Micro-Electro-Mechanical Systems/Near-Infrared Validation of Different Sampling Modes and Sample Sets Coupled with Multiple Models

ARTICLE *in* PLANTA MEDICA · JANUARY 2015

Impact Factor: 2.15 · DOI: 10.1055/s-0034-1396152 · Source: PubMed

6 AUTHORS, INCLUDING:

Zhisheng Wu

Beijing University of Chinese Medicine and …

50 PUBLICATIONS **108** CITATIONS

SEE PROFILE

Xinyuan Shi

Beijing University of Chinese Medicine and …

54 PUBLICATIONS **108** CITATIONS

SEE PROFILE

Manfei xu

Beijing University of Chinese Medicine and …

6 PUBLICATIONS **1** CITATION

SEE PROFILE

Yanjiang Qiao

Beijing University of Chinese Medicine and …

82 PUBLICATIONS **150** CITATIONS

SEE PROFILE

Micro-Electro-Mechanical Systems/Near-Infrared Validation of Different Sampling Modes and Sample Sets Coupled with Multiple Models

Authors Zhisheng Wu[1,2,3], Xinyuan Shi[1,2,3], Guang Wan[1], Manfei Xu[1], Xueyan Zhan[1], Yanjiang Qiao[1,2,3]

Affiliations
[1] Beijing University of Chinese Medicine, Beijing, China
[2] Key Laboratory of TCM-information Engineering of State Administration of TCM, Beijing, China
[3] Beijing Key Laboratory for Basic and Development Research on Chinese Medicine, Beijing, China

Key words
- near-infrared
- accuracy profile
- integrating sphere
- process analytical technology
- fiber optic sensor

received June 26, 2014
revised Nov. 14, 2014
accepted Nov. 23, 2014

Bibliography
DOI http://dx.doi.org/10.1055/s-0034-1396152
Planta Med 2015; 81: 167–174

ISSN 0032-0943

Correspondence
Yanjiang Qiao
School of Chinese Pharmacy
Beijing University of Chinese Medicine
South of Wangjing Middle Ring Road, Chaoyang District
Beijing City 100102
China
Phone: + 86 10 84 73 86 21
Fax: + 86 10 84 73 86 61
yjqiao@263.net

Correspondence
Xinyuan Shi
School of Chinese Pharmacy
Beijing University of Chinese Medicine
South of Wangjing Middle Ring Road, Chaoyang District
Beijing City 100102
China
Phone: + 86 10 84 73 86 21
Fax: + 86 10 84 73 86 61
shixinyuan01@163.com

Abstract

The aim of the present study was to demonstrate the reliability of micro-electro-mechanical systems/near-infrared technology by investigating analytical models of two modes of sampling (integrating sphere and fiber optic probe modes) and different sample sets. Baicalin in Yinhuang tablets was used as an example, and the experimental procedure included the optimization of spectral pretreatments, selection of wavelength regions using interval partial least squares, moving window partial least squares, and validation of the method using an accuracy profile. The results demonstrated that models that use the integrating sphere mode are better than those that use fiber optic probe modes. Spectra that use fiber optic probe modes tend to be more susceptible to interference information because the intensity of the incident light on a fiber optic probe mode is significantly weaker than that on an integrating sphere mode. According to the test set validation result of the method parameters, such as accuracy, precision, risk, and linearity, the selection of variables was found to make no significant difference to the performance of the full spectral model. The performance of the models whose sample sets ranged widely in concentration (i.e., 1–4%) was found to be better than that of models whose samples had relatively narrow ranges (i.e., 1–2%). The establishment and validation of this method can be used to clarify the analytical guideline in Chinese herbal medicine about two sampling modes and different sample sets in the micro-electro-mechanical systems/near-infrared technique.

Supporting information available online at http://www.thieme-connect.de/products

Introduction

So far, a variety of wet chemistry methods, including HPLC and MS, have been discussed. These methods are subject to many disadvantages, such as time, cost, and labor-intensive sample preparation and the use of an organic solvent [1]. Vibrational spectroscopy techniques are faster and have nondestructive advantages. Near-infrared (NIR) is a vibrational spectroscopy technique due to molecular vibrations of hydrogen bonds, like C–H, N-H, O-H, and S-H. The vibrational energy of hydrogen bonds comes from overtones and combinations of the fundamental mid-infrared bands [2–4]. Due to the principle of overtones and combinations, it can gather information rapidly with little or no sample preparation. It has been widely used in the quantitative and qualitative analysis of Chinese herbal medicines (CHM) [5–9]. NIR has been widely regarded as an excellent online/in-line tool for process monitoring in CHM [10].

Micro-electro-mechanical systems (MEMS)-NIR are a new generation of optical spectrometers, which adopt a novel approach to improving the precision and accuracy of a measurement [11]. Two sets of sampling technology, a fiber optic probe and an integrating sphere, are widely used in diffuse reflectance mode by MEMS-NIR systems [12–14]. The reflectance integrating sphere is designed for the collection of optical radiation, which has been reflected from a sample for the purpose of measuring reflectance. It is quick and convenient and has been used with NIR reflectance spectroscopy to assess total quantities of polysaccharides and triterpenoids in *Ganoderma lucidum* and *Ganoderma atrum* from different origins [15]. Quantitative analysis of free amino acid content was demonstrated in Radix Pseudostellariae [16]. The NIR data on Rhizoma et Radix Ba-

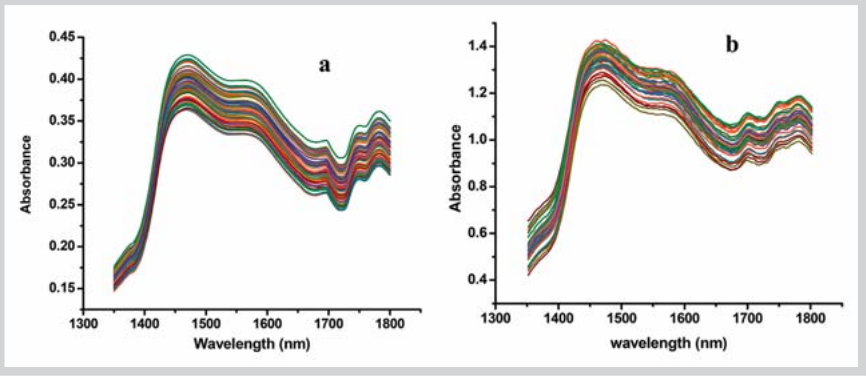

Fig. 1 Raw near-infrared spectra using integrating sphere and fiber optic probe modes. **a** Integrating sphere mode; **b** fiber optic probe mode. (Color figure available online only.)

phicacanthis Cusia were used to evaluate the quantitative composition [17]. A method of detecting total flavonoid content was established in fresh *Ginkgo biloba* leaves using NIR technology [18].

In a fiber optic probe, the light travels from the spectrometer to the optical probe and then back to the spectrometer for detection. The advantages of fiber optic probes are the modularity of the system. The measurement speed allows for online analysis, and the use of common detectors reduces costs. Some studies of the diffuse reflectance mode of fiber optic probes have been published. In one, the concentration of tagitinin C in *Tithonia diversifolia* leaves was determined using a fiber optic probe coupled with Fourier transform infrared spectroscopy [19]. An online NIR model of chlorogenic acid was produced in the extraction process of *Lonicera japonica* [9].

In all cases of NIR analysis, algorithms are required; these include multiple linear regression (MLR), partial least squares regression (PLS), interval partial least-squares (iPLS), moving window partial least squares (mwPLS), and support vector regression (SVR). Model performance was assessed according to the following criteria: low standard error of calibration (SEC), low root mean standard error of cross validation (RMSECV), low root mean square errors of prediction (RMSEP), high coefficient of determination (R^2), and low root mean squared error (RMSE) [4,20].

Once a model of calibration is developed, it must be validated before it can be accepted for routine use. For method validation, there are several methodology parameters that should be determined in order to keep practices consistent with the recommendations of the International Conference on Harmonization (ICH) and other guidelines: accuracy, precision (repeatability and intermediate precision), specificity, linearity, and range of application [21].

Based on β-expectation tolerance intervals, the accuracy profile provides the NIR model with a reliable validation and actual future performance. It fully complies with the ICH Q2(R1) regulatory documents (Validation of analytical procedures: text and methodology) in that it integrates all the required methodology parameters, i.e., accuracy, precision, lower limit of quantification (LLOQ), upper limit of quantization (ULOQ), range, and linearity [22–24]. Detailed information about the accuracy profile theory was discussed in our previous papers [5,7–9,21].

However, there is no previous report on the development and validation of the NIR method using two sampling modes in CHM with an accuracy profile based on the different samples sets. There has never been any case of model development or validation of MEMS-NIR. The formulation of the Yinhuang tablet (0.3 g per tablet) includes *Lonicera japonica* extract (0.1 g per tablet), *Scutellaria baicalensis georgi* extract (0.04 g per tablet), and starch (0.16 g per tablet) [25].

Table 1 Experimental design based on the different contents of the sample sets.

Acquisition mode	Different contents of sample set	Sample number
Integrating sphere	1% ~ 4.5%	65
Integrating sphere	2% ~ 4.5%	57
Integrating sphere	3% ~ 4.5%	34
Integrating sphere	1% ~ 2%	31
Fiber optic	1% ~ 4.5%	65
Fiber optic	2% ~ 4.5%	57
Fiber optic	3% ~ 4.5%	34
Fiber optic	1% ~ 2%	31

In the present work, Yinhuang tablets were taken as an example, and the NIR quantitative model of baicalin was investigated based on different calibration samples using fiber optic probes and an integrating sphere. The accuracy profile approach based on test sets was followed to validate the quantitative results of the calibration model using a fiber optic probe and an integrating sphere.

Results and Discussion

The NIR spectra of Yinhuang tablets were collected using two sampling modes. There were less burrs in the raw spectra (**Fig. 1**). The phenomenon was due to the lack of interference of size nonuniformity and tightness inconformity in the tablets. The results demonstrated that NIR spectra using the integrating sphere mode were significantly better than those using the fiber optic probe mode.

The NIR spectra that used the fiber optic probe mode presented more burrs, which showed that the fiber optic probe mode involved more interference. Raw spectra performed better than derivative spectra. The preprocessing spectra of the first derivative (1 d) were better than that of the second derivative (2 d) spectra (**Figs. 1S** and **2S**, Supporting Information).

PLS models based on different calibration samples sets were established (**Table 1**). As shown in **Table 2**, the standard normal variate (SNV) was the best preprocessing method for 1–4.5% of the calibration sample sets whose data was from the integrating sphere or fiber optic probe modes. This model also turned out to have the lowest RMSE, RMSEP, and bias values.

Table 2 Predicting results of different preprocessing methods using 1–4.5% of each sample set (unit, %).

Pretreatments	Latent factors		Calibration set				Validation set				Prediction set					
			R^2		RMSE		R^2		RMSE		R^2		RMSEP		Bias	
	I	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
Raw	3	7	0.9561	0.9804	0.0019	0.0013	0.9524	0.9713	0.0020	0.0015	0.9289	0.8009	0.0024	0.0040	0.0004	0.0005
1D	5	5	0.9945	0.9922	0.0007	0.0008	0.9914	0.9848	0.0009	0.0011	0.9238	0.6125	0.0025	0.0056	0.0002	0.0017
2D	3	6	0.9828	0.9843	0.0012	0.0011	0.9750	0.9630	0.0014	0.0018	0.7842	NOD	0.0042	0.0099	0.0001	0.0014
Baseline	3	6	0.9571	0.9843	0.0019	0.0011	0.9542	0.9782	0.0020	0.0014	0.9419	0.8022	0.0022	0.0040	0.0002	0.0002
SG	5	5	0.9939	0.9906	0.0007	0.0009	0.9901	0.9865	0.0009	0.0011	0.9178	0.6418	0.0026	0.0054	0.0002	0.0013
SNV	2	5	0.9635	0.9760	0.0017	0.0014	0.9620	0.9676	0.0018	0.0017	0.9503	0.8336	0.0020	0.0037	0.00003	0.0002
SNV + 1D	4	4	0.9892	0.9930	0.0009	0.0008	0.9849	0.9874	0.0011	0.0010	0.9252	0.7214	0.0025	0.0048	−0.001	0.0008

*Integrating sphere mode (I) and fiber optic mode (F)

The PLS model was performed using a 1–2% calibration sample set and integrating sphere mode for the SNV preprocessing method (**Table 3S**, Supporting Information). As for the calibration sample set from 3–4.5%, raw pretreatment was the optimal preprocessing method (**Table 5S**, Supporting Information). It is noteworthy that the PLS model based on the 1–2% and 3–4.5% calibration sample sets using the fiber optic probe mode could not satisfy the analytical requirements no matter what preprocessing method was used (**Tables 4S** and **6S**, Supporting Information). As shown in **Table 2**, the integrating sphere mode showed better prediction results than the fiber optic probe mode. The results of the other calibration sets were consistent (see Supporting Information).

Different methods of selecting wavelengths, iPLS and mwPLS, were investigated. The principle of the iPLS algorithm is to split the spectra into smaller equidistant regions and then to develop PLS regression models for each of the subintervals using the same number of latent factors. Thereafter, the region with the lowest error is chosen. mwPLS is a variable selection technique based on all continuous windows size "H" in the spectral data set. The RMSECV value of cross-validation was further used to find the best spectral region(s) of size H. Detail descriptions about iPLS and mwPLS theory and algorithms were reported in our previous references [8, 9].

iPLS model performance did not perform significantly better than the full-spectrum PLS model in either of the two sampling modes (results not shown). Characteristic multi-regions for calibration models were selected by the mwPLS method. As shown in **Fig. 2**, the mwPLS method was established using a size 9 window because of the minimum RMSECV value in these characteristic wavelength multi-regions.

The results produced using all of the calibration sample sets and the integrating sphere mode were collected. They demonstrated that characteristic wavelength multi-regions could be selected by calibration sample sets with a wide concentration range (i.e., 1–4% of the sample set and 2–4% of the sample set), as shown in **Fig. 2**. Characteristic wavelength multi-regions could not be used in the calibration sample sets with narrow concentration ranges (i.e., 1–2% of the sample set and 3–4% of the sample set).

The results also showed that characteristic wavebands were not selected in all calibration sample sets using the fiber optic probe mode. The phenomenon was due to the characteristics of the fiber optic probe mode. The intensity of the incident light in the fiber optic probe mode was significantly weaker than that in the integrating sphere mode. Each wavelength region of diffuse reflectance spectra tended to be more susceptible to interference (i.e., noise) when using the fiber optic probe mode, and it may be difficult to extract characteristic bands using the variable selection method.

Finally, calibration sample sets with relatively narrow concentration ranges and two sampling modes were also more susceptible to interference than those with wide ranges and one sampling mode, and the characteristic bands were more difficult to extract. Although iPLS and mwPLS were applied to wavelength selection, the results demonstrated that model performance using sample sets with a relatively narrow concentration range was still not able to satisfy the analytical requirements.

An accuracy profile was used to validate the NIR method in order to evaluate two sampling modes based on different calibration samples sets. Method parameters were compared to each other using a sample set from 1–4.5% as an example.

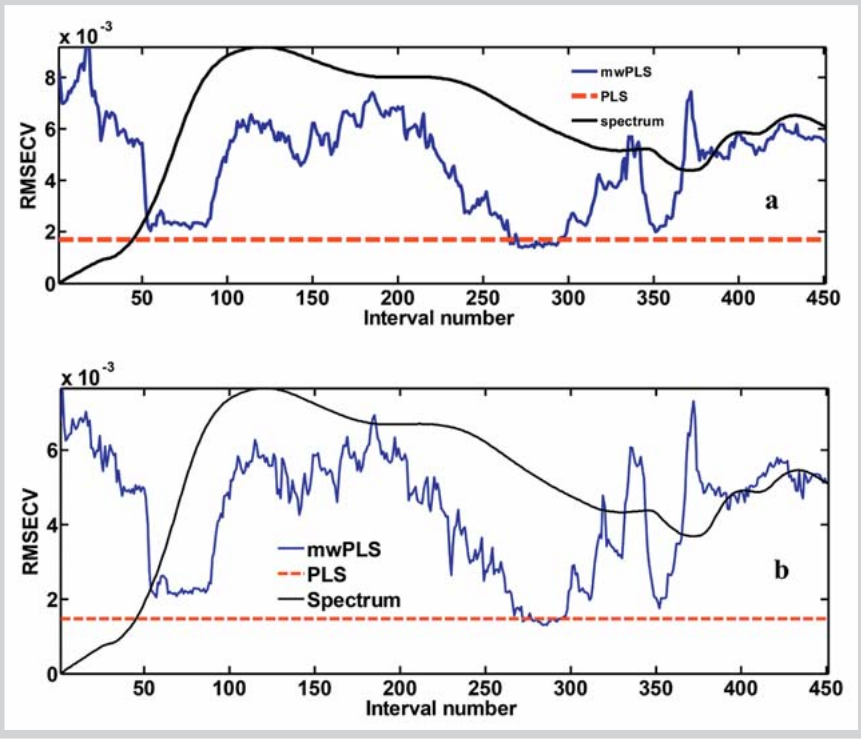

Fig. 2 Root mean standard error of the cross validation value of the moving window partial least squares method by optimal window size 9 (**a** %–4.5% of the sample set using the integrating sphere mode; **b** 2–4.5% of the sample set using the integrating sphere mode). (Color figure available online only.)

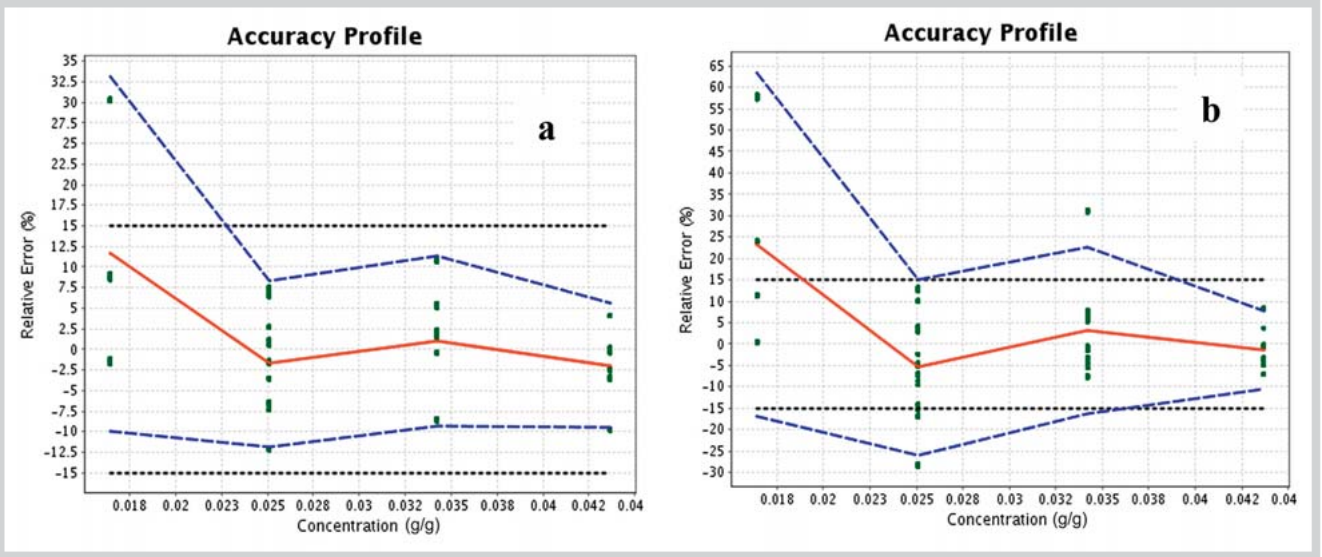

Fig. 3 Accuracy profile figure (**a** integrating sphere mode; **b** fiber optic probe mode). Black line: accepted limits; blue line: 90% β-expecting tolerance interval; red line: relative bias. (Color figure available online only.)

Fig. 3 depicts the accuracy profile results obtained with two sampling modes. The acceptance limits were set to ± 15%, and the maximum risk to obtain results outside these acceptance limits was set to 10%. As shown in **Fig. 3**, the relative prediction error using the integrating sphere mode with the lowest concentration level (0.015 g/g) was beyond the accepted limit, 15%, but that of other concentrations was within the accepted limits.

However, 90% β-expectation tolerance limits were included within the ± 15% accepted limit for only the highest concentration (0.04 g/g) using the fiber optic probe mode. The relative prediction error in other concentrations was outside of the accepted limit, 15%. In this way, it demonstrated that NIR quantitative results collected using the integrating sphere mode were distinctly better than those produced using fiber optic probe modes.

Table 3 shows the method parameters of the NIR model using two different sampling modes, including accuracy, precision, and uncertainty. However, the method parameters with the concentration level of 0.018 g/g did not satisfy the analytical requirements. Analytical requirements were satisfied with three other levels of concentration, 0.025 g/g, 0.035 g/g, and 0.043 g/g. The risk was below 2.1% for all three concentration levels, and repeatability and intermediate precision were below 6.05%.

With the fiber optic probe mode, repeatability and intermediate precision with the highest concentration level (0.043 g/g) were low, and the risk was 69.1%. The results with other concentration

Table 3 Method parameters of the near-infrared model based on ICH Q2(R1) using two sampling modes.

Mode	LEV	Accuracy			Precision		Uncertainty	
		REB	RTL	RIS	REP	INP	UB	REU
JFQ	0.018	11.59	[−10, 33.2]	41.66	11.94	11.94	0.0005	24.60
	0.025	−1.75	[−11.8, 8.3]	2.013	5.93	5.93	0.0002	11.99
	0.035	1.04	[−9.3, 11.4]	2.106	6.05	6.05	0.0003	12.26
	0.043	−1.95	[−9.5, 5.6]	0.403	4.33	4.33	0.0004	8.84
GX	0.018	23.36	[−16.78, 63.50]	69.73	22.19	22.19	0.0009	45.75
	0.025	−5.50	[−26.12, 15.13]	27.24	12.13	12.13	0.0005	24.53
	0.035	3.09	[−16.33, 22.51]	21.58	11.33	11.33	0.0006	22.98
	0.043	−1.45	[−10.49, 7.59]	1.11	5.17	5.17	0.0005	10.55

* JFQ: integrating sphere mode; GX: fiber optic probe mode; LEV: level (g/g); REB: relative bias (%); RTL: relative expectation tolerance limits (%); RIS: risk (%); REP: repeatability (RSD %); INP: intermediate precision (RSD %); UB: uncertainty of the bias (g/g); REU: relative expanded uncertainty (%)

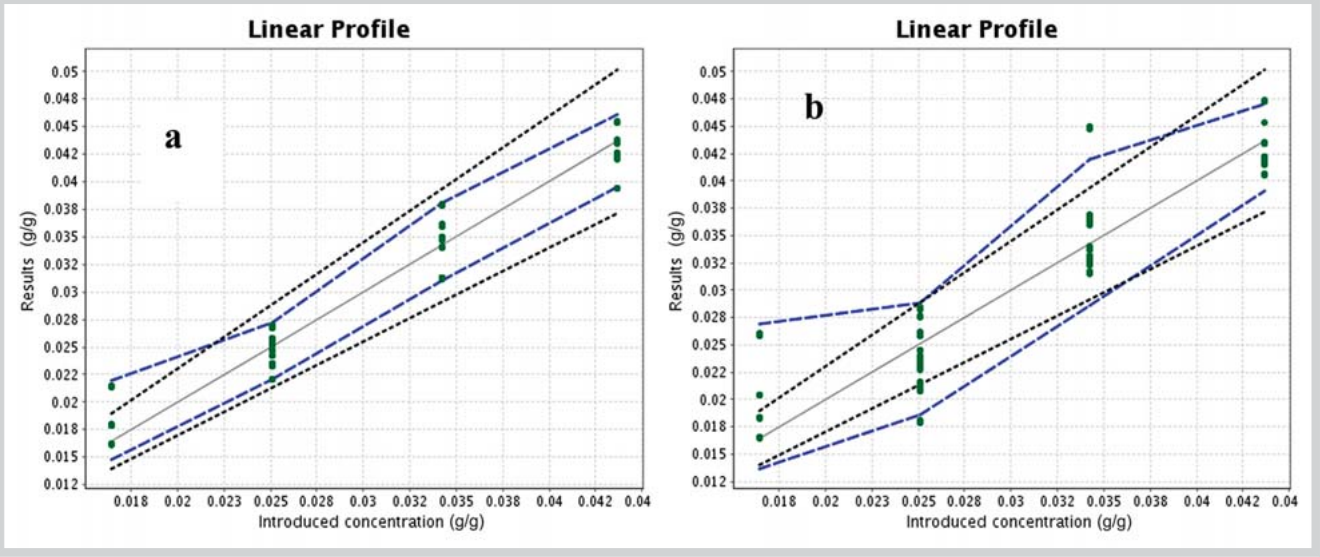

Fig. 4 Near-infrared predicted linear profile (**a** integrating sphere mode; **b** fiber optic probe mode). Black line: accepted limits; blue line: 90% β-expecting tolerance interval. (Color figure available online only.)

levels were beyond the analytical requirements. The results further showed that full spectral PLS models established in the integrating sphere mode are better than those established in the fiber optic probe mode.

Fig. 4 shows the linear results of the full spectral PLS model using two sampling modes. In the integrating sphere mode, the regression equation was $Y = 0.001\,850 + 0.9401X$, $R^2 = 0.9480$. In the fiber optic probe mode, the regression equation was $Y = 0.002173 + 0.9350X$, $R^2 = 0.8354$. When the intercept is 0 and the slope is 1, the total prediction error of the model is 0. The results show that the prediction total error of the full spectral PLS model in the integrating sphere mode is less pronounced than in the fiber optic probe mode.

Accuracy profile figures are displayed in **Fig. 5**. These data were collected using the integrating sphere mode with two different sample sets. The relative prediction error of the mwPLS model with sample sets from 2–4.5% remained within the accepted limits of 15%, which turned out to be similar to those of the full spectral PLS model. The method parameters of the mwPLS model are shown in **Table 4**. These included accuracy, precision, and uncertainty. The risk with three concentration levels was below 2.0%, and intermediate precision and repeatability were below 6.83%. Chemometric indicators, RMSE, RMSEP, bias, and other matters clearly showed the superiority of mwPLS over full spectral PLS as a method of selecting wavelength regions for calibration models. However, method validation presented analytical results similar to those of the performance of full spectral PLS models.

The full-spectrum PLS model did not show a significantly better prediction than the iPLS or mwPLS model but was still good for further analysis, as shown in **Fig. 3**. Chemometric indicators do not allow full assessment of the abilities of the PLS model, as reported by Hubert [23,24]. Method validation using accuracy profiles is crucial to the practical use of the PLS model, which scrutinizes the accuracy of the results by considering both systematic and random errors in terms of accuracy, precision, risk, and linearity.

The accuracy profile figures of each full-spectrum PLS model are shown in **Fig. 6**. In the integrating sphere sampling mode, the relative prediction error of the calibration sample set at any concentration was within the accepted limits of 15%. However, in the fiber optic sampling mode, the relative prediction error of the first two concentration levels (0.025 g/g and 0.035 g/g) exceeded the accepted limits of 15%.

Based on the different contents of the calibration samples sets, PLS model performance using the integrating sphere mode with

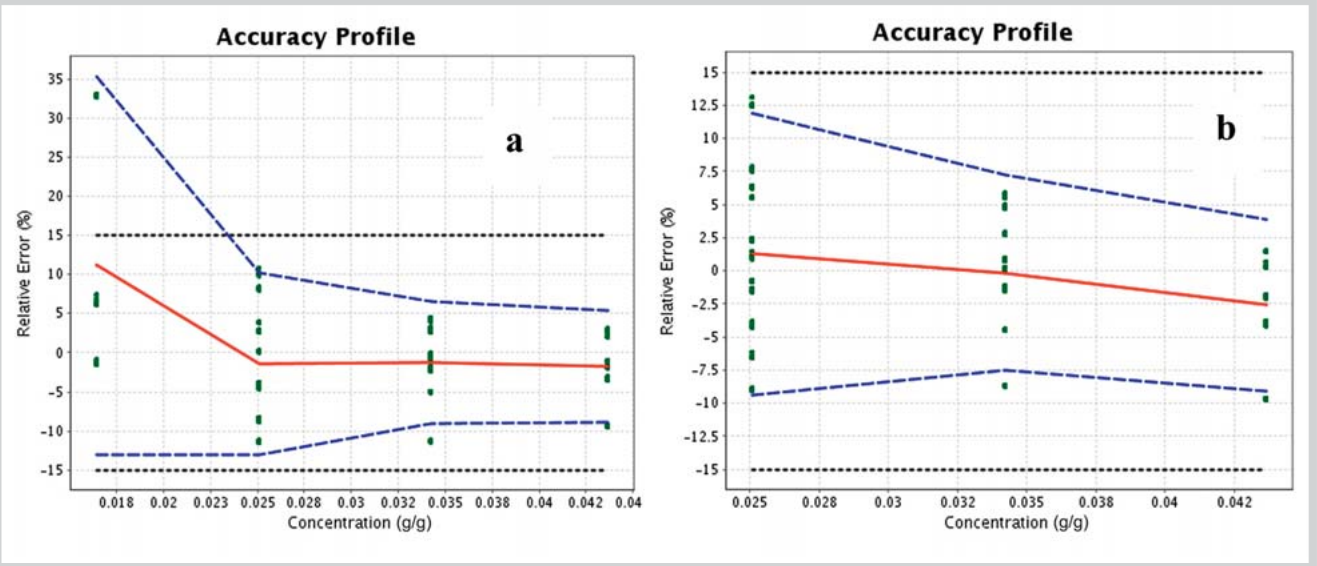

Fig. 5 Accuracy profile figure. **a** 1–4.5% of the Sample set using the integrating sphere mode; **b** 2–4.5% of the sample set using the integrating sphere mode sample. Black line: accepted limits; blue line: 90% β-expecting tolerance interval; red line: relative bias. (Color figure available online only.)

Table 4 Method parameters of near-infrared moving window partial least squares model based on ICH Q2(R1).

LEV	Accuracy			Precision		Uncertainty	
	REB	RTL	RIS	REP	INP	UB	REU
0.018	11.17	[−13.01, 35.34]	43.1	13.37	13.37	0.0005	27.55
0.025	−1.44	[−13.05, 10.18]	3.90	6.83	6.83	0.0003	13.82
0.035	−1.27	[−9.08, 6.53]	0.33	4.56	4.56	0.0003	9.24
0.043	−1.80	[−8.97, 5.38]	0.25	4.10	4.10	0.0004	8.37
0.025	1.28	[−9.34, 11.89]	2.44	6.24	6.24	0.0002	12.63
0.035	−0.16	[−7.55, 7.22]	0.16	4.31	4.31	0.0002	8.74
0.043	−2.60	[−9.08, 3.87]	0.17	3.70	3.70	0.0003	7.55

the calibration sample sets (2–4% and 3–4%) was better than that of the fiber optic probe mode. Comparison of PLS models with four different concentrations of calibration sample sets showed the accuracy profile results with the sample sets from 2–4% to be more accurate than those from 3–4%. When concentrations of the sample sets ranged from 1–2%, model performance could not satisfy the analytical requirements using both sampling modes (data not shown). It is here concluded that model performance is better when the sample sets have a relatively wide range of concentrations.

Quantitative analysis of baicalin in Yinhuang tablets was established by MEMS-NIR using two different sampling modes and different contents of samples sets. The experimental procedure included optimization of spectral pretreatments, selection of wavelength regions using iPLS and mwPLS, and validation of the method using an accuracy profile. The following factors were subjected to methodology validation: accuracy, precision, linearity, and risk. The selection of the wavelength region was found to make no significant difference to the model performance of full-spectrum PLS. The results demonstrated that quantitative models using the integrating sphere mode were better than those in the fiber optic mode. With respect to sample sets, model performance with a relatively wide concentration range was better than that with a narrow range.

The method has proved the reliability of the MEMS-NIR technique using two sampling modes in CHM. The validation process was based on different concentrations of calibration samples sets and assessed the phenomenon of model performance in different sample sets using two different sampling modes. Final results provide data support for the establishment of a steady model in CHM.

Materials and Methods

Materials

Three production batches of Yinhuang particles were purchased from Jimin pharmaceutical Co., Ltd. and deposited in the Key Laboratory of TCM-information Engineering of State Administration of Traditional Chinese Medicine (lot number 1311081, lot number 1311082, and lot number: 1311083). Baicalin reference standard (purity above 98%, lot number 624–200104) was supplied by the National Institute for the Control of Pharmaceutical and Biological Products. HPLC grade methanol was purchased from Tedia. Deionized water was purified usuing a Milli-Q water system (Millipore Corp.).

Experimental design

Different amounts of three batches of Yinhuang particles were mixed with starch in a mortar to produce a mixture of 1–4% baicalin by mass. The mixture was compressed into tablets as guided

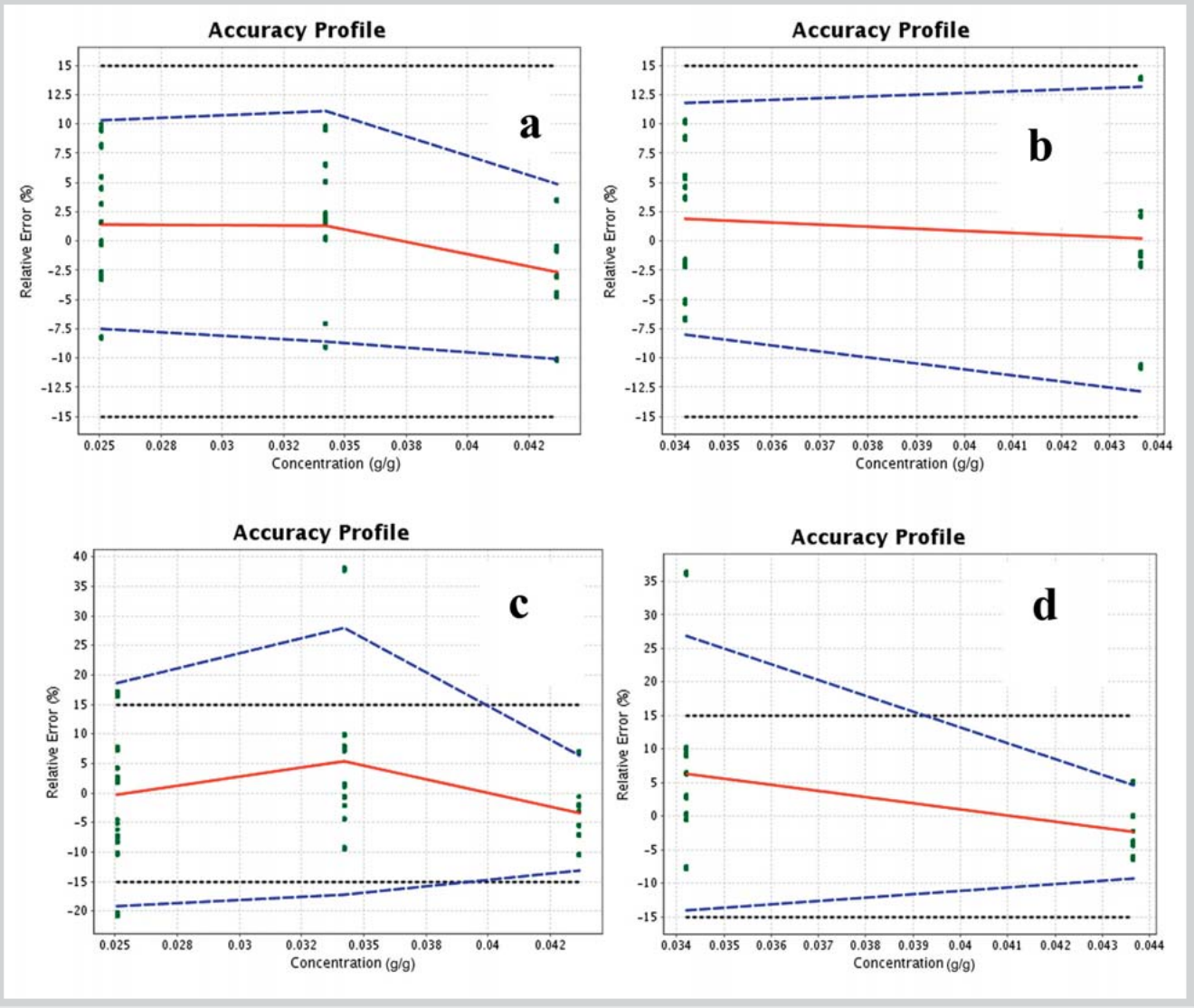

Fig. 6 Accuracy profile figure. **a** 2–4.5% of the Sample set using the integrating sphere mode; **b** 3–4.5% of the sample set using the integrating sphere mode; **c** 2–4.5% of the sample set using the fiber optic mode; **d** 3–4.5% of the sample set using the fiber optic mode. Black line: accepted limits; blue line: 90% β-expecting tolerance interval; red line: relative bias. (Color figure available online only.)

by the Chinese Pharmacopoeia (Ch.P) [25]. Details about the experiment design are shown in **Table 1**.

Near-infrared equipment and software

The NIR spectra were collected with a 410 NIR system (Axsun Technologies, Inc.) using the diffuse reflectance fiber optic probe mode (probe material: 316 stainless steel; Axsun Technologies, Inc.) and the integrating sphere mode (Axsun Technologies, Inc.). Each spectrum was the result of 16 scans in the range between 1350 nm and 1800 nm using 1 nm resolution. To prevent outer environmental conditions such as temperature and humidity from affecting the results, the temperature was controlled at 25 °C and the humidity was kept at constant level.

Every sample was scanned three times, and the final spectrum used for each sample was an average of the three results. Different operators (two individuals) collected the NIR spectra in two days in order to produce accurate profiles. All NIR spectra were collected and archived using Axsun Result software.

Data analysis was performed by using homemade routines programmed in MATLAB code (MATLAB, The MathWorks). The toolbox used to select the most informative variables, called iPLS and MWPLS, was downloaded from http://www.models.kvl.dk/. The calculation of the accuracy profile was adopted using e. noval V3.0 software (Arlenda).

High-performance liquid chromatography method

The Agilent 1100 HPLC system, with a vacuum degasser, a quaternary pump, an auto-sampler, a thermostatic column compartment, and a diode array detector (DAD), was used. Separation was performed on a Dikma Diamonsil C18 column (250 mm × 4.6 mm with a 5-µm particle size) with isocratic elution of the mobile phase consisting of methanol, water, and phosphoric acid (50 : 50 : 0.2, v/v) at a flow rate of 1.0 mL/min. Column temperature was kept constant and the detection wavelength was set to 274 nm.

Fig. 7 shows a typical HPLC chromatogram of the baicalin reference standard and a Yinhuang tablet. The methodological parameters of validation were linearity, precision (repeatability and intermediate precision), stability, and accuracy (recovery), which

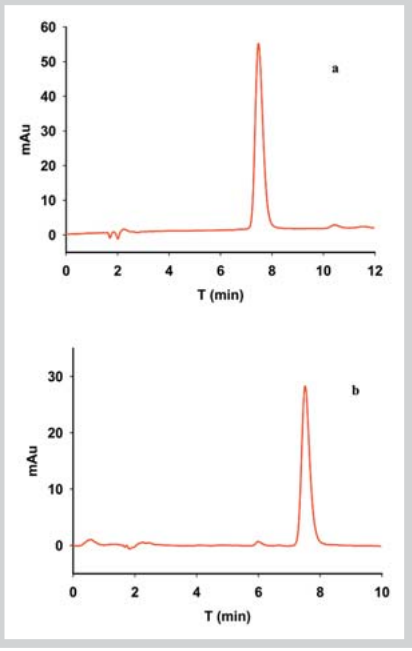

Fig. 7 Chromatograms of the baicalin reference standard (**a**) and Yinhuang tablets (**b**). (Color figure available online only.)

were consistent with the description in Ch. P [22] and previous reports on Yinhuang oral solution [7].

Supporting information

The preprocessing spectra using the integrating sphere and fiber optic probe modes are shown in **Figs. 1S** and **2S**. In addition, PLS models of different sampling modes and sample sets for different preprocessing methods are shown in **Tables 1S-6S**.

Acknowledgements

This work was financially supported by the Beijing Municipal Government for the university affiliated with the Party Central Committee (Prof. Shi), National Natural Science Foundation of China (81303218), and Doctoral Fund of Ministry of Education of China (20130013120006) and the Innovation Team Foundation of Beijing University of Chinese Medicine, Beijing (Research Centre of TCM-information Engineering), and are all gratefully acknowledged.

Conflict of Interest

The authors declare that there is no conflict of interest with any financial organization regarding material discussed in this manuscript.

References

1 *Qiao X, Liu CF, Ji S, Lin XH, Guo DA, Ye M.* Simultaneous determination of five minor coumarins and flavonoids in *Glycyrrhiza uralensis* by solid-phase extraction and high-performance liquid chromatography/electrospray ionization tandem mass spectrometry. Planta Med 2014; 80: 237–242
2 *Weeranantanaphan J, Downey G, Allen P, Sun DW.* A review of near infrared spectroscopy in muscle food analysis: 2005–2010. J Near Infrared Spec 2011; 19: 61–104
3 *Rasanen E, Sandler N.* Near infrared spectroscopy in the development of solid dosage forms. J Pharmaceut Biomed 2007; 59: 147–159
4 *Luypaert J, Massart DL, Heyden YV.* Near-infrared spectroscopy applications in pharmaceutical analysis. Talanta 2007; 72: 865–883
5 *Wu ZS, Du M, Sui CL, Shi XY, Qiao YJ.* Development and validation of NIR model using low-concentration calibration range: rapid analysis of *Lonicera japonica* solution in ethanol precipitation process. Anal Methods 2012; 4: 1084–1088
6 *Xu B, Wu ZS, Lin ZZ, Sui CL, Shi XY, Qiao YJ.* NIR analysis for batch process of ethanol precipitation coupled with a new calibration model updating strategy. Anal Chim Acta 2012; 720: 22–28
7 *Wu ZS, Xu B, Du M, Sui CL, Shi XY, Qiao YJ.* Validation of a NIR quantification method for the determination of chlorogenic acid in *Lonicera japonica* solution in ethanol precipitation process. J Pharmaceut Biomed 2012; 62: 1–6
8 *Wu ZS, Ma Q, Lin ZZ, Ai L, Shi XY, Qiao YJ.* A novel model selection strategy using total error concept. Talanta 2013; 107: 248–254
9 *Wu ZS, Sui CL, Xu B, Ai L, Ma Q, Shi XY, Qiao YJ.* Multivariate detection limits of on-line NIR model for extraction process of chlorogenic acid from *Lonicera japonica*. J Pharmaceut Biomed 2013; 77: 16–20
10 *Sun S, Chen J, Zhou Q, Lu G, Chan K.* Application of mid-infrared spectroscopy in the quality control of traditional Chinese medicines. Planta Med 2010; 76: 1987–1996
11 *Cozzolino D.* Near infrared spectroscopy in natural products analysis. Planta Med 2009; 75: 746–756
12 *Lu HY, Wang SS, Cai R, Meng Y, Xie X, Zhao WJ.* Rapid discrimination and quantification of alkaloids in Corydalis Tuber by near-infrared spectroscopy. J Pharmaceut Biomed 2012; 59: 44–49
13 *González-Caballero V, Pérez-Marín D, López MI, Sánchez MT.* Optimization of NIR spectral data management for quality control of grape bunches during on-vine ripening. Sensors 2011; 11: 6109–6124
14 *Ni YN, Song RM, Kokot S.* Discrimination of Radix Isatidis and Rhizoma et Radix Baphicacanthis Cusia samples by near infrared spectroscopy with the aid of chemometrics. Spectrochim Acta A 2012; 96: 252–258
15 *Chen Y, Xie MY, Zhang H, Wang YX, Nie SP, Li C.* Quantification of total polysaccharides and triterpenoids in *Ganoderma lucidum* and *Ganoderma atrum* by near infrared spectroscopy and chemometrics. Food Chem 2012; 135: 268–275
16 *Lin H, Chen QS, Zhao JW, Zhou P.* Determination of free amino acid content in Radix Pseudostellariae using near infrared spectroscopy and different multivariate calibrations. J Pharmaceut Biomed 2009; 50: 803–808
17 *Lai YH, Ni YN, Kokot S.* Discrimination of Rhizoma Corydalis from two sources by near-infrared spectroscopy supported by the wavelet transform and least-squares support vector machine methods. Vib Spectrosc 2011; 56: 154–160
18 *Shi JY, Zou XB, Zhao JW, Holmes M, Wang KL, Wang X, Chen H.* Determination of total flavonoids content in fresh *Ginkgo biloba* leaf with different colors using near infrared spectroscopy. Spectrochim Acta A 2012; 94: 271–276
19 *Ziemons E, Barillaro V, Rozet E, Mbakop NW, Lejeune R, Angenot L, Thunus L, Hubert P.* Direct determination of tagitinin C in *Tithonia diversifolia* leaves by on-line coupling of supercritical carbon dioxide extraction to FT-IR spectroscopy by means of optical fibres. Talanta 2007; 71: 911–917
20 *Roggo Y, Chalus P, Maurer L, Lema Martinez C, Edmond A, Jent N.* A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. J Pharmaceut Biomed 2007; 44: 683–700
21 *Wu ZS, Du M, Xu B, Lin ZZ, Shi XY, Qiao YJ.* Absorption characteristics and quantitative contribution of overtones and combination of NIR: method development and validation. J Mol Struct 2012; 1019: 97–102
22 *Gonzalez AG, Herrador MA.* A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. Trac-Trend Anal Chem 2007; 26: 227–238
23 *Mantanus J, Ziemons E, Lebrun P, Rozet E, Klinkenberg R, Streel B, Evrard B, Hubert P.* Active content determination of non-coated pharmaceutical pellets by near infrared spectroscopy: method development, validation and reliability evaluation. Talanta 2010; 80: 1750–1757
24 *Bouabidi A, Rozet E, Fillet M, Ziemons E, Chapuzet E, Mertens B, Klinkenberg R, Ceccato A, Talbi M, Streel B, Bouklouze A, Boulanger B, Hubert P.* Critical analysis of several analytical method validation strategies in the framework of the fit for purpose concept. J Chromatogr A 2010; 1217: 3180–3192
25 *Chinese Pharmacopoeia Commission.* Pharmacopeia of People's Republic of China. Beijing: China Medical Science Press; 2010: 1082–1084