Department of Analytical Chemistry of Medicines[1], Showa Pharmaceutical University, Division of Formulation Science[2], Towa Pharmaceutical Co, Ltd., Division of Drugs[3], National Institute of Health Sciences, Tokyo, Japan

Strategic development of a multivariate calibration model for the uniformity testing of tablets by transmission NIR analysis

D. Sasakura[1], K. Nakayama[2], T. Sakamoto[3], T. Chikuma[1]

Received October 22, 2014, accepted December 16, 2014

Daisuke Sasakura, Department of Analytical Chemistry of Medicines, Showa Pharmaceutical University, 3-3165 Higashi-tamagawagakuen, Machida, Tokyo 194-8543, Japan
saru_dai@hotmail.com

Pharmazie 70: 289–295 (2015) *doi: 10.1691/ph.2015.4161*

The use of transmission near infrared spectroscopy (TNIRS) is of particular interest in the pharmaceutical industry. This is because TNIRS does not require sample preparation and can analyze several tens of tablet samples in an hour. It has the capability to measure all relevant information from a tablet, while still on the production line. However, TNIRS has a narrow spectrum range and overtone vibrations often overlap. To perform content uniformity testing in tablets by TNIRS, various properties in the tableting process need to be analyzed by a multivariate prediction model, such as a Partial Least Square Regression modeling. One issue is that typical approaches require several hundred reference samples to act as the basis of the method rather than a strategically designed method. This means that many batches are needed to prepare the reference samples; this requires time and is not cost effective. Our group investigated the concentration dependence of the calibration model with a strategic design. Consequently, we developed a more effective approach to the TNIRS calibration model than the existing methodology.

1. Introduction

Content uniformity testing in tablet manufacture is of critical importance. According to regulations, such as United States Pharmacopeia (UP), 10–30 tablets should be sampled from 1 to 3 million tablets and then tested by chromatographic methods with a statistical deviation approach. The analysis of 0.001% or less of the tablets manufactured is considered representative of the batch. High performance liquid chromatography (HPLC) and ultra high performance chromatography (UPLC) are the preferred methods of analysis. However, to analyze a greater number of tablets requires additional equipment, personnel, time, and cost. Near infrared spectroscopy (NIRS) is particularly attractive as alternative for the pharmaceutical industry as it does not require sample preparation and can easily analyze more than 30 tablets in an hour at the production line.

Transmission and Diffuse Reflection are the most common observation methods for NIRS. Transmission NIRS (TNIRS) is more sensitive and is easily repeatable as the components are observed in the tablet in this non destructive approach. The drawbacks of TNIRS include a narrow spectrum range and the overlapping of overtone vibrations in the regions of interest. Furthermore, physical information, such as the packing density, hardness (based on the characteristics of the particles of each component), and any variety in manufacture, has to be included in the analysis. To perform content uniformity testing in a tablet by TNIRS, this variety of information in the tableting process needs to be included by multivariate prediction models, such as a partial least square regression (PLSR) modeling. Therefore, as NIRS grows in popularity in the pharmaceutical industry, interest in the development of multivariate prediction models has also increased. Typically, the creation of a multivariate calibration model (e.g. PLSR) requires ongoing work and must be updated often to guarantee that a low prediction error is obtained and maintained for new batches. One issue in the application of TNIRS for content uniformity testing of tablets is that the multivariate model is developed empirically rather than strategically constructed. Furthermore, in a real tablet manufacturing process, the model should include the formulation, chemical and quality of the raw materials in the tablet. Thus, it should be possible to include in the calibration model any factors that influence the composition of the tablet by designing the experiment strategically. Our group investigated the development of a multivariate calibration model is the for the analysis of content uniformity in tablets by TNIRS with a strategic design approach.

One of the most onerous aspects in the development of a TNIRS calibration model preparation of reference samples. In particular, 20–30 points are required for the concentration range reference samples to give a model with sufficient accuracy. However, this means that more experiments are to be done than there would be in 10–20 batches just to prepare the reference samples. Decreasing the number of concentration reference samples would rapidly reduce the requirements for modeling. This can be achieved by using a more scientific methodology, which is less reliant on empirical measurements. This study reports on the number of points required in the concentration range to develop a calibration model using a more strategic approach.

2. Investigations and results

2.1. Candidate components for a calibration model based on a strategic design

Theophylline (TEO) as an active pharmaceutical ingredient (API); lactose, corn starch, micr cellulose (MCC) and magne-

Table 1: Target concentration and the range of concentrations for the calibration model of each component

Component	Target volume (per dose)	Concentration range
TEO	10 mg	5–15 mg
Lactose	104 mg	90–120 mg
Com starch	45 mg	35–55 mg
MCC	40 mg	3 0–50 mg
StMg	1 mg	0.5–1.5 mg

sium stearate (StMg) as excipients; were used as ocrystalline chemical components of a model tablet for this study. The target concentration and the range of concentration for calibration model of each component is shown in Table 1. The target concentration of the API was 10 mg/200 mg (5%/dose). Typically, a well-controlled practical tableting process may have slight deviations in the concentrations of each component, but not in the physical properties such as the tableting pressure, temperature, humidity and particle size. To develop a multivariate calibration model, the possible co-linearity of each component must be considered to reduce shifts in the prediction value based on component dependence in the calibration models. The use of a randomized approach in the calculation of candidate concentrations is important to ensure the non-collinearity of each component. Thirty candidate concentrations of each component were used as samples for this randomized method. To investigate the concentration level dependence of the candidate samples, they were classified as either Case 1 or Case 2. The number of samples was the same in both cases. The overall results of the candidate component, by case, are shown in Table 2. The non - linear relationship of each component was confirmed by a scatter plot with between each component, as shown in Fig. 1. Consequently, the candidate components used in this study had no co-linearity between API and any other component.

Table 2: a: Overall results of the candidate component in (a) Case 1

Sample	API mg/dose	Lactose mg/dose	Corn Starch mg/dose	MCC mg/dose	StMG mg/dose	# of samples
1	6.42	115.76	38.68	38.42	0.72	1
2	10.70	91.08	50.34	47.06	0.82	1
3	11.32	93.14	49.40	44.98	1.16	1
4	14.92	100.90	41.42	41.74	1.04	1
5	11.98	107.84	43.62	35.46	1.10	1
6	7.82	117.98	39.62	33.50	1.08	1
7	12.64	96.40	51.08	38.54	1.34	1
8	13.94	93.62	43.42	48.12	0.90	1
9	5.16	110.88	52.44	30.88	0.66	1
10	8.04	101.78	42.80	46.20	1.18	1
11	9.18	96.16	47.18	46.88	0.62	1
12	5.28	107.98	45.44	39.84	1.46	1
13	14.04	111.90	41.18	32.04	0.84	1
14	9.08	119.04	38.98	31.60	1.28	1
15	10.14	100.72	38.62	49.26	1.26	1
16	11.38	115.20	41.88	30.72	0.82	1
17	7.96	95.84	53.52	41.34	1.34	1
18	13.14	95.96	43.72	46.22	0.96	1
19	13.60	116.98	38.26	30.44	0.72	1
20	6.40	99.74	52.92	40.36	0.58	1
21	7.28	104.96	45.98	40.90	0.88	1
22	9.98	92.64	49.36	46.94	1.10	1
23	8.72	104.32	42.78	43.50	0.68	1
24	10.90	118.72	35.06	34.34	0.98	1
25	5.16	104.00	49.62	40.68	0.56	1
26	11.92	102.74	52.32	31.82	1.18	1
27	11.68	94.62	46.32	46.26	1.12	1
28	10.58	96.72	51.60	40.50	0.60	1
29	8.50	117.70	38.32	34.12	1.36	1
30	5.06	103.40	54.12	36.54	0.90	1
					Total	30

Table 2b: Overall results of the candidate component in (b) Case 2

Sample	API mg/dose	Lactose mg/dose	Com Starch ms/dose	MCC ma/dose	StMg mg/dose	# of samples
1	10.70	91.08	50.34	47.06	0.82	5
2	14.92	100.90	41.42	41.74	1.04	5
3	12.64	96.40	51.08	38.54	1.34	5
4	8.04	101.78	42.80	46.20	1.18	5
5	5.28	107.98	45.44	39.84	1.46	5
6	9.08	119.04	38.98	31.60	1.28	5
					Total	30

2.2. Spectrum band assignment of TEO in the NIR region

TNIRS can be used to observe the second and third overtones regions from 7000 to 11000 cm^{-1}. Searching the NIR spectrum of TEO found a stretching vibration of the terminal aromatic methyl at 8800 to 9200 cm^{-1}, which was characteristic of the target compound, as shown in Fig. 2.

2.3. Development of a PLSR calibration model based on a strategic design

The results of spectrum pretreatment on multiple samples for the creation of a calibration model by PLSR modeling, using HPLC method as the reference method, are presented in Table 3. This method of validating PLSR does not allow for cross validation. The correlation coefficients (R^2) and factor of the PLS model were 99.4 to 99.9 and 2 to 4, respectively. These results are robust and have a precision similar to those used in practice. The root mean square of error of cross validation (RMSECV) were 0.129 to 0.259 in Case 1 and 0.112 to 0.192 in Case 2. This means Case 1 gave a slightly lower accuracy model than Case 2. As mentioned above, these results have a low error when compared with the reference values and high accuracy of prediction over a range of concentrations, rather than the more normal, and less challenging, repetition at the same value. Furthermore, the first derivative preprocessing may have a high contribution to the error of prediction and the normalization preprocessing gave low correlation and high RMSECV. Therefore, the candidate preprocessing methods, based on RMSECV, were first derivative + SLS for Case 1 and first derivative + SNV for Case 2. The overall results of each calibration curve after cross validation are shown in Fig. 3.

2.4. Specificity of PLSR model

The comparison of the correlation coefficient vector profile (CCVP) from the PLSR and the spectra of each component with pretreatment, is shown in Fig. 4. Each analysis gave a similar profile with TEO by CCVP, which was different from every other component. Consequently, this result gave evidence that this calibration model is specific for TEO.

2.5. Comparison of performance of the PLSR calibration model to the number of points in the concentration level

To ensure the performance of each calibration model, each model had to be able to perform to external standard samples to

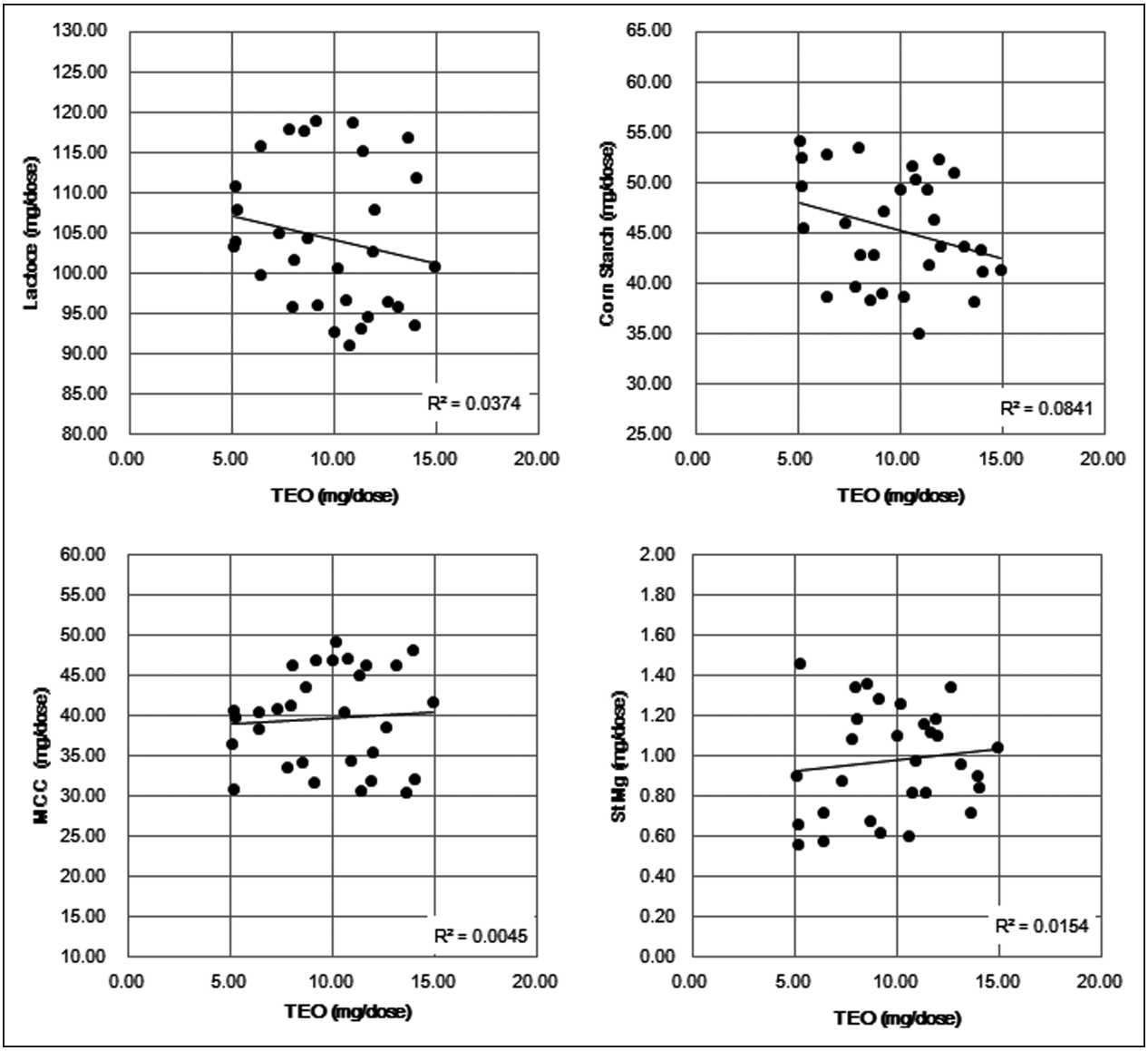

Fig. 1: a Definition of the candidate components for a calibration model by the randomized approach showing (a) as Case 1. Definition of the candidate components for a calibration model by the randomized approach showing (b) as Case 2.

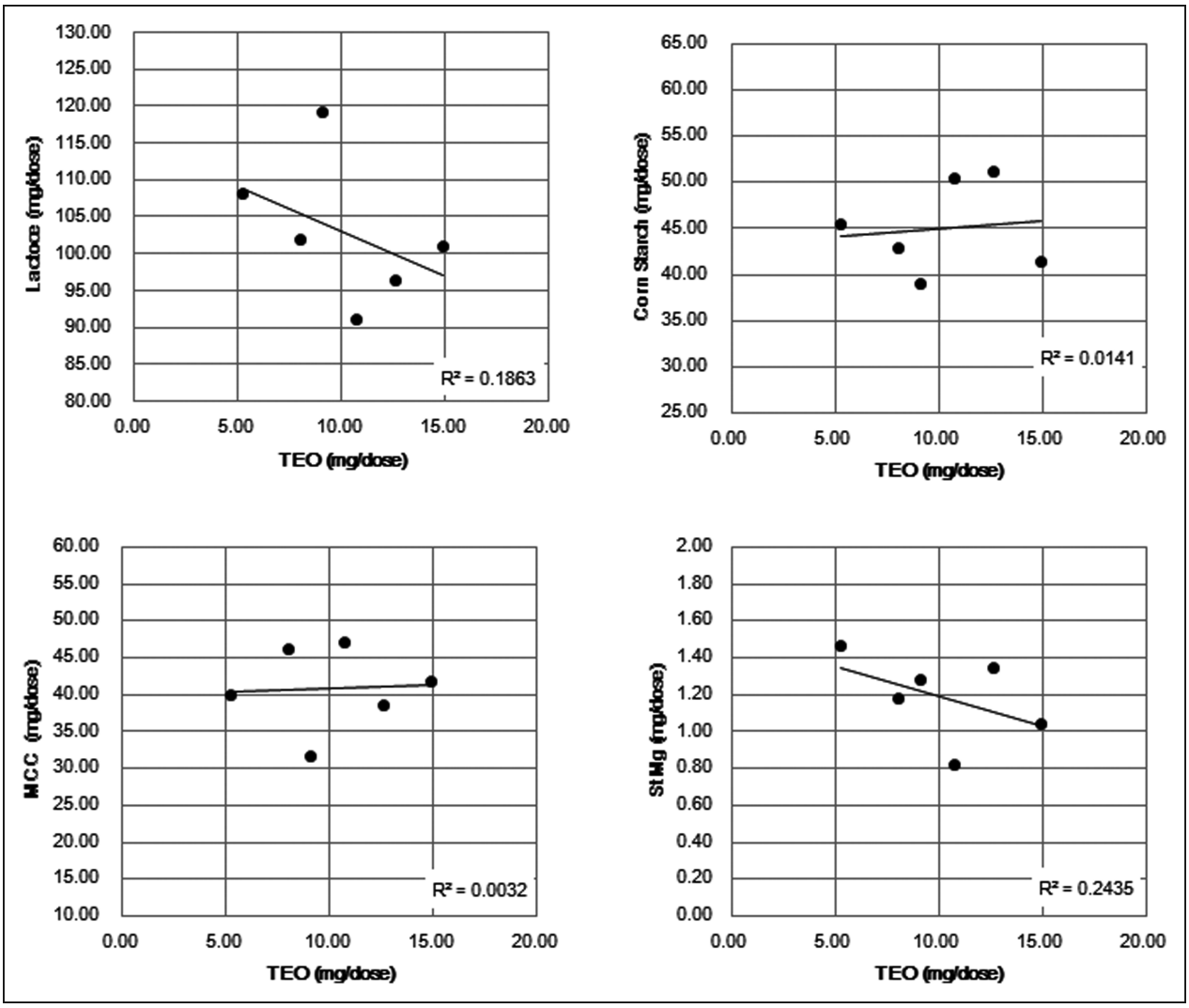

Fig. 1: *(Continued).*

within 10.00% of the concentration. The validation result of the performance of the PLS modeling by TNIRS compared with the HPLC result is shown in Table 4. In Case 1, the NIR prediction value was lower than the HPLC value, with an error range of 0 to 0.09. In Case 2, the NIR value was higher than the HPLC value, with an error range of − 0.09 to 0.01. The 95% confidence intervals of each NIR calibration model were 9.79 to 9.98 (average 9.89) and 9.93 to 10.12 (average 10.03), meaning both models had a similar prediction ability. The 95% confidence interval of the HPLC calibration model was 9.79 to 10.01 (average 9.93); hence, the Case 2 model overestimated the value.

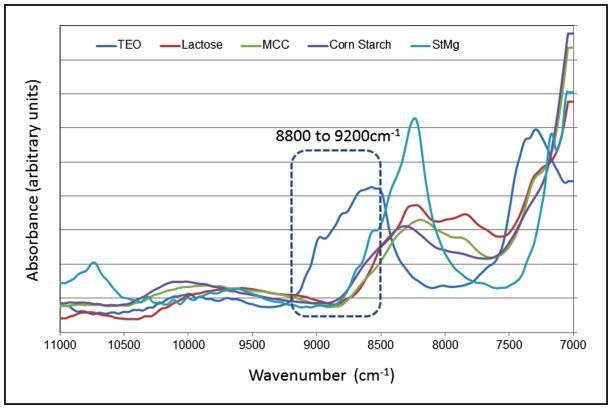

Fig. 2: Near infrared spectrum of each component of the raw material of the tablets.

3. Discussion

This study describes the strategic design and verification of a PLSR model for the content uniformity testing of tablets. Multivariate analysis has been used to create a high-performance calibration model by verified experimental design and the preprocessing of the spectra, using low numbers of concentration level samples. The number of points used for the concentration range of the PLSR modeling was limited with the use of a strategic design. Calibration models were developed that had the same total amount of samples, but the number of samples for the concentration range was either 30 or 5 points. Both cases gave similar performance as PLSR calibration models. Furthermore, this approach has shown that the possibility of capability to develop other components such as the excipients, another API and any chemical component in the tablet by TNIRS. In that case, it is necessary to carefully consider the concentration of an each target component in the tablet and the spectrum assignment of each chemical component.

In conclusion, this article presents an approach that is more rapid, more cost effective, and a more scientific way to develop a

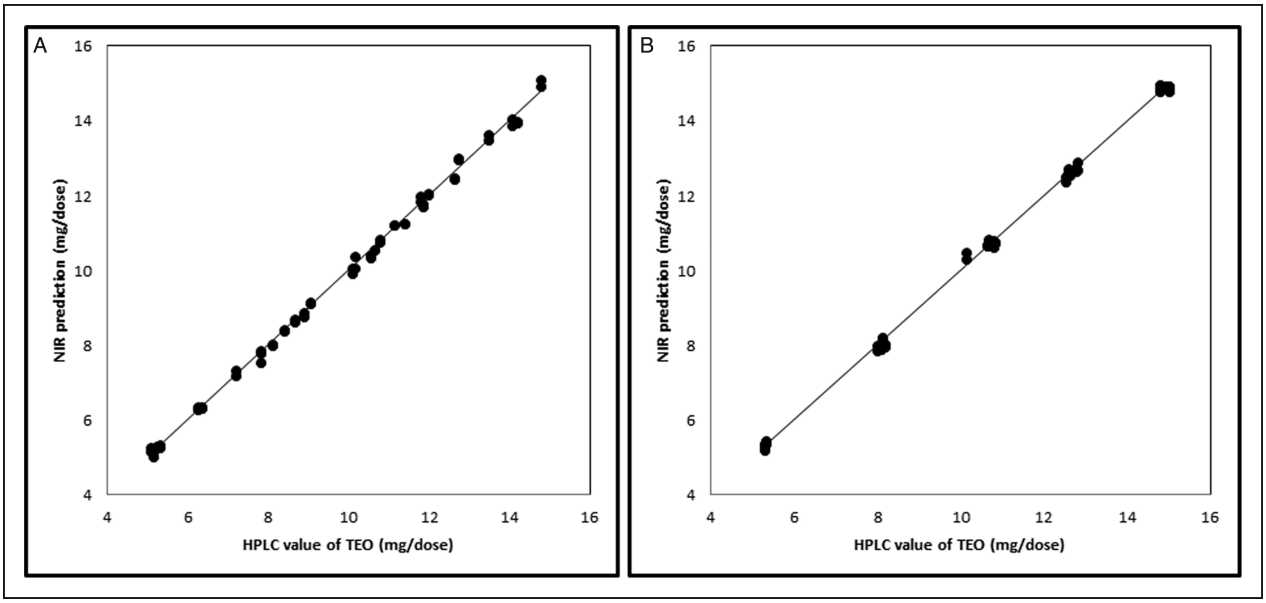

Fig. 3: a/b: Calibration model by cross validation; HPLC result was used as a reference showing (a) Case 1 and (b) Case 2.

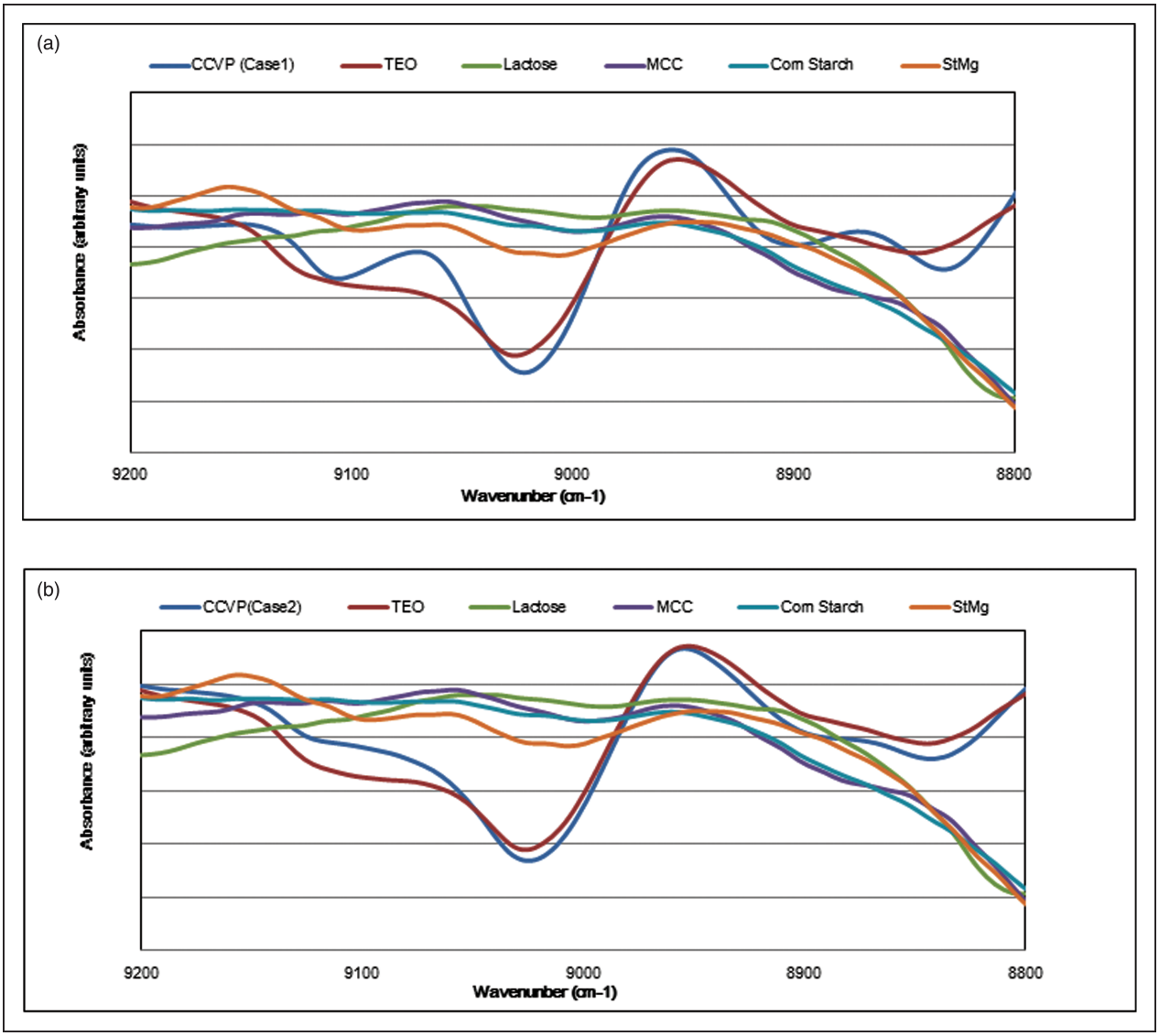

Fig. 4: a/b: Correlation coefficient vector profile (CCVP) of PLSR and the spectrums of each component showing (a) Case 1 and (b) Case 2.

Table 3: Standard errors and the number of factors for the calibration models evaluated with pre-treatment of spectr

Spectrum pre-treatment	RMSECV	Case 1 PLS factor	R^2	RMSECV	Case 2 PLS factor	R^2
First derivative	0.154	4	99.9	0.116	4	99.9
First derivative + SLS	0.19	3	99.8	0.112	2	99.9
First derivative + SNV	0.129	2	99.9	0.113	2	99.9
Min-max normalization	0.19	4	99.8	0.167	4	99.8
Offset correction	0.165	4	99.9	0.12	4	99.8
Second derivative	0.204	3	99.8	0.122	2	99.8
SLS	0.154	2	99.9	0.119	3	99.9
SNV	0.259	4	99.4	0.192	3	99.4

Table 4: a/b: Standard errors and number of factors for the calibration models' evaluation of the test samples showing (a) Case 1 and (b) Case 2

Sample	HPLC(mg/dose)	NIR prediction (mg/dose)	Error (mg/dose)
1	9.99	9.94	0.05
2	9.95	9.86	0.09
3	9.78	9.71	0.07
4	9.95	9.93	0.02
5	9.99	9.99	0
Mean			0.046
SDV			0.04

Table 4(b)

Sample	HPLC	NIR prediction	Error
	(mg/dose)	(mg/dose)	(mg/dose)
1	9.99	10.08	−0.09
2	9.95	9.94	0.01
3	9.78	9.81	−0.03
4	9.95	10.04	−0.09
5	9.99	10.01	−0.02
Mean			−0.044
SDV			0.07

multivariate calibration model for the content uniformity testing of tablets by TNIRS.

4. Experimental

4.1. Acquisition of transmission near infrared spectra for the PLSR calibration model

Transmission spectra were recorded in a multipurpose analyzer (MPA) A Fourier transform near infrared (FT-NIR) spectrometer (Bruker Optick GmbH, Germany) equipped with a 30 position sample wheel and a room temperature–indium gallium arsenide (RT-InGaAs) external detector was positioned above the tablet. The spectra were collected with Opus 6.0 software (Bruker Optick GmbH). Each spectrum was an average of 256 scans at a resolution of 8 cm^{-1}, over the range of 11000 to 7500 cm^{-1}. Both directions of tablet movement were collected to compensate for the effects of any directionality in the samples.

4.2. Acquisition of near infrared spectra for the assignment of TEO spectrum

The spectra for each component were recorded in a MPA FT-NIR spectrometer (Bruker Optick GmbH) equipped with an integration sphere and a lead sulfide (PbS) detector. Each spectrum was collected with Opus 6.0 software (Bruker Optick GmbH), in a 22 mm glass vial from the bottom of the vial. Each spectrum was an average of 64 scans at a resolution of 8 cm^{-1}, over the range of 11000 to 7500 cm^{-1}.

4.3. Quantitative analysis of TEO by HPLC

All measurements for the quantitative analysis of the reference data of the calibration model were done by the HPLC facility in the laboratory of Towa Pharmaceutical Company. The test protocol was based on the regulation requirements of the Japanese Pharmacopeia.

4.4. Tablet manufacturing

All samples were produced by the laboratory of Towa Pharmaceutical Company. Each concentration for direct compression was manufactured with the appropriate candidate concentrations of each component. Tablets were compresed at a maximum load of 80 kN (0.8 t). Tablets obtained were round at 200 mg/dose, with an average diameter surface of 8 mm, a thickness of approximately 3.0 mm, and with a secant on one side.

4.5. Development and design of a PLSR coefficient calibration model

The recorded TNIR spectra, together with the results from the HPLC reference analysis, were analyzed using PLSR. PLS calibration development and cross validation were performed with the QUANT 2 module of the Opus 6.0 software (Bruker Optick GmbH), for all spectra pretreatments tested. The derivative method used was the Savitky-Golay algorithm with 25 point smoothing. The calibration designing function module of Opus 6.0 software (Bruker Optick GmbH) was used to calculate the concentration of each component for each calibration model.

References

Blanco M, Coello J, Eustaquio A, Iturriaga H, Maspoch S (1999) Development and validation of methods for the determination of miokamycin in various pharmaceutical preparations by use of near infrared spectroscopy. Analyst 124: 1089–1092.

Blanco M, Coello J, Iturriaga H, Maspoch S, Pezuela C (1997) Effect of data preprocessing methods in near-infrared diffuse reflectance spectroscopy for the determination of the active compound in a pharmaceutical component. Appl Spectrosc 51: 240–246.

Blanco M, Colleo J, Iturriiaga H, Maspoch, and Pou N (2000) Development and validation of near infrared method for the analytical control of a pharmaceutical preparation in tree steps of manufacturing process. Fres J Anal Chem 368: 534–539.

Broad N, Jee R, Moffat C, Smith M (2001) Application of transmission near infrared spectroscopy to uniformity of content testing of intact steroid tablets. Analyst 126: 2207–2211.

Carlos PM, María AS, Rodolfo JR (2006) Quantitation of drug content in a low dosage formulation by transmission near infrared spectroscopy. AAPS Pharm Sci Tech 7: E206-E214.

Gottfries J, Depui H, Fransson M (1996) Vibrational spectrometry for the assessment of active substance in metoprololtablets: a comparison between transmission and diffuse reflectance near infrared spectrometry. J Pharm Biomed Anal 14: 1495–1503.

Kemper MS, Magnuson EJ, Lowry SR. (2002) Use of FT-NIR transmission spectroscopy for the quantitative analysis of an active ingredient in a

translucent pharmaceutical topical gel formulation. AAPS Pharm Sci 3: E23.

Romanach RJ, Santos MA (2003) Content uniformity testing with near infrared spectroscopy. Am Pharm Anal 11: 1227–1231.

Savitky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 36: 1627–1639.

Sigrid P, Wouter S, Tom V K, Mohammad G, Mario H, Thomas D, Yvan V H (2013) Robust calibrations on reduced sample sets for API content prediction in tablets: Definition of a cost-effective NIR model development strategy. Anal Chim Acta 761: 62–70.

United States Pharmacopeial Convention (2005) USP 28-NF19. US General Chapter 905, Uniformity of Dosage Forms. United States Pharmacopeial Convention, Inc.