specificity values can be calculated for these models [56]. When a model was used for predictions using a test set of samples outside the target class, the specificity value for this model and this test set was calculated as follows [38]:

$$\text{Specificity } (\%) = 100\ \% \times (n_{\text{external objects}})/n_{\text{samples}})$$

New samples were classified as external objects if at least half of their spectra were projected outside of the acceptance area of the respective DD-SIMCA model.

3. Results

3.1 Visual comparison of acquired NIR spectra

Fig. 1 shows the NIR spectra of different brands of penicillin V,

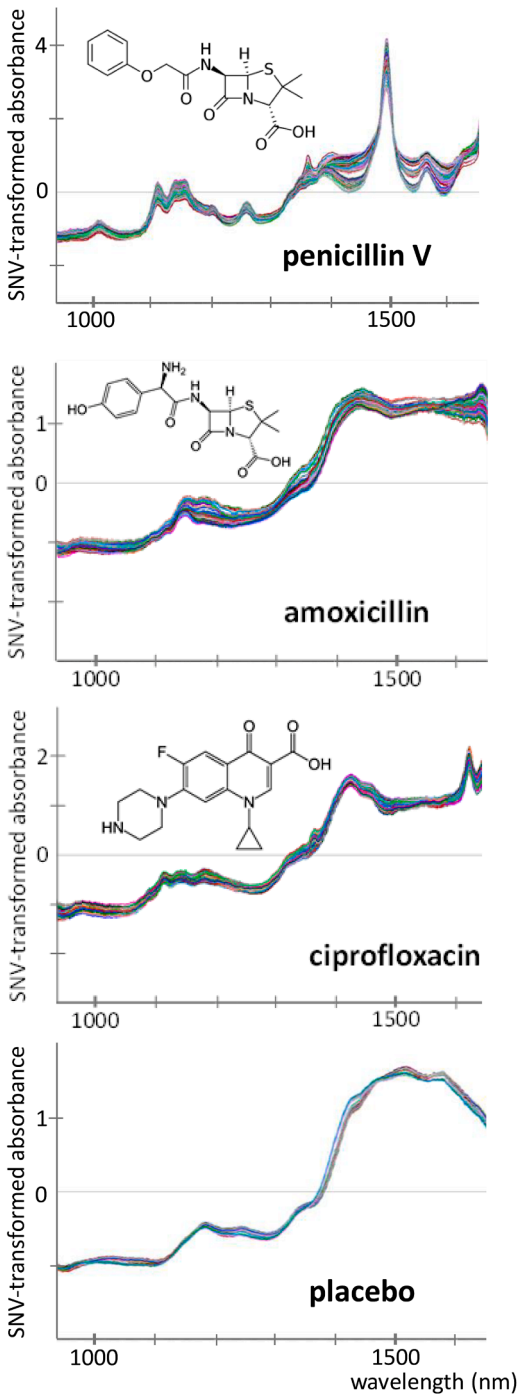

Fig. 1. Near-infrared spectra (950 - 1650 nm) of penicillin V tablets (11 products), amoxicillin tablets (10 products), ciprofloxacin tablets (10 products) and placebo tablets (4 products).

amoxicillin, ciprofloxacin, and placebo tablets. NIR spectra of tablets depend both on the chemical composition, i.e. on the APIs as well as on the excipients, and on physicochemical characteristics such as crystallinity, particle size, moisture etc. [19]. Excipients and physicochemical characteristics are different between products from different manufacturers, and this results in differences between the NIR spectra of different brands. Nevertheless, Fig. 1 clearly suggests that for tablets containing e.g. penicillin V or ciprofloxacin, the spectral features in the investigated wavelength range predominantly depend on the respective API. Amoxicillin tablets showed fewer distinctive absorption peaks in this wavelength range, and the investigated placebo tablets showed even less. Fig. S2 depicts the spectra of all products included into the calibration set of this study. Most of them show API-related differences to each other and to the spectra of the placebo tablets. This suggests that spectra recorded on the low-cost NIR-S-G1 device may allow verification of the presence and type of APIs contained in tablet products, provided the respective API shows absorption bands in the range of 950–1650 nm. Interestingly, the spectra of penicillin V and amoxicillin tablets are very different, despite the considerable chemical similarity between these two beta-lactam antibiotics (Fig. 1).

3.2 Principal component analysis (PCA)

For a first exploratory analysis, the spectra of all products in the calibration set, recorded with device C, were investigated by PCA. Fig. 2A displays the result as a PC-1 vs. PC-2 scores plot. Five of the ten APIs or API combinations were quite well separated from each other. The spectra of those products were removed from the data set and a second PCA (Fig. 2B) was carried out. This led to a separation of most of the remaining APIs. However, the data points of amoxicillin and amoxicillin/clavulanic acid tablets still overlapped. Even when a PCA with spectra only of these two types of tablets was conducted, and up to seven principal components were investigated, their separation could not be achieved.

As expected, the NIR spectra of different tablet brands (containing the same API in the same amount) still showed differences from each other in their PC scores, prominent especially in the case of hydrochlorothiazide and metformin tablets. For metformin, six brands were comprised in the calibration set. All contained 500 mg metformin per tablet, but the average tablet weight of the brands varied from 533.4 to 649.1 mg, due to different amounts of excipients used by the manufacturers. Fig. 2B shows for each of the six brands the mass percentage of metformin in the total tablet weight. As clearly visible, the PC-1 scores were correlated to the mass percentage of the API in the tablets. For hydrochlorothiazide, only three brands were available for inclusion in the calibration set. All three brands contained 50 mg API per tablet, but they showed extremely different API mass percentages in the total tablet weight: 15.9 %, 35.0 %, and 50.3 %, respectively. Correspondingly the spectra of three brands showed large differences in their PC scores (Fig. 2A). These observations exemplify that API mass percentage is an important, albeit not the only factor influencing the NIR spectra of different commercial products containing the same API.

No or only very small differences were observed between spectra recorded from different sides of the tablets, even when the tablets had embossings or score lines on one side. Notably, even coloured tablets (e.g. of furosemide or metronidazole; Table S1) did not show noticeable differences from white tablets in their PC scores.

The spectra of the calibration set had been recorded using three separate NIR-S-G1 devices (see Section 2.6). For a further analysis, all spectra recorded with devices B, C, and D from the calibration set products and placebos were pooled and investigated by PCA. The results are shown in Figs. 2C and 2D. As expected, the resulting model showed more variation than the model obtained with device C alone, but it still allowed a similar separation of the ten APIs and the placebos in the same two PCA steps. However, the variation introduced into the model by the use of different devices suggests that an adjustment of the spectra from different devices with calibration transfer methods, e.g. by piecewise direct standardization, may be useful to improve the accuracy of the model [61]. In contrast, no influence was detected regarding the investigator operating the device.

A different approach which led to a somewhat better separation of the spectra of the different APIs recorded with the three devices is illustrated and explained in Fig. S3. In that approach, nine consecutive PCA steps were used instead of two, and after each step the spectra of one of the APIs were removed from the dataset.

In summary, these PCA analyses suggested that a verification of the presence of the declared API in different tablet brands using the low-cost

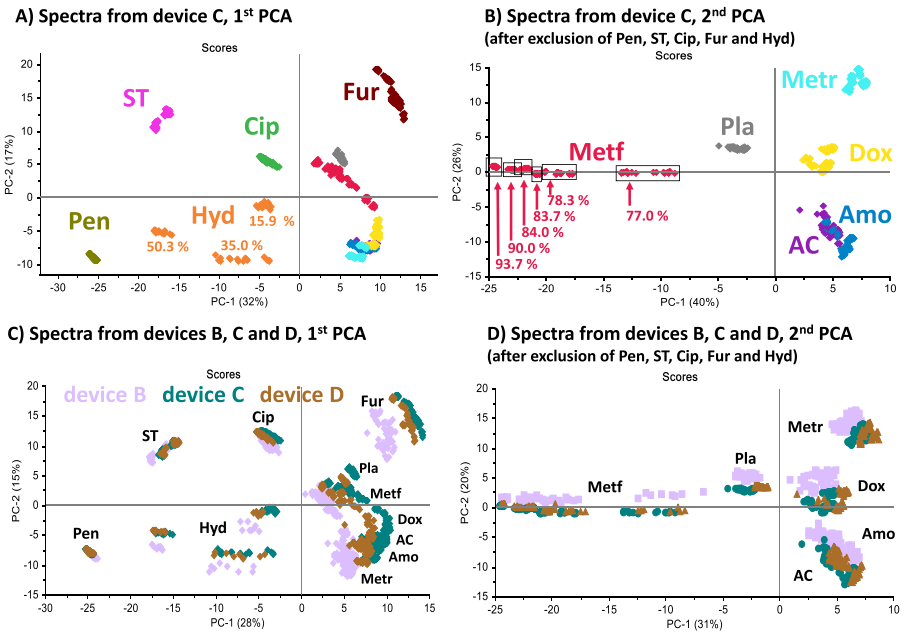

Fig. 2. Principal component analysis of the spectra of tablets from all active pharmaceutical ingredients (APIs) in the calibration set. A) and B): Spectra recorded with device C. Different colours denote different APIs. C) and D): Spectra recorded with devices B, C and D. Different colours denote different spectrometric devices. For hydrochlorothiazide and metformin tablets, the mass percentage of the API in the total tablet weight is indicated for each of the investigated brands. Amo, amoxicillin; AC, amoxicillin/clavulanic acid; Cip, ciprofloxacin; Dox, doxycycline; Fur, furosemide; Hyd, hydrochlorothiazide; Metf, metformin; Metr, metronidazole; Pen, penicillin V; ST, sulfamethoxazole/trimethoprim; Pla, placebo.

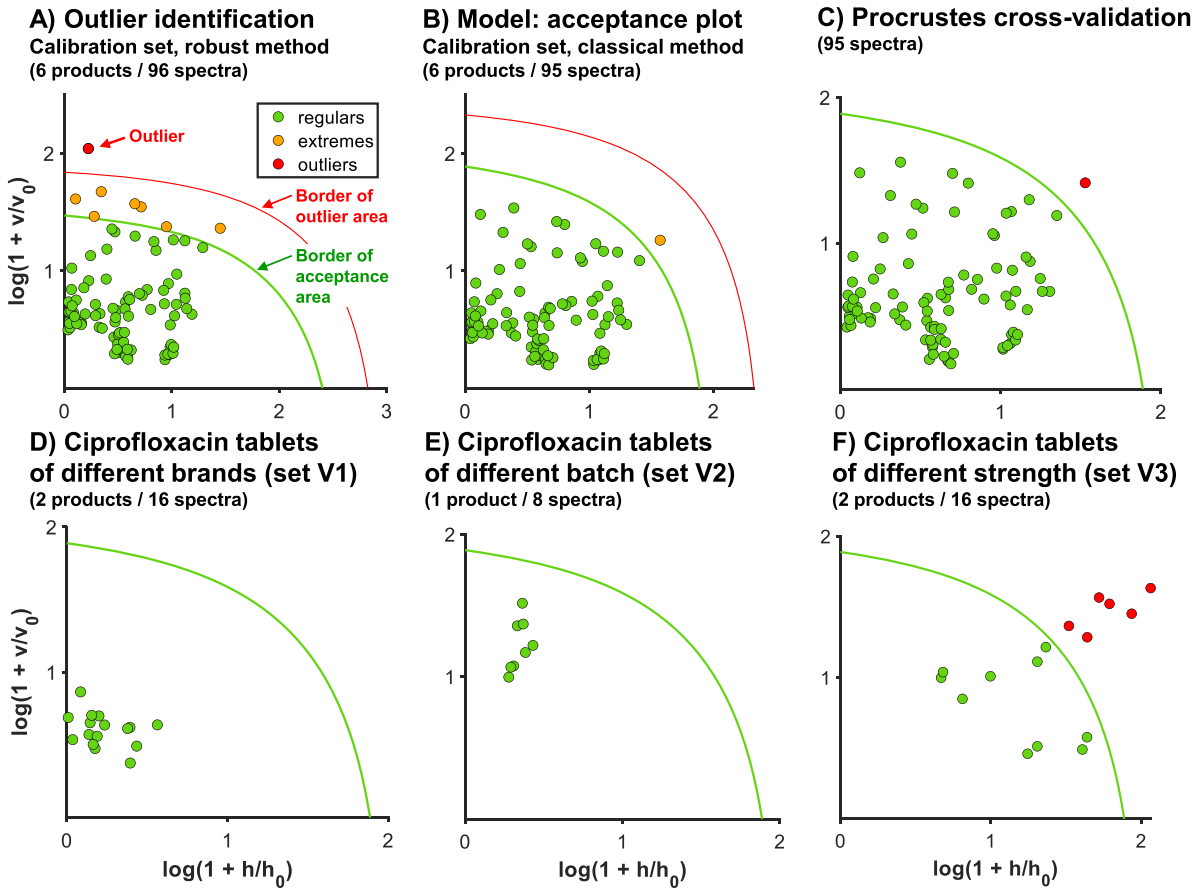

Fig. 3. DD-SIMCA analysis of the spectra of ciprofloxacin tablets. Spectra were recorded with device C. $\alpha = 0.01$ and two prinicipal components were used for model building with the "classical" method [37]. Panels D-F: projection of spectra of further ciprofloxacin products onto the DD-SIMCA model.

Table 3
Parameters for DD-SIMCA models, their sensitivity, and their specificity.

	Pre-processing	α	No. of PCs	No. of outliers removed	Sensitivity calibration (% of spectra)	Sensitivity PCV (% of spectra)	Sensitivity different brand (% of products)	Sensitivity different batch (% of products)	Sensitivity different strength(% of products)	Specificity different or no API (% of products)
Penicillin V	SG (2,2,13) + SNV	0.01	2	0	100.0 %	100.0 %	2/2 (100 %)	1/2 (50 %)	1/3 (100 %)	137/137 (100 %)
Sulfamethoxazole/ trimethoprim	SG (2,2,13) + SNV	0.01	2	0	100.0 %	100.0 %	1/2 (50 %)	1/1 (100 %)	1/1 (100 %)	140/140 (100 %)
Ciprofloxacin	SG (**1**,2,**15**) + SNV	0.01	2	1	98.9 %	98.9 %	2/2 (100 %)	1/1 (100 %)	1/2 (50 %)	139/139 (100 %)
Furosemide	SG (2,2,13) + SNV	0.01	1	0	100.0 %	100.0 %	1/2 (50 %)	2/2 (100 %)	n.d.	140/140 (100 %)
Metronidazole	SG (2,2,13) + SNV	0.01	1	1	100.0 %	100.0 %	2/2 (100 %)	2/3 (67 %)	1/3 (33 %)	136/136 (100 %)
Metformin	SG (2,2,13) + SNV	0.01	1	0	100.0 %	100.0 %	1/2 (50 %)	2/2 (100 %)	3/4 (100 %)	136/136 (100 %)
Hydrochlorothiazide	SG (2,2,13) + SNV	0.01	2	2	95.7 %	95.7 %	1/1 (100 %)	1/1 (100 %)	1/5 (20 %)	140/140 (100 %)
Doxycycline	SG (2,2,13) + SNV	0.01	2	0	100.0 %	100.0 %	2/2 (100 %)	2/3 (67 %)	n.d.	139/139 (100 %)
Amoxicillin & amoxicillin/ clavulanic acid	SG (**1**,2,**15**) + SNV	0.01	4	0	99.4 %	98.9 %	3/4 (75 %)	3/3 (100 %)	1/3 (33 %)	129/129 (100 %)
Total							**15/19 (79 %)**	**15/18 (83 %)**	**9/21 (43 %)**	**100 %**

For specificity calculations a test set comprising the spectra of all other APIs (and placebos) in this study was used. PC, principal component; PCV, Procrustes cross-validation; SG, Savitzky-Golay algorithm; SNV, standard normal variate. SG parameters: first digit, derivative; second digit, polynomial order; third digit, window size.

NIR-S-G1 device may be possible, and encouraged us to expand the investigations using further, more adapted chemometric methods.

3.3 DD-SIMCA model building and Procrustes cross-validation

While PCA is an unsupervised method which investigates the data without prior knowledge on their class membership, data-driven soft independent modelling of class analogy (DD-SIMCA) is a one-class classification method. It distinguishes objects from one target class (in this case, tablets containing a certain API) from all other objects and classes, without the necessity of known samples of the other classes. This is important in the screening for substandard and falsified (SF) medicines, as the composition of such SF medicines is not known *a priori*. DD-SIMCA is particularly useful in the verification of the identity of products [43,57]. The basic principles of DD-SIMCA are explained in Section 2.9.

Models were computed for each API using the spectra of the calibration set obtained with NIR-S-G1 device C. Fig. 3 shows as an example the resulting plots for ciprofloxacin. Initial analysis using the "robust" method (based on median and interquartile range statistics) [37] identified one outlier (Fig. 3A). This was removed from the dataset, and subsequently the optimal number of principal components (PCs) was selected comparing DD-SIMCA models built using the "classical" method (based on mean and variance statistics) [37,38].

Fig. 3B shows the acceptance plot of the final model, built using two PCs. As visible, all but one of the spectra are inside the acceptance area, i.e. sensitivity is 98.9 % which is good since the expected sensitivity is 99 % (for $\alpha = 1$ %) (Table 3).

DD-SIMCA models were internally validated using Procrustes cross-validation (PCV), a novel approach for the validation of chemometric models. PCV has been suggested especially for short datasets as a bridge between conventional cross-validation and the use of independent validation sets [51,58]. Fig. 3C shows the result of the PCV for ciprofloxacin. The agreement between the acceptance plot and the PCV plot is very good.

The quality of models built with different numbers of PCs was also investigated for each API using and comparing extreme plots [38] for the models and the PCVs (see Section 2.9). As shown in Fig. S4, in all cases the extreme plots for the model and the PCV were very similar, and the observed numbers of extremes were within the tolerance limits for all values of α. Table 3 shows that for all APIs it was possible to build models (using $\alpha = 0.01$) showing an *a posteriori* sensitivity close to the expected value of $100 \times (1 - \alpha)$% both for the model and the PCV, and this further confirmed the quality of the models. As described below, spectra of amoxicillin and amoxicillin/clavulanic acid tablets could not be reliably separated from each other and were therefore combined into a single class, later to be separated by partial least squares discriminant analysis.

3.4 External validation of the models

In the present study, for most APIs eight commercial products of the same strength could be obtained for inclusion in the calibration and validation sets. For each API, two of these products were selected randomly for the validation set V1 and six for the calibration set (see Methods). As an example, the results for ciprofloxacin external validation set V1 are shown in Fig. 3D. All eight spectra of each of the two products of this validation set were projected into the acceptance area of the ciprofloxacin model, therefore both products were classified correctly. Obviously, six calibration set products are not enough to represent the variability of all commercial tablet products of this API on the market, therefore a limited sensitivity of this external validation was to be expected. As an example, for metformin the spectra of only one of the two products of validation set V1 were projected into the acceptance area (Fig. S5). The spectra of the other product were projected outside of the acceptance area resulting in a sensitivity of 50 % for the external

validation of the metformin model. The results for all APIs are summarized in Table 3 and visualized in Fig. S5. Out of 19 products in the external validation sets V1 for the different APIs, 15 (79 %) were correctly classified as members of the respective target class. Out of the four products which were classified as non-members, three showed API mass percentages out of the range represented in the respective calibration set. This exemplifies that, as mentioned above, API mass percentage is an important, albeit not the only factor influencing the NIR spectra of different commercial products containing the same API.

For 16 of the 19 products of validation sets V1, their classification as members or non-members of the target class was supported by the projected position of all eight spectra of this product, and for the other three products by seven out of eight spectra (Fig. S5).

3.5 Testing of the models with different batches of brands contained it the calibration set

Eighteen samples were available which represented the same brands and strengths as contained in the calibration set, but different batches thereof (= external validation set V2, Table 2). Fifteen (83 %) of these samples were correctly classified as members of the target class (Table 3). For ciprofloxacin, this is exemplified in Fig. 3E.

In striking contrast, all eight spectra of a metronidazole sample (stated manufacturer CSPC Ouyi Pharmaceutical Co. Ltd., China; batch no. 825160701) were projected to a position quite far outside the acceptance area (Fig. S6), very different from the corresponding sample of the same brand in the calibration set (batch no. 825170302). Both samples had been collected in the DR Congo, and both had passed analysis for correct identity, content, and dissolution of the API in our laboratory. However, the average tablet weight of the two batches was very different (293.7 mg and 340.6 mg, respectively). Therefore, the API mass percentage of the first-mentioned batch resulted as 85.1 %, which was higher than for any of the six samples of the calibration set. The most likely explanation for this observation is that the manufacturer had used different manufacturing protocols for the two different batches of the same brand. A similar, though less pronounced observation was made for a doxycycline sample, again from CSPC Ouyi Pharmaceutical Co. Ltd. Also in this case the average tablet weight of the batch in the external validation set V2 was different from that of the batch in the calibration set (173.0 mg and 186.5 mg, respectively), and four of its eight spectra were projected to a position just outside of the acceptance area (Fig. S6).

Finally, a penicillin V sample from external validation set V2 was classified as non-member of the target class, since six of its eight spectra were projected to a position just outside of the acceptance area (Fig. S6). This sample, and the corresponding batch in the calibration set, showed the lowest API mass percentage of all products used for model building. It appears possible that small differences in composition or physicochemical properties of this penicillin V sample may have resulted in projection of its spectra outside of the acceptance area.

3.6 Testing of the models with tablets of different strength than those comprised it the calibration set

Into the calibration and validation sets, for each API only products of identical strength had been included. For further testing, a total of 21 products were available containing the same API in different strength than the products in the calibration set (external validation set V3, Table 2). 17 of these 21 products were from different manufacturers than those in the calibration set (Table S1). 9 of the 21 products (43 %) were classified as members of the target class of the respective API, by projection of their spectra onto the DD-SIMCA models (Table 3 and Fig. S7). Notably, for many of the 21 products, the position of the projections inside or outside of the acceptance area was related to the mass percentage of the API in the total tablet weight in comparison to the calibration set. E.g. in case of ciprofloxacin (Fig. 3F), the six spectra projected outside of the acceptance area belonged to a sample with a mass percentage outside of the range covered in the calibration set.

3.7 Testing of the models with tablets containing different APIs

Next, the DD-SIMCA model for each API was tested by projecting the spectra of products of all other APIs (and placebos) onto this model, to investigate whether these products would be correctly identified as non-members of the respective target class. The projected spectra were obtained from the products included in the calibration set and the validation sets V1, V2 and V3 of the other investigated APIs (Tables 1 and 2), but also from products of 30 additional APIs (= external validation set V4, Table 2), not comprised in the calibration set. In total, >1.250 spectra from ≥129 products representing 40 different APIs or API combinations were projected onto each model. The results are shown in Fig. 4. Without a single exception, all spectra were projected to positions outside of the acceptance areas of the respective models, in most cases very far outside.

External validation set V4 also comprised tablets of moxifloxacin hydrochloride, an API which is chemically closely related to ciprofloxacin hydrochloride. These two APIs were clearly distinguishable in the DD-SIMCA analysis (Fig. 4), while they had not been separated using projections of moxifloxacin hydrochloride spectra onto the initial, exploratory PCA models (data not shown). As further shown in Fig. 4, tablets containing free moxifloxacin base were clearly separated from those containing moxifloxacin hydrochloride. (All “ciprofloxacin” samples in the calibration set contained ciprofloxacin in form of its hydrochloride, see Table S1.) External validation set V4 also comprised tablets of xipamide, which is chemically closely related to furosemide. The xipamide spectra were projected to a position clearly outside of the furosemide acceptance area, though not very far outside (Fig. 4). Furthermore, external validation set V4 comprised one product representing doxycycline monohydrate tablets. These were very clearly separated from the calibration set tablets which contained doxycycline hyclate, i.e. doxycycline hydrochloride hemiethanolate hemihydrate (Fig. 4).

3.8 Discrimination of spectra of amoxicillin and amoxicillin/clavulanic acid tablets by partial least squares discriminant analysis (PLS-DA)

As already observed in the initial exploratory PCA analysis, the separation of spectra of amoxicillin tablets from those of amoxicillin/clavulanic acid tablets was problematic. Also in the DD-SIMCA analysis, spectra of amoxicillin/clavulanic acid tablets were projected to positions close to the acceptance area of the amoxicillin model, and some spectra of amoxicillin tablets were projected even into the acceptance area of the model for amoxicillin/clavulanic acid (Supp. Fig. S8). Very similar problems have been reported previously [40]. Therefore, in the DD-SIMCA analyses described above, we had combined these APIs into a single class. Their separation was now attempted using partial least squares discriminant analysis (PLS-DA). PLS-DA typically outperforms SIMCA in classification rates, provided that within-class variability is low [62]. The classification was performed using two classes, i.e. amoxicillin and amoxicillin/clavulanic acid tablets. The calibration set spectra from device C were used to compute the PLS-DA, and the models were then tested by Procrustes cross-validation as well as against the spectra of the external validation set V1 (recorded with device C). The results are shown in Table 4: all spectra from the calibration set were classified correctly, and so were the spectra from the three products of validation set V1 which had been identified by DD-SIMCA as members of the combined amoxicillin and amoxicillin/clavulanic acid class. One further validation set product of amoxicillin/clavulanic acid had been classified, by projection onto the DD-SIMCA model, as a non-member of that combined class (Fig. S5) and was therefore not included into the PLS-DA.

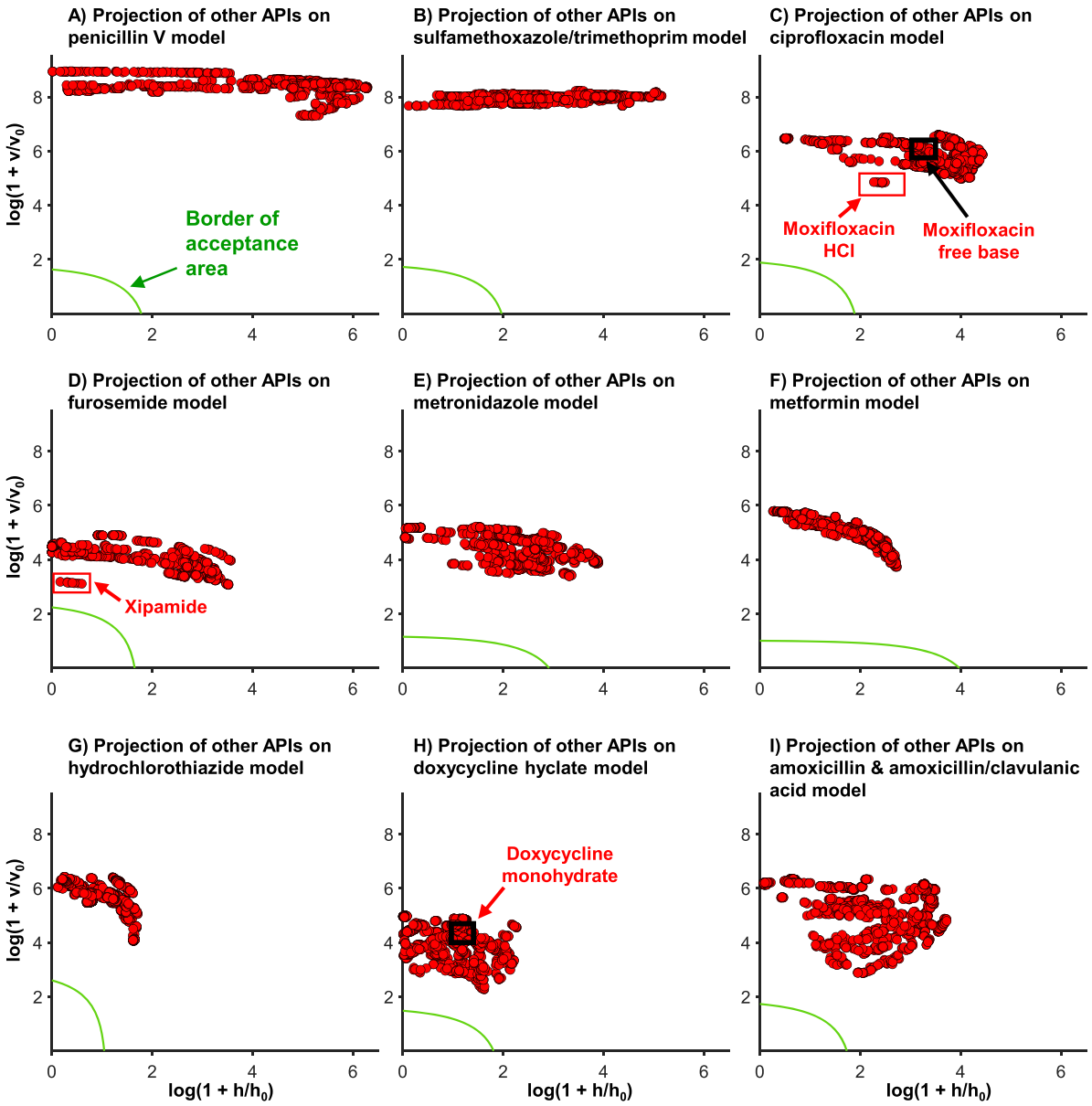

Fig. 4. Testing of DD-SIMCA models for each active pharmaceutical ingredient (API) by projecting the spectra of all other APIs (and placebos) used in this study (40 APIs or API combinations, i.e. ≥129 products resulting in >1.250 spectra) onto the respective model. It should be noted that all scales are logarithmic. Models were built using spectra recorded with device C.

Table 4
Results obtained from the partial least squares discriminant analysis (PLS-DA) of the spectra of amoxicillin and amoxicillin/clavulanic acid tablets (recorded with device C).

API	Calibration set			External validation set V1	
	No. of spectra	Calibration set correct identification	PCV correct identification	No. of spectra	Correct identification
Amoxicillin	96	100 %	100 %	16	100 %
Amoxicillin/ clavulanic acid	80	100 %	100 %	8	100 %

PCV, Procrustes cross-validation.

3.9 Testing of the models with falsified medicines

The key purpose of the present study was to examine whether the developed method would be useful for the identification of falsified (and possibly also substandard) medicines in low-resource settings. Twenty samples of falsified medicines were available (Table 5). They had been collected in African countries and analysed according to the Unites States Pharmacopeia in our laboratory. The spectra of all 20 falsified products were projected onto each of the DD-SIMCA models. The results are shown in Fig. 5 and were consistent with the expectations.

The spectra of seven falsified products which had been found to contain the correct amounts of the labelled active ingredients but to

Table 5
Falsified medicines investigated in this study.

Product code	Stated API	Stated strength [mg]	Determined API content [mg]	Validation set (Table 2) and observed deficiency	Country of collection	Reference
Pen_Oxf	Penicillin V	500	50 mg paracetamol	V6 - incorrect API	Cameroon	[35]
ST_Cit	Sulfamethoxazole/ trimethoprim	400/80	0	V5 - no API	Nigeria	this study
ST_Opt	Sulfamethoxazole/ trimethoprim	400/80	0	V5 - no API	Chad	[9]
ST_Rot	Sulfamethoxazole/ trimethoprim	400/80	0	V5 - no API	Nigeria	this study
ST_Wel	Sulfamethoxazole/ trimethoprim	400/80	27 mg paracetamol	V6 - incorrect API	Nigeria	this study
ST_Spr	Sulfamethoxazole/ trimethoprim	400/80	as stated on label	V7 - false manufacturer stated	Cameroon	[60]
ST_Gau	Sulfamethoxazole/ trimethoprim	400/80	414/19	V8 - insufficient API	Nigeria	this study
ST_Mao	Sulfamethoxazole/ trimethoprim	400/80	191/17	V8 - insufficient API	Chad	[9]
ST_Zun	Sulfamethoxazole/ trimethoprim	400/80	273/54	V8 - insufficient API	Nigeria	this study
Fur1_Mic	Furosemide	40	as stated on label	V7 - false expiry date stated	Cameroon	[60]
Fur2_Mic	Furosemide	40	as stated on label	V7 - false expiry date stated	Cameroon	[60]
Fur3_Mic	Furosemide	40	as stated on label	V7 - false expiry date stated	Cameroon	[60]
Fur4_Mic	Furosemide	40	as stated on label	V7 - false expiry date stated	Cameroon	[60]
Metr_Mac	Metronidazole	200	93 mg metronidazole benzoate	V6 - incorrect API	DR Congo	[35]
ChF_Daw	Chloroquine	250	126 mg metronidazole	V6 - incorrect API	DR Congo	[42]
Chl_Jia	Chloroquine	100	14 mg metronidazole	V6 - incorrect API	Cameroon	[42]
Metr_Mem	Metronidazole	250	as stated on label	V7 - false expiry date stated	DR Congo	[60]
Hyd_Ste	Hydrochlorothiazide	50	5 mg glibenclamide	V6 - incorrect API	Cameroon	[9,70]
AC_Gla	Amoxicillin/clavulanic acid	500/125	0	V5 - no API	Cameroon	[35]
AC_Med	Amoxicillin/clavulanic acid	500/125	as stated on label	V7 - false manufacturer stated	Cameroon	[60]

state a false manufacturer or a false expiry date on their label were projected into the acceptance area of the respective DD-SIMCA model. All the other 13 falsified products were correctly identified as non-members of any of the investigated target classes. Furthermore, as shown in Fig. 5B, three sulfamethoxazole/trimethoprim products which contained insufficient amounts of the APIs were projected to positions outside of the acceptance area. All three products showed API mass percentages out of the range covered in the calibration set. Likewise, as shown in Fig. 5E, spectra of two falsified medicines labelled as chloroquine but containing a low dose of metronidazole (126 mg and 14 mg, respectively, rather than the 250 mg contained in the metronidazole tablets of the calibration set) were projected to a position near to the acceptance area of the metronidazole model. Also, a falsified product labelled to contain (free) metronidazole but found to contain metronidazole in form of its benzoate ester was clearly distinguished from authentic products (Fig. 5E).

In the initial PCA analyses (Fig. 2 and Fig. S3), projections of several falsified tablets labelled as sulfamethoxazole/trimethoprim but found to contain no API at all had not been separated from the spectra of the doxycycline tablets in the calibration set (data not shown). In the DD-SIMCA analysis, this separation was achieved (Fig. 5H).

3.10 Calibration transfer between the different spectroscopic devices by piecewise direct standardization

PCA analysis had shown obvious differences between the spectra recorded with three different NIR-S-G1 devices (Fig. 2C and 2D). This is a common problem in NIR spectroscopy [61,63] and was also observed in DD-SIMCA analysis. An example is shown in Fig. 6A and 6B for ciprofloxacin tablets: in the acceptance plot of the model built using the calibration set spectra recorded with device C, all but one spectrum from device C were correctly located within the acceptance area. In contrast, all calibration set spectra recorded with device B were projected to positions outside the acceptance area of this model, i.e. were not identified as members of the ciprofloxacin class.

One possible approach to deal with this problem is to include the calibration set spectra from some or all instruments into the model building, forcing the calibration algorithm to try to find an equation that is robust to between-instrument variation [61]. Therefore, calibration set spectra recorded with devices B, C and D were pooled and used for building DD-SIMCA models for each API. However, as expected this resulted in a loss of accuracy of the models. This was most strikingly observed in the case of the doxycycline model (Fig. S9, panels A and B). External validation set V4 contained products of 30 different APIs; seven of these were now projected into the acceptance area of the doxycycline model (five with all their eight spectra, and two with some of the eight spectra). Therefore, these products could not be differentiated from doxycycline tablets any more. The loss of accuracy from pooling the spectra of three devices before model building is further illustrated and explained in Fig. S10.

A more suitable approach to address the observed problem of between-instrument variation may therefore be the application of a transfer method, to match the signals from secondary (“slave”) instruments to those of the primary (“master”) instrument. Piecewise direct standardization (PDS) is the most commonly used method for such a transfer [61,63-65]. For PDS, a few samples with high leverage are selected from the set of samples measured on the "master", and remeasured on the “slave” instrument. The differences between the subset spectra obtained on both instruments are then used to compute the transfer parameters [64]. It is obviously of crucial importance that the samples selected for this subset are representative for the whole dataset, and a higher number of samples in this subset leads to better results in the correction of between-instrument variation [64]. In the present study, only six products for each API were available, allowing only a first pilot experiment of PDS. This is exemplified for ciprofloxacin in Fig. 6D-F. Device C was chosen as “master” instrument, and devices B and D as “slaves”. Using PCA, the three calibration set products with the most different spectra were identified (Fig. 6D) and used to compute transfer parameters. These parameters were applied to transfer the spectra of the remaining calibration set products obtained with devices B and D, which were then projected onto the original model (built with the calibration set spectra recorded on device C). As shown in Figs. 6E and

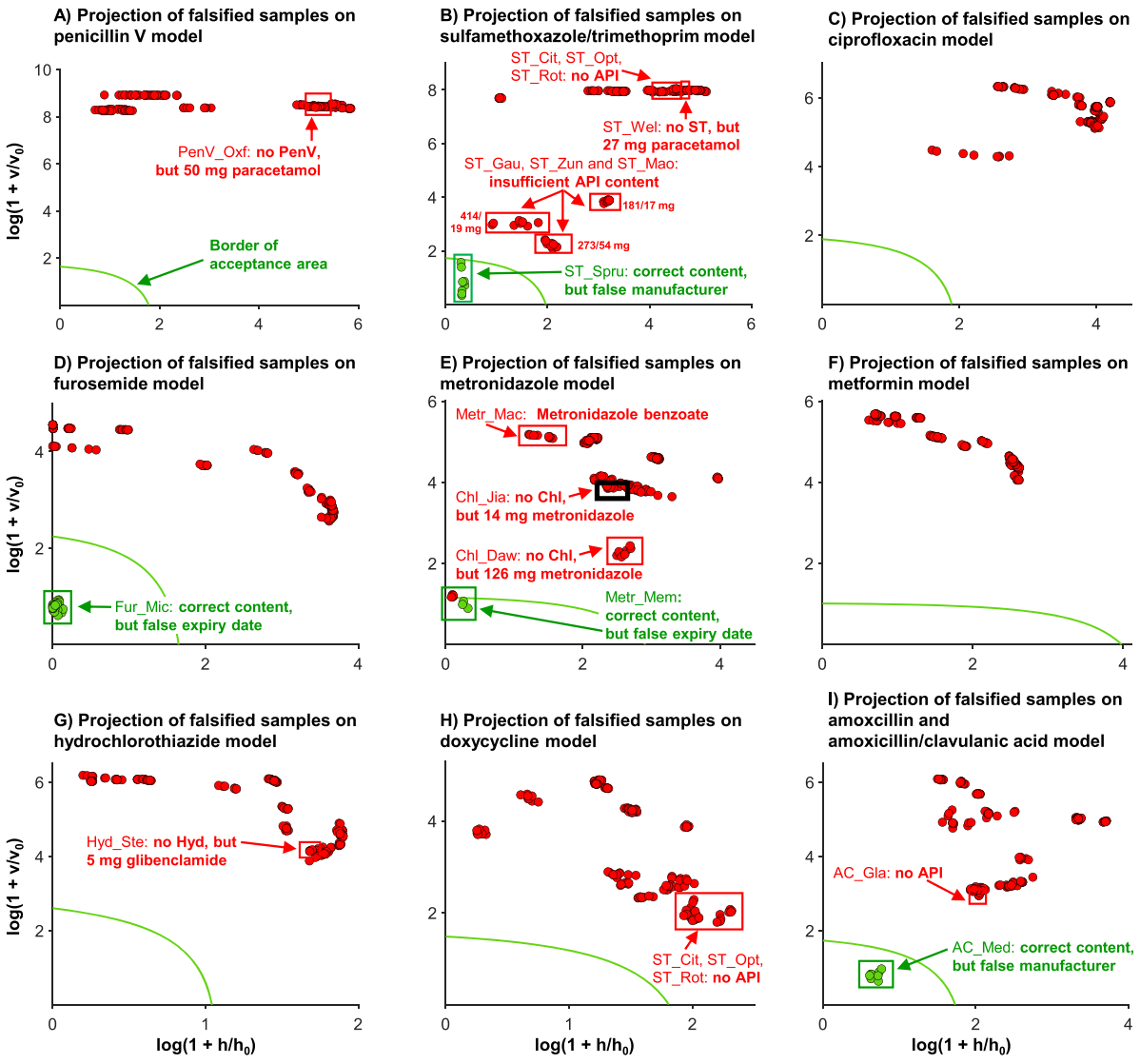

Fig. 5. Projection of the spectra of 20 falsified medicines onto each of the DD-SIMCA models for the active pharmaceutical ingredients (APIs) or API combinations in the calibration set. Chl, chloroquine; see legend of Fig. 2 for abbreviations for the other APIs. Extensions after the API abbreviations (such as PenV_Oxf) refer to the manufacturer stated on the label. Details on the 20 falsified medicines are listed in Table 5. In panel B, the sulfamethoxazole/trimethoprim assay results for the three products with insufficient API content are stated (label claim: 400/80 mg). Product ST_Gau contained an extremely insufficient amount of trimethoprim but a slightly excessive amount of sulfamethoxazole. Due to its low tablet weight, the mass percentage of sulfamethoxazole in the total tablet weight (80 %) was higher than in the six products comprised in the respective calibration set (66–71 %). For products ST_Zun and ST_Mao, the API mass percentage was lower than in the calibration set products.

6F, the transferred spectra were recognized as members of the target class with 100 % sensitivity.

Table 6 shows the parameters used and results obtained for the PDS of all investigated APIs. In most cases, the transferred spectra were recognized as members of the target class with good sensitivity. The only exception were metformin tablets, which had already shown a high variability in the initial exploratory PCA (Fig. 2). Hydrochlorothiazide tablets could not be included into the PDS, as only three products had been available for the calibration set.

These results indicate that the PDS may indeed be suitable to reduce the problem of between-instrument variation in the present investigations, though experiments with a larger dataset will be necessary to provide conclusive evidence.

4. Discussion

The present study investigated the possibilities and limitations of using a low-cost NIR spectrometer combined with chemometric methods for the verification of the presence of the declared APIs in tablet formulations, especially for medicine screening studies in low-resource settings.

Clearly, the NIR spectra of the investigated medicines showed API-specific differences (Fig. 1 and Suppl. Fig. S2), despite of different types and amounts of excipients used by different manufacturers, and despite the limited wavelength range and resolution of the employed NIR-S-G1 spectrometer. DD-SIMCA analysis of these spectra allowed to classify the medicines according to their APIs. When 170 different commercial products representing 40 different APIs or API combinations, as well as placebos, were tested using a single NIR-S-G1 instrument, non-members of the target classes could be identified by DD-SIMCA analysis with 100 % specificity. However, amoxicillin and amoxicillin/clavulanic acid tablets could not be reliably separated from each other by DD-SIMCA analysis. Better results for their separation were achieved by PLS-DA, after the members of a combined amoxicillin and amoxicillin/clavulanic acid class had been identified by DD-SIMCA analysis. Difficulties in the separation of these two medicine types have

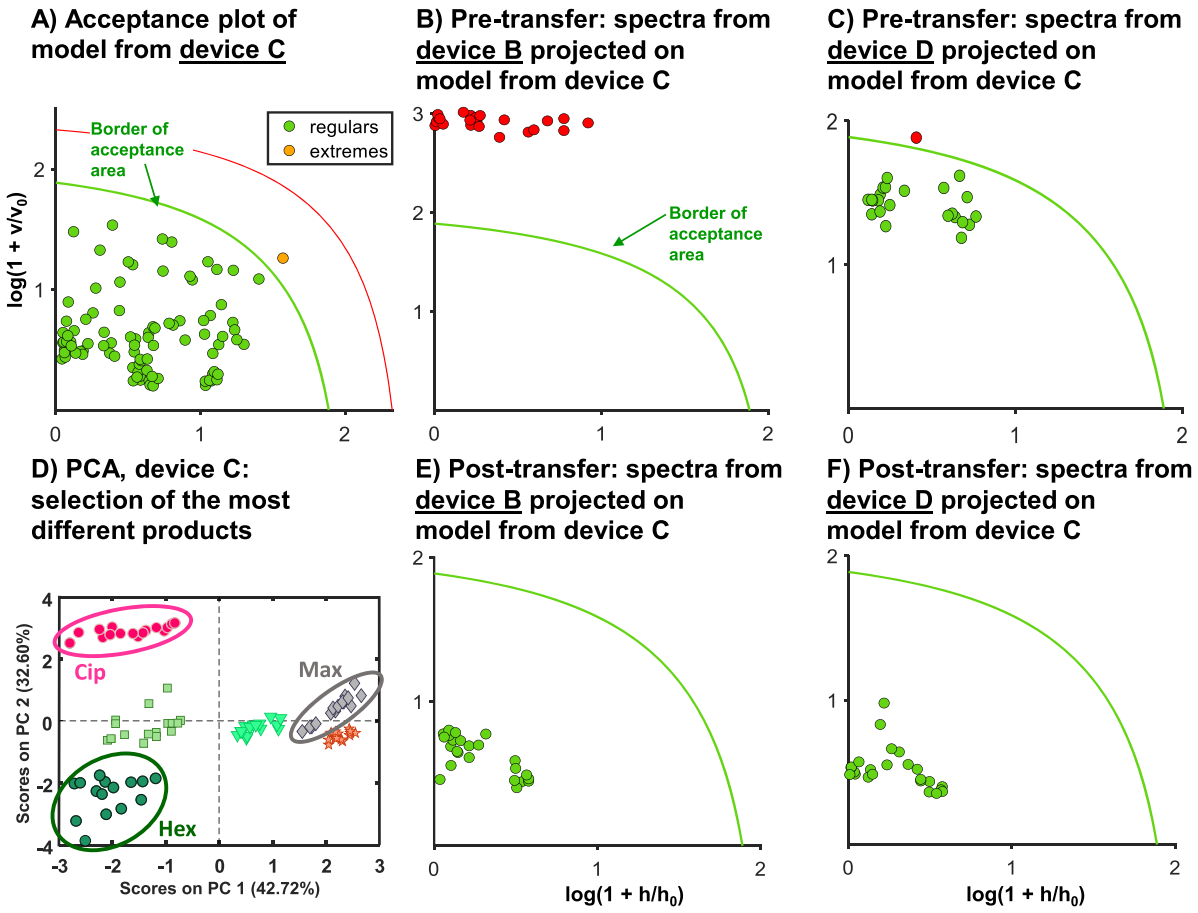

Fig. 6. Testing a possible calibration transfer for spectra of ciprofloxacin tablets (calibration set) by piecewise direct standardization. Using PCA, the three products from the calibration set with the most different spectra (recorded with device C) were identified (panel D) and used to compute the parameters for a subsequent transfer of spectra of the other calibration set products (recorded with devices B and D; panels E and F). Abbreviations for manufacturers: Cip, Cipla Ltd.; Hex, Hexal/SANDOZ S.R.L.; Max, Maxheal Laboratories Pvt. Ltd.

Table 6
Parameters and results of piecewise direct standardization (PDS).

API	Number of transferred spectra (for device B/D)	Device B		Device D	
		Window width	Sensitivity post-transfer	Window width	Sensitivity post-transfer
Penicillin V	24/24	1	100 %	3	100 %
Sulfamethoxazole/trimethoprim	24/24	3	100 %	3	100 %
Ciprofloxacin	23/24	3	100 %	3	100 %
Furosemide	24/24	3	100 %	3	100 %
Metronidazole	24/23	3	100 %	3	96 %
Metformin	24/24	1	67 %	3	79 %
Doxycycline	24/24	3	100 %	3	96 %
Amoxicillin and amoxicillin/clavulanic acid	56/56	1	93 %	3	98 %

Device C was used as “master” instrument, and devices B and D as “slaves”. For each API three calibration set products were used to compute transfer parameters (for Amo and Amo/AC: four calibration set products). These parameters were applied to transfer the spectra of the other calibration set products (excluding outliers).

been reported previously, even using spectrometers of different wavelength ranges and better resolution [40].

In the present study, in most cases the spectra of non-members of the respective target class were projected to positions far from the acceptance areas of that target class (Fig. 4), indicating that the chance of a wrong acceptance of a non-member of the target class (e.g. a falsified medicine containing no or wrong APIs) was very small. This was confirmed by the investigation of falsified medicines which had been previously identified in medicine quality studies in African countries (Fig. 5). The observed performance of the present method was even better than reported from NIR spectroscopic investigations by Tie et al. [40] and by Zambrzycki et al. [13] The latter authors reported that 91.5 % of medicines containing no or incorrect APIs were classified correctly, using the NIR-S-G1 device coupled to a smartphone app termed PillScanNIR. This app was at that time under development by a non-profit organization (Zambrzycki et al. [13], S4 Appendix). Details on the chemometric method employed by this app have not been published.

In the present study, only six different commercial products of each API were used for DD-SIMCA model building, resulting in limited sensitivity: 15 out of 19 products of external validation set V1 (79 %) were correctly identified as members of the target class. As already observed by Ciza et al. [32], a DD-SIMCA model might reject a sample even if it contains the correct API but differs in the quantitative composition or other properties. The present study showed that for the

products of external validation sets V1 and V3 (Supp. Figs. S5 and S7, respectively) many of the rejected samples showed a mass percentage of the API in the total tablet weight that was out of the range covered in the respective calibration set. Therefore, larger and more representative spectral libraries are required for model building, covering products with different API mass percentages and other differences in their excipient composition and physicochemical properties, in order to achieve better sensitivities in the correct classification of members of the target class. For some medicine types, such as hydrochlorothiazide and metformin tablets (Fig. 2A and B), it may even be necessary to build separate models covering products of different API mass percent ranges. In medicine quality studies, the API mass percentage can be readily calculated from the tablet weight and the API content stated on the label.

It may appear tempting to speculate that the differences observed between the NIR spectra of tablets with different API mass percentages could be used to obtain quantitative information, e.g. for the identification of substandard medicines with insufficient API content. However, in diffuse reflectance spectroscopy quantitative API determination requires a precise calibration using tablets with identical excipient composition and physicochemical properties, but with different API contents. Accurate results are not likely to be obtained in the investigation of substandard and falsified medicines, due to the inherent unpredictability of their precise composition. Zambrzycki et al. [13] reported that using the NIR-S-G1 device and the PillScanNIR app, only 30.6 % of medicines containing insufficient quantities of APIs were detected correctly as being substandard. The present method will therefore only detect gross deviations from the declared API content (see Fig. 5B and E). For more precise quantitative NIR investigations of substandard and falsified medicines, transmission spectroscopy using solutions of the investigated products is likely to be more appropriate [43].

For the (qualitative) verification of the presence of the declared API in medicine samples, the method presented in this study offers the typical advantages of NIR spectroscopy, e.g. simplicity and speed of data acquisition, independence from solvents and reagents, and performance without sample destruction. The employed NIR-S-G1 spectrometer was purchased for 1.600 USD per unit for the present study, and even lower prices have been reported in the literature [13]. However, also the limits of the present method must be recognized. E.g. for fixed combinations of two or more APIs, it needs to be investigated to what extent the presence or absence of individual components can be verified. In this study, the presence of the minor component clavulanic acid in fixed amoxicillin/clavulanic acid combinations could not be readily confirmed by DD-SIMCA analysis but required an additional PLS-DA. Further, the present method cannot be expected to be suitable for medicines with very low API amounts, such as oral contraceptives. In the present study, only samples with $\geq$12.5 mg API per dosage unit were included. In the core list of the 22nd WHO Model List of Essential Medicines 2021 [47], 71.8 % of the medicines used as solid oral dosage forms (excluding vitamins and minerals) have an API content of $\geq$ 12.5 mg per unit. Another 10.0 % of the medicines are listed with strengths both higher and lower than 12.5 mg API per unit, and for 18.2 %, all listed solid oral dosage forms contain less than 12.5 mg. Furthermore, the present method requires that the investigated APIs show NIR absorption bands in the investigated wavelength range (950–1650 nm). Doxycycline shows only weak absorption bands in this range (Supp. Fig. S2), and possibly for that reason doxycycline was less perfectly separated from external products than other investigated APIs (Supp. Fig. S10). Further research is required to establish the applicability of the present method to different dosage forms, beyond tablets. In conclusion, the present method can be suitable for many but certainly not for all medicines of interest.

A common problem in NIR spectroscopy is variation between spectra recorded on different spectroscopic devices, even of the same type [61,63]. This was clearly observed in the present study, and led to a loss of accuracy of the chemometric analysis if spectra from different devices were pooled (Figs. 2C and 2D; Supp. Figs. S9 and S10). Notably, with the exception of doxycycline, a complete separation of the investigated APIs from all other APIs was still possible (Supp. Fig. S10). Therefore, at least for APIs with characteristic absorption bands in the investigated wavelength range, using different instruments even without calibration transfer may be possible. However, better accuracy can be achieved using a suitable method for calibration transfer. Fig. 6 and Table 6 indicate that piecewise direct standardization (PDS) [61,63-65] is likely to be effective for such a transfer. However, ideally the determination of the transfer parameters is based on measurement of exactly the same samples under exactly the same conditions on the different instruments [61]. This can be very challenging, especially if the instruments are to be operated in different countries or even different continents. A study on performance qualification procedures for NIR-S-G1 devices used at different geographic locations has recently been published [50], although it does not address the problem of calibration transfer.

The present investigation confirms the results of previous studies, conducted with smaller numbers of APIs, that DD-SIMCA as a one-class classifier (OCC) method is valuable for authentication studies in the screening for substandard and falsified medicines [39,43]. DD-SIMCA has been reported to result in similar sensitivity and specificity as classical SIMCA analysis [39]. (Note that for OCC models, specificity values can only be determined for a given dataset of non-members of the target class). However, using DD-SIMCA in the “compliant” approach [38,57] readily allows to optimize the values for type I and type II errors (i.e. α and β values) for a given model and a given non-member dataset, and such optimizations will be useful once larger spectral databases, representative for a large part of the products on the market, have been established. Procrustes cross-validation (PCV), a novel approach for the validation of chemometric models which is useful especially for small datasets [58,59], was used in this study, and it proved helpful especially in the selection of the optimal number of PCs for each DD-SIMCA model, based on a comparison of sensitivities and extreme plots between model and PCV datasets.

While this manuscript was in preparation, Waffo Tchounga et al. [46] published a study of 292 samples of ciprofloxacin and metronidazole tablets (total 64 brands) collected in Cameroon, which were investigated using a single NIR-S-G1 device followed by chemometric analysis using DD-SIMCA. The results are in very good agreement with those of the present study. The spectroscopic analysis allowed to reliably distinguish tablets containing either one of the two APIs from each other and also from tablets containing chemically related APIs (three nitroimidazoles and four fluoroquinolones). That study also documented the importance of sufficiently large and representative spectral libraries for the sensitive identification of members of the target class, noted especially in case of the metronidazole samples.

The primary use of the method described in this study will be in screening investigations in the post-marketing surveillance of medicines in low-resource settings. Samples failing the screening analysis will have to be further analysed by compendial methods [60]. In this situation, the DD-SIMCA model should reliably classify samples containing no or incorrect APIs as non-members of the target class, i.e. should have a small type II error. In contrast, occasional misclassifications of target class medicines as externals, i.e. type I errors, may be acceptable, as they will be corrected in subsequent compendial analysis.

NIR spectroscopic methods have been explored not only for the authentication of the APIs present in medicines but also for the authentication of specific brands of medicines [29,31,32,66]. Obviously, brand authentication requires a complete and up-to-date library of all commercial products which are present on the respective market. In many low- and middle-income countries, which import medicines from a multitude of sources and have relatively weak medicine regulatory agencies, such complete libraries may be very difficult to establish and to maintain [33]. Reportedly, the governments of China and Russia, i.e. countries with well-controlled domestic markets and strong governance, have established surveillance systems for falsified and substandard

medicines based on NIR-spectroscopic verification of brand authenticity [6,67].

In future applications in low-resource settings, spectral libraries which are not yet complete enough to allow brand authentication may already be sufficient for API verification, especially if such libraries are combined from different countries. International organizations like WHO may try to co-ordinate the establishment of such libraries and provide guidance to different stakeholders e. g. regarding methods for spectra recording, wavelength ranges, suitable instruments and instrument settings, data pre-processing etc. to make spectral data collected by different researchers and institutions comparable worldwide. Ideally, easy-to-use chemometric software tools for field application should be developed, most feasibly for use on laptop computers, and should be made available to government and non-government stakeholders in LMICs.

Like most screening technologies, the present method cannot provide definitive evidence of compliance or non-compliance of a medicine with pharmacopeial specifications. In practice it may best be employed together with other screening methods, such as visual inspection [68, 69] and the GPHF Minilab [9,11] as well as together with confirmatory analysis by compendial methods for samples which fail in the screening [60]. The speed of analysis by NIR spectroscopy may allow the screening of a much larger number of medicine samples compared to previously employed technologies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We are grateful to Benedikt Mannstadt and Katharina Beck for assistance in the recording of the spectra. We would also like to thank LuxFlux GmbH, Reutlingen, Germany, and Roland Winzen, nir-support, Bad Neuenahr-Ahrweiler, Germany, for advice and assistance in the initial steps of this study. The contribution by J.G. and G.G. were supported by PhD scholarships from the Dr. Hilmer Foundation, Deutsches Stiftungszentrum, Essen, Germany, and from the Cusanuswerk e.V., Bonn, Germany, respectively.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.talo.2023.100270.

References

[1] World Health Organization, A Study on the Public Health and Socioeconomic Impact of Substandard and Falsified Medical Products, World Health Organization, Geneva, Switzerland, 2017. https://www.who.int/publications-detail-redirect/9789241513432 (accessed 28 October 2023).
[2] World Health Organization, WHO Global Surveillance and Monitoring System for Substandard and Falsified Medical Products, World Health Organization, Geneva, Switzerland, 2017. https://apps.who.int/iris/handle/10665/326708 (accessed 28 October 2023).
[3] World Health Organization, Assessment of Medicines Regulatory Systems in Sub-Saharan African Countries. An Overview of Findings from 26 Assessment Reports, World Health Organization, Geneva, Switzerland, 2010. http://tropicaldoctor.alertvista.org/wp-content/uploads/2013/07/2010-WHO-Assessment26African_countries.pdf (accessed 28 October 2023).
[4] G.J. Buckley, L.O. Gostin (Eds.), Countering the Problem of Falsified and Substandard Drugs, National Academies Press, Washington, USA, 2013. https://www.ncbi.nlm.nih.gov/books/NBK202530/(accessed 28 October 2023).
[5] P. Nkansah, K. Smine, V. Pribluda, S. Phanouvong, C. Dunn, S. Walfish, F. Umaru, A. Clark, G. Kaddu, M. Hajjou, J. Nwokike, L. Evans, Guidance for Implementing Risk-Based Post-Marketing Quality Surveillance in Low- and Middle-Income Countries, U.S. Pharmacopeial Convention. The Promoting the Quality of Medicines Program, Rockville, USA, 2017. https://www.usp-pqm.org/sites/default/files/pqms/article/risk-based-post-marketing-surveillance-feb-2018.pdf (accessed 28 October 2023).
[6] C. Hu, Y. Feng, L. Yin, Review of the characteristics and prospects of near infrared spectroscopy for rapid drug-screening systems in China, J. Near Infrared Spectrosc. 23 (2015) 271–283, https://doi.org/10.1255/jnirs.1154.
[7] L. Höllein, E. Kaale, Y.H. Mwalwisi, M.H. Schulze, U. Holzgrabe, Routine quality control of medicines in developing countries: analytical challenges, regulatory infrastructures and the prevalence of counterfeit medicines in Tanzania, TrAC, Trends Anal. Chem. 76 (2016) 60–70, https://doi.org/10.1016/j.trac.2015.11.009.
[8] L. Roth, A. Nalim, B. Turesson, L. Krech, Global landscape assessment of screening technologies for medicine quality assurance: stakeholder perceptions and practices from ten countries, Global. Health 14 (2018) 43, https://doi.org/10.1186/s12992-018-0360-y.
[9] G. Gnegel, C. Häfele-Abah, R. Neci, Difäm-EPN Minilab Network, L. Heide, Surveillance for substandard and falsified medicines by local faith-based organizations in 13 low- and middle-income countries using the GPHF Minilab, Sci. Rep. 12 (2022) 13095, https://doi.org/10.1038/s41598-022-17123-0.
[10] S. Vickers, M. Bernier, S. Zambrzycki, F.M. Fernandez, P.N. Newton, C. Caillet, Field detection devices for screening the quality of medicines: a systematic review, BMJ Glob. Health 3 (2018), e000725, https://doi.org/10.1038/s41598-022-17123-0.
[11] U.S. Pharmacopeial Convention, USP Technology Review: Global Pharma Health Fund (GPHF) – Minilab™, Technology Review Program, Rockville, USA, 2020. https://www.usp.org/sites/default/files/usp/document/our-work/global-public-health/2020-usp-technology-review-global-pharma-health-fund-minilab.pdf (accessed 28 October 2023).
[12] L. Roth, K.B. Biggs, D.K. Bempong, Substandard and falsified medicine screening technologies, AAPS Open 5 (2019) 2, https://doi.org/10.1186/s41120-019-0031-y.
[13] S.C. Zambrzycki, C. Caillet, S. Vickers, M. Bouza, D.V. Donndelinger, L.C. Geben, M.C. Bernier, P.N. Newton, F.M. Fernandez, Laboratory evaluation of twelve portable devices for medicine quality screening, PLoS Negl. Trop. Dis. 15 (2021), e0009360, https://doi.org/10.1371/journal.pntd.0009360.
[14] R. Deidda, P.-Y. Sacré, M. Clavaud, L. Coïc, H. Avohou, P. Hubert, E. Ziemons, Vibrational spectroscopy in analysis of pharmaceuticals: critical review of innovative portable and handheld NIR and Raman spectrophotometers, TrAC, Trends Anal. Chem. 114 (2019) 251–259, https://doi.org/10.1016/j.trac.2019.02.035.
[15] T. Sanada, M. Ohnishi, N. Yoshida, K. Kimura, H. Tsuboi, Quality assessment of Diflucan(®) tablets distributed online, Med. Access Point Care 5 (2021), 23992026211002089, https://doi.org/10.1177/23992026211002089.
[16] T. Sanada, N. Yoshida, R. Matsushita, K. Kimura, H. Tsuboi, Falsified tadalafil tablets distributed in Japan via the internet, Forensic Sci. Int. 307 (2020), 110143, https://doi.org/10.1016/j.forsciint.2020.110143.
[17] O.Y. Rodionova, L.P. Houmøller, A.L. Pomerantsev, P. Geladi, J. Burger, V. L. Dorofeyev, A.P. Arzamastsev, NIR Spectrometry for counterfeit drug detection: a feasibility study, Anal. Chim. Acta 549 (2005) 151–158, https://doi.org/10.1016/j.aca.2005.06.018.
[18] S. Kovacs, S.E. Hawes, S.N. Maley, E. Mosites, L. Wong, A. Stergachis, Technologies for detecting falsified and substandard drugs in low and middle-income countries, PLoS ONE 9 (3) (2014) e90601, https://doi.org/10.1371/journal.pone.0090601.
[19] G. Reich, Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications, Adv. Drug Deliv. Rev. 57 (2005) 1109–1143, https://doi.org/10.1016/j.addr.2005.01.020.
[20] K.J. Siebert, Using chemometrics to classify samples and detect misrepresentation, Progress in Authentication of Food and Wine, ACS Sympos. Ser. 1081 (2011) 39–65, https://doi.org/10.1021/bk-2011-1081.ch004.
[21] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812–2831, https://doi.org/10.1039/C3AY41907J.
[22] W. Kessler, Multivariate Datenanalyse für die Pharma-, Bio- und Prozessanalytik, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2006.
[23] A.Y. Miró Vera, M. Alcalà Bernàrdez, Near-infrared Spectroscopy in Identification of Pharmaceutical Raw Materials, Encyclopedia of Analytical Chemistry - Pharmaceuticals and Drugs, 2017, https://doi.org/10.1002/9780470027318.a9619.
[24] 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 Rev2, 2014. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-use-near-infrared-spectroscopy-pharmaceutical-industry-data-requirements-new-submissions_en.pdf (accessed 28. October 2023).
[25] P.Y. Sacre, E. Deconinck, T. De Beer, P. Courselle, R. Vancauwenberghe, P. Chiap, J. Crommen, J.O. De Beer, Comparison and combination of spectroscopic techniques for the detection of counterfeit medicines, J. Pharm. Biomed. Anal. 53 (2010) 445–453, https://doi.org/10.1016/j.jpba.2010.05.012.
[26] I. Storme-Paris, H. Rebiere, M. Matoga, C. Civade, P.A. Bonnet, M.H. Tissier, P. Chaminade, Challenging near infrared spectroscopy discriminating ability for

counterfeit pharmaceuticals detection, Anal. Chim. Acta 658 (2010) 163–174, https://doi.org/10.1016/j.aca.2009.11.005.
[27] R. da Silva Fernandes, F.S. da Costa, P. Valderrama, P.H. Março, K.M. de Lima, Non-destructive detection of adulterated tablets of glibenclamide using NIR and solid-phase fluorescence spectroscopy and chemometric methods, J. Pharm. Biomed. Anal. 66 (2012) 85–90, https://doi.org/10.1016/j.jpba.2012.03.004.
[28] Y.V. Zontov, K.S. Balyklova, A.V. Titova, O.Y. Rodionova, A.L. Pomerantsev, Chemometric aided NIR portable instrument for rapid assessment of medicine quality, J. Pharm. Biomed. Anal. 131 (2016) 87–93, https://doi.org/10.1016/j.jpba.2016.08.008.
[29] N. Fuenffinger, S. Arzhantsev, C. Gryniewicz-Ruzicka, Classification of ciprofloxacin tablets using near-infrared spectroscopy and chemometric modeling, Appl. Spectrosc. 71 (2017) 1927–1937, https://doi.org/10.1177/0003702817699624.
[30] M. Eady, M. Payne, S. Sortijas, E. Bethea, D. Jenkins, A low-cost and portable near-infrared spectrometer using open-source multivariate data analysis software for rapid discriminatory quality assessment of medroxyprogesterone acetate injectables, Spectrochim. Acta, Part A 259 (2021), 119917, https://doi.org/10.1016/j.saa.2021.119917.
[31] M. Yabre, L. Ferey, A.K. Sakira, C. Bonmatin, C. Faure, T.I. Some, K. Gaudin, Green analytical methods of antimalarial artemether-lumefantrine analysis for falsification detection using a low-sost handheld NIR spectrometer with DD-SIMCA and drug quantification by HPLC, Molecules 25 (2020) 3397, https://doi.org/10.3390/molecules25153397.
[32] P.H. Ciza, P.-Y. Sacré, C. Waffo, L. Coic, H. Avohou, J. Mbinze, R. Ngono, R. D. Marini, P. Hubert, E. Ziemons, Comparing the qualitative performances of handheld NIR and Raman spectrophotometers for the detection of falsified pharmaceutical products, Talanta 202 (2019) 469–478, https://doi.org/10.1016/j.talanta.2019.04.049.
[33] O. Awotunde, N. Roseboom, J. Cai, K. Hayes, R. Rajane, R. Chen, A. Yusuf, M. Lieberman, Discrimination of substandard and falsified formulations from genuine pharmaceuticals using NIR spectra and machine learning, Anal. Chem. 94 (2022) 12586–12594, https://doi.org/10.1021/acs.analchem.2c00998.
[34] A. Petersen, N. Held, L. Heide, Difäm-EPN-Minilab Survey Group, surveillance for falsified and substandard medicines in Africa and Asia by local organizations using the low-cost GPHF Minilab, PLoS ONE 12 (2017), e0184165, https://doi.org/10.1371/journal.pone.0184165.
[35] S. Schäfermann, C. Hauk, E. Wemakor, R. Neci, G. Mutombo, E. Ngah Ndze, T. Cletus, F. Nyaah, M. Pattinora, D. Wistuba, I. Helmle, C. Häfele-Abah, H. Gross, L. Heide, Substandard and falsified antibiotics and medicines against noncommunicable diseases in western Cameroon and northeastern Democratic Republic of Congo, Am. J. Trop. Med. Hyg. 103 (2020) 894–908, https://doi.org/10.4269/ajtmh.20-0184.
[36] S. Ozawa, H.H. Chen, Y.A. Lee, C.R. Higgins, T.T. Yemeke, Characterizing medicine quality by active pharmaceutical ingredient levels: a systematic review and meta-analysis across low- and middle-income countries, Am. J. Trop. Med. Hyg. 106 (2022) 1778–1790, https://doi.org/10.4269/ajtmh.21-1123.
[37] A.L. Pomerantsev, O.Y. Rodionova, Concept and role of extreme objects in PCA/SIMCA, J. Chemom. 28 (2014) 429–438, https://doi.org/10.1002/cem.2506.
[38] Y.V. Zontov, O.Y. Rodionova, S.V. Kucheryayskiy, A.L. Pomerantsev, DD-SIMCA - A MATLAB GUI tool for data driven SIMCA approach, Chemom. Intell. Lab. Syst. 167 (2017) 23–28, https://doi.org/10.1016/j.chemolab.2017.05.010.
[39] H. Chen, Z. Lin, C. Tan, Application of near-infrared spectroscopy and class-modeling to antibiotic authentication, Anal. Biochem. 590 (2020), 113514, https://doi.org/10.1016/j.ab.2019.113514.
[40] Y. Tie, C. Duchateau, S. Van de Steene, C. Mees, K. De Braekeleer, T. De Beer, E. Adams, E. Deconinck, Spectroscopic techniques combined with chemometrics for fast on-site characterization of suspected illegal antimicrobials, Talanta 217 (2020), 121026, https://doi.org/10.1016/j.talanta.2020.121026.
[41] K. Degardin, A. Guillemain, N.V. Guerreiro, Y. Roggo, Near infrared spectroscopy for counterfeit detection using a large database of pharmaceutical tablets, J. Pharm. Biomed. Anal. 128 (2016) 89–97, https://doi.org/10.1016/j.jpba.2016.05.004.
[42] G. Gnegel, C. Hauk, R. Neci, G. Mutombo, F. Nyaah, D. Wistuba, C. Häfele-Abah, L. Heide, Identification of falsified chloroquine tablets in Africa at the time of the COVID-19 pandemic, Am. J. Trop. Med. Hyg. 103 (2020) 73–76, https://doi.org/10.4269/ajtmh.20-0363.
[43] P.H. Ciza, P.-Y. Sacre, M.R. Kanyonyo, C.T. Waffo, M.A. Borive, L. Coïc, J. K. Mbinze, P. Hubert, E. Ziemons, R.D. Marini, Application of NIR handheld transmission spectroscopy and chemometrics to assess the quality of locally produced antimalarial medicines in the Democratic Republic of Congo, Talanta Open 3 (2021), 100025, https://doi.org/10.1016/j.talo.2020.100025.
[44] W. Wang, M.D. Keller, T. Baughman, B.K. Wilson, Evaluating low-cost optical spectrometers for the detection of simulated substandard and falsified medicines, Appl. Spectrosc. 74 (2020) 323–333, https://doi.org/10.1177/0003702819877422.
[45] C. Caillet, S. Vickers, S. Zambrzycki, F.M. Fernandez, V. Vidhamaly, K. Boutsamay, P. Boupha, P. Peerawaranun, M. Mukaka, P.N. Newton, A comparative field evaluation of six medicine quality screening devices in Laos, PLoS Neglected Trop. Dis. 15 (2021), e0009674, https://doi.org/10.1371/journal.pntd.0009674.
[46] C.A. Waffo Tchounga, P.Y. Sacre, R. Ravinetto, M. Lieberman, P. Hamuli Ciza, R. Ngono Mballa, E. Ziemons, P. Hubert, R. Djang'eing'a Marini, Usefulness of medicine screening tools in the frame of pharmaceutical post-marketing surveillance, PLoS ONE 18 (2023), e0289865, https://doi.org/10.1371/journal.pone.0289865.
[47] World Health Organization, Model List of Essential Medicines - 22nd List, WHO/MHP/HPS/EML/2021.02, 2021. https://www.who.int/publications/i/item/WHO-MHP-HPS-EML-2021.02 (accessed 28. October 2023).
[48] R.A. Crocombe, Portable spectroscopy, Appl. Spectrosc. 72 (2018) 1701–1751, https://doi.org/10.1177/0003702818809719.
[49] InnoSpectra Corporation, InnoSpectra NIRScan. Spectrometer & Module, 1.2, User Manual. Ver, 2021, 2021/09/06, https://github.com/InnoSpectra/ISC-NIRScan-GUI (accessed 28 October 2023).
[50] M. Eady, M. Payne, C. Changpim, M. Jinnah, S. Sortijas, D. Jenkins, Establishment of instrument operation qualification and routine performance qualification procedures for handheld near-infrared spectrometers used at different locations within a laboratory network, Spectrochim. Acta, Part A 267 (2022), 120512, https://doi.org/10.1016/j.saa.2021.120512.
[51] S. Kucheryavskiy, O. Rodionova, A. Pomerantsev, Procrustes cross-validation of multivariate regression models, Anal. Chim. Acta 1255 (2023), 341096, https://doi.org/10.1016/j.aca.2023.341096.
[52] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, Validation of Analytical Procedures Q2(R2), Draft Version. ICH Harmonised Guidline, 2022, in: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-q2r2-validation-analytical-procedures-step-2b_en.pdf (accessed 28 October 2023).
[53] United States Pharmacopeia, USP general chapter <1850>evaluation of screening technologies for assessing medicines quality, 2020. https://www.usp.org/sites/default/files/usp/document/our-work/global-public-health/usp-general-chapter-1850-evaluation-of-screening-technologies-for-assessing-medicine-quality.pdf (accessed 28 October 2023).
[54] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639, https://doi.org/10.1021/ac60214a047.
[55] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (1989) 772–777, https://doi.org/10.1366/0003702894202201.
[56] A.L. Pomerantsev, O.Y. Rodionova, New trends in qualitative analysis: performance, optimization, and validation of multi-class and soft models, TrAC, Trends Anal. Chem 143 (2021), 116372, https://doi.org/10.1016/j.trac.2021.116372.
[57] O.Y. Rodionova, P. Oliveri, A.L. Pomerantsev, Rigorous and compliant approaches to one-class classification, Chemom. Intell. Lab. Syst. 159 (2016) 89–96, https://doi.org/10.1016/j.chemolab.2016.10.002.
[58] S. Kucheryavskiy, S. Zhilin, O. Rodionova, A. Pomerantsev, Procrustes cross-validation-A bridge between cross-validation and independent validation sets, Anal. Chem. 92 (2020) 11842–11850, https://doi.org/10.1021/acs.analchem.0c02175.
[59] A.L. Pomerantsev, O.Y. Rodionova, Procrustes cross-validation of short datasets in PCA context, Talanta 226 (2021), 122104, https://doi.org/10.1016/j.talanta.2021.122104.
[60] C. Hauk, N. Hagen, L. Heide, Identification of substandard and falsified medicines: influence of different tolerance limits and use of authenticity inquiries, Am. J. Trop. Med. Hyg. 104 (2021) 1936–1945, https://doi.org/10.4269/ajtmh.20-1612.
[61] T. Fearn, Standardisation and calibration transfer for near infrared instruments: a review, J. Near Infrared Spectrosc. 9 (2001) 229–244, https://doi.org/10.1255/jnirs.309.
[62] M. Bylesjo, M. Rantalainen, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg, OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification, J. Chemom. 20 (2006) 341–351, https://doi.org/10.1002/cem.1006.
[63] P.H. Ciza, P.Y. Sacre, C. Waffo, T.M. Kimbeni, B. Masereel, P. Hubert, E. Ziemons, R.D. Marini, Comparison of several strategies for the deployment of a multivariate regression model on several handheld NIR instruments. Application to the quality control of medicines, J. Pharm. Biomed. Anal. 215 (2022), 114755, https://doi.org/10.1016/j.jpba.2022.114755.
[64] E. Bouveresse, D. Massart, Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration, Chemom. Intell. Lab. Syst. 32 (1996) 201–213, https://doi.org/10.1016/0169-7439(95)00074-7.
[65] C. Mees, J.-M. Kauffmann, J.A.F. Pierna, K. De Braekeleer, Benchtop NIR data standardization on handheld spectrometers to identify paracetamol tablets, J. Chemom. 36 (2022) e3389, https://doi.org/10.1002/cem.3389.
[66] M. Yabre, A.K. Sakira, M. Bande, B.W.F. Goumbri, S.M. Ouattara, S. Fofana, T. I. Some, Detection of falsified antimalarial sulfadoxine-pyrimethamine and dihydroartemisinin-piperaquine drugs using a low-cost handheld near-infrared spectrometer, J. Anal. Methods Chem. 2022 (2022), 5335936, https://doi.org/10.1155/2022/5335936.
[67] O.Y. Rodionova, A.V. Titova, K.S. Balyklova, A.L. Pomerantsev, Detection of counterfeit and substandard tablets using non-invasive NIR and chemometrics - a conceptual framework for a big screening system, Talanta 205 (2019), 120150, https://doi.org/10.1016/j.talanta.2019.120150.

[68] B. Schiavetti, E. Wynendaele, V. Melotte, J. Van der Elst, B. De Spiegeleer, R. Ravinetto, A simplified checklist for the visual inspection of finished pharmaceutical products: a way to empower frontline health workers in the fight against poor-quality medicines, J. Pharm. Policy Pract. 13 (2020) 9, https://doi.org/10.1186/s40545-020-00211-9.

[69] World Health Professions Alliance. BeAware: tool for visual inspection of medicines, 2011, https://www.whpa.org/sites/default/files/2018-12/Toolkit_BeAware_Inspection.pdf (accessed 28 October 2023).

[70] World Health Organization, Medical Product Alert N° 6/2019: falsified hydrochlorothiazide (containing glibenclamide) in Cameroon, 2019. https://www.who.int/news/item/17-04-2019-medical-product-alert-n-6-2019-(english-version) (accessed 28 October 2023).