AAPS PharmSciTech **2003; 4 (2) Article 24 (http://www.pharmscitech.org).**

Evaluation of the Conformity Index and the Mahalanobis Distance as a Tool for Process Analysis: A Technical Note

Submitted: February 11, 2003; Accepted: April 29, 2003

Gary E. Ritchie[1], Howard Mark[2], and Emil W. Ciurczak[3]

[1]Purdue Pharma LP, Ardsley, NY10502
[2]Mark Electronics, Suffern, NY 10901
[3]Integrated Technical Solutions, Golden's Bridge, NY 10526

KEYWORDS: conformity index (CI), Mahalanobis distances (MD), near-infrared analysis (NIR), process analytical technology (PAT), validation

INTRODUCTION

If a fully functioning process is controlled by sensors placed at critical points along the process stream, no one sensor need be a fully independent analyzer. Since the analysis of the final product begins with raw material testing and continues through the entire stream, no one test stands alone, but each is a part of the analysis continuum. With this in mind, the question arises whether a relatively simple test, such as the Conformity Index (CI) (as proposed in 1993 by W. Plugge and C. Van Der Vlies[1]) may be used in lieu of a full chemometric analysis. Similarly, Mahalanobis distances (MD)[2] (proposed slightly earlier, in 1992, as a similar form of quality control procedure[3] although not applied to pharmaceutical analysis at that time) also seems eminently suited for this purpose.

In the study by Plugge and Van Der Vlies, the authors proposed the use of near infrared spectroscopy (NIR) as an alternative to several compendial test methods such as identification, water content, and assay for ampicillin trihydrate.[1] The key question that arises is whether or not such a method can be validated to satisfy current FDA requirements, as expressed in publicly available documents outlining the standards with which all methods of pharmaceutical analysis are compared.[4-8] At the time Plugge and Van Der Vlies developed their method, no reference guidelines for NIR were in place to use in evaluating their method. Therefore, this paper examines the question of whether or not such a method, based on their proposed statistic or on MD, could be accepted today.

Considerable progress has been made since 1993 in developing method validation parameters for use in pharmaceutical laboratories and especially in developing guidelines for validating methods utilizing NIR spectrophotometry.[4-14]

The requirements for validation of an analytical method as expressed in the official documentation[4-8] includes, for major constituents, verifying the following characteristics: (1) accuracy, (2) intermediate precision, (3) linearity, (4) range, (5) repeatability, (6) robustness, and (7) specificity. These characteristics are required to be tested, not only for NIR but also for any method developed for the analysis of a major ingredient, whether for assay or for content uniformity.

Thus, while the method of Plugge and Van Der Vlies was revolutionary and acceptable a decade ago, this study will evaluate whether or not that same work would be accepted today, in the light of the progress that has been made in the development of validation methodology. In this study, this question is examined in detail, beginning with an examination of the Conformity Index proposed by Plugge and Van Der Vlies as well as MD, starting with a detailed summary of the methodology used by Plugge amd Van Der Vlies.

Corresponding Author: Howard Mark, Mark Electronics, 21 Terrace Ave Suffern, NY 10901. Phone: (845) 357-9496; Fax: Not Forwarded; Email: hlmark@nearinfrared.com.

METHODOLOGY

Reference to compendial test methods refers to the Pharmeuropa (Ph Eur) and British Pharmacopeia (BP) methods for identification (compared with a European Pharmacopeial Chemical Reference Substance by mid-infrared [mid-IR]), water content (Ph Eur BP, Karl Fisher [KF]), and assay (hydroxylamine colorimetric

assay method). Plugge et al used the following validation parameters to verify method suitability for the use of Near-Infrared Spectroscopy (NIRS):

1. Identification based on spectral match value

2. Water Content (based on NIR wavelengths specific for O-H absorptions attributed to water). The validation of this analysis included the following parameters:

A. Accuracy

1. The accuracy of the assay method for water content was established by selecting 10 batches outside of the compendial limits for water content (12.0%-15.0%) that ranged from 7.1% to 11.6% and by constructing a calibration curve based on 2 NIRS wavelengths, 1642 nm and 1930 nm. The rest of the outlier batches were reserved for validation of the calibration curve.

2. Four-thousand-nine hundred and fifty-two batches were found to have an average of 13.1% with a standard deviation (SD) of ±0.2%.

3. Four-hundred and seventy-four out of 4952 (9.57%) batches were scanned by NIR and yielded an average of 13.2% with a SD of ±0.2%. Since the average NIR value is close to the average reference value indicating low bias, and the SDs agree, Plugge and Van Der Vlies conclude that the NIR measurement is satisfactorily accurate.

B. Precision

1. The precision was established by filling 10 individual cells with powder from a well homogenized batch sample and measuring each cell.

2. One cell was filled 10 times with the same powder and measured each time.

3. One cell was filled once and measured 10 times.

4. The Relative Standard Deviation (RSD) was calculated for each experiment above.

C. Precision

1. Seventeen batches were used to compare the KF versus the NIR water results by constructing a linear plot. The least squares equation is given as follows:

$$KF = b + mx \quad (1)$$

where,

KF = the predicted water value

b = the y-intercept

m = the slope

x = the NIR water value

2. The SD for the slope and intercept were also reported.

3. The correlation coefficient was reported.

An assessment for the SD of the intercept and slope were given in which the reported SD for the intercept is a measure of the bias and is assessed from 0, whereas the SD for the slope is a measure of its deviation from unity.

D. Ruggedness

1. The authors recommended that if one model is used for more than one instrument, bias corrections should be made to adjust the model to each instrument.

2. Slope adjustments between instruments are not usually performed.

3. Operator variability was present in the cell preparation; therefore this source of variability was built into the model in the form the spectra selected to be used for the reference spectra.

4. Assay based on the use of a newly proposed CI

(a) The use of an average value called "standard activity" was assigned to every batch in which the assay value lies within a range of ±3 times the SD of the assay method. These batches were declared to be of "standard quality." For example, the standard activity of 85.5% was assigned to each batch of ampicillin trihydrate for which the assay result is found to be between 84.2% and 86.8%, those values being the 99.9% confidence limits of the standard activity.

(b) The justification for the above is based on the assay results of 4952 production batches, and 388 assay results of the same control samples over a 2-year period.

E. Conformity

Plugge and Van Der Vlies introduced the CI as a measure of the "degree of conformity" of a batch with samples of standard quality.[1] MD have also been proposed as a means of assessing whether or not a sample is consistent with a set of data known to represent satis-

factory product.[3] The CI was calculated as described above (the MD calculation is described below):

1. The authors then calculated the second derivatives of each of the 30 spectra obtained as the identification reference spectra. While the parameters for the derivative calculation are not given, the note that "... information on 17 wavelengths at both ends of the spectrum gets lost" leads us to surmise that the authors used a 35-point Savitsky-Golay calculation for the second derivative. This same parameter was therefore applied to the MD calculation. Also, despite their statement that they computed the average absorbance and SD, it appears from their description that they actually computed the average second derivative and SD of the second derivatives. Therefore, these same calculations were also performed to compute the CI for the tablet data.

2. They then calculated the average value of these second derivative spectra ($\bar{X}_{abs}$) and their SD at each wavelength.

3. The average reference spectrum and the SD spectrum were calculated.

4. The unknown spectrum is measured and its second derivative calculated.

5. A Q_w value is calculated at each wavelength by dividing the absolute value of the difference between the derivative value at each wavelength of the sample spectrum and the reference spectrum by the SD of the reference spectra at the same wavelength.

6. The CI is the maximum Q_w found for the unknown sample.

7. CIs for 324 approved production samples were calculated.

The results found from applying this CI calculation to various samples were as follows:

1. For the 17 batches that had CIs greater than ±5 SD, a deviation from the normal process had been reported, although the samples passed the traditional (compendial) tests.

2. CI sensitivity was demonstrated by evaluating 17 out-of-spec batches that had anhydrous crystalline form contamination or a high level of water content.

3. Competitor lots were also tested against the established reference spectrum for CI.

4. Physical effects were evaluated. The particle size (micronized versus compacted material were evaluated for CI) and contamination from the addition of 1% by weight of magnesium stearate.

5. Recrystallization of the trihydrate also generated a different impurity profile and different particle size distribution; this was detected by the CI.

6. The authors proposed the use of the CI instead of the construction of an assay calibration curve based on a mixture of regular (nominal), treated, aged, or blends made with sodium carbonate.

The authors' recommendations were as follows:

1. The adoption of alternate test methods does not supersede the compendial tests in instances of dispute of the contested test result.

2. The operative premise that supports the use of the proposed alternate test method is that the process is well established and that it can be demonstrated that the day-to-day variation observed in the release values can be ascribed to unavoidable variations of the analytical test method itself.

3. Samples that exceed the 5.0 SD level should be analyzed by the official method before that batch can be released.

Skip lot testing by KF and the hydroxylamine test should be performed on every 10th batch to verify that the nominal 13.4% water value and the standard activity value previously established still holds.

Mahalanobis Distances

MD have been described in detail,[2,15] and several potential applications have been described.[3,16] MD are computed from the matrix equation:

$$D = \sqrt{(X - \bar{X})M(X - \bar{X})'} \qquad (2)$$

where,

D = Mahalanobis distances

X = a spectrum (a vector quantity)

$\bar{X}$ = the mean spectrum from a set of samples (a vector quantity)

M = the matrix inverse of the pooled variance-covariance matrix from a set of samples

A constant value of D defines an n-dimensional surface enclosing a specified proportion of the data, essentially the multivariate equivalent of a confidence interval. A simplified explanation is that data representing known good material (eg, the samples used by Plugge and Van Der Vlies to calculate the criteria for the CI) can be separated from the nonconforming material by surrounding the good material with a surface enclosing all

Table 1. A Summary of the Assay Values for the Different Sample Sets

	Number of Samples	Minimum Assay Value	Maximum Assay Value	Low Sub Range	High Sub Range
Production samples	260	181.3	210.7	N/A	N/A
In-spec development samples	141	181.9	209.7	N/A	N/A
Out-of-spec development samples	254	151.6	239.1	151.6 to 181.1	213.8 to 239.1

the good material. Thus, anything inside the surface is good, and anything outside the surface is not. The 2 calculations can, in fact, be seen to have a similar goal with the main difference being the method used to calculate the region separating the good versus not–good materials.

Principal Component Analysis (PCA) is a mathematical transformation of spectral data that has enjoyed a long and largely successful tenure in NIR spectroscopy, being applied to pharmaceutical analysis as long ago as 1989.[17] PCA is also widely used today and is known to often address some problems of NIR analysis. Because PCA thoroughly reorganizes the information in a set of spectral data, applying it to a set of data before computing the MD makes it tantamount to a separate and independent method of analyzing the data, while still legitimately falling under the heading of MD.

MATERIALS AND METHODS

Comparison of the methods was performed using data previously reported.[14] The tablet data from that study are available online at: http://www.idrc-chambersburg .org/Shootout.html.

The data represent 13 production lots plus a set of laboratory development samples. The production samples have a limited range of the analyte, as normally expected from a production process in good control. The development samples were specifically created to extend the range of analyte beyond the values available in the production samples in order to facilitate the work reported.[14] For the previously reported work, the samples were divided into calibration, validation, and test sets. While these data do not include the wide range of variability in the Ampicillin samples used by Plugge and Van Der Vlies, they include the extraneous variation of "production" versus "synthetic" samples.

For the purpose of the current evaluation the samples were separated differently. The entire sample from the 3 sets described was combined and regrouped into 2 sets: (1) all the production samples with their limited range of assay values; and (2) all the other samples, which represent samples both within and beyond the range of values in the production set.

The nonproduction samples were further subdivided into 2 subsets: those within the production range (in-spec) and those with assay values beyond the production set range (out-of-spec). A summary of the characteristics of each subset is listed in **Table 1**.

The production samples were used to develop the criteria for calculating both the Conformity Index and the MD methods for testing the agreement of the samples with the known-good production data. For the Conformity Index, the second derivatives were calculated using, as described above, a quadratic polynomial fitting function and 35 points to create the fit.

MD were originally explained as using the original raw absorbance data.[2] In the referenced study describing the data,[14] the calibrations were performed using the Multiple Linear Regression (MLR) algorithm, after transforming the data using multiplicative scatter correction (MSC) followed by a first derivative calculation. Therefore, for the current study, the data were subjected to the same transformation, and the wavelengths used in the referenced study, which were shown to be characteristic of the active ingredient in these tablets, were used to compute MD.

In order to investigate the effect of performing PCA prior to calculating MD, a small and prespecified set of transformation conditions was applied to the data prior to performing the PCA and MD computations.

To perform this type of data analysis while retaining the concept of basing the data analysis on the previously-validated results[14] the following appropriate methodology was used:

Table 2. Summary Results for CI and MD Computed from the Tablet Data*

		Conformity Index			MD Computations from Wavelength Data			MD Computations From PCA (5 factors) with No Data Transform			MD Computations From PCA (5 factors) with Data Transform (See Text)		
	# of samples	Min CI	Max CI	% > limit	Min MD	Max MD	% > limit	Min MD	Max MD	% > limit	Min MD	Max MD	% > limit
Production Samples	260	1.21	6.94	N/A	0.633	4.03	N/A	0.874	4.99	N/A	0.941	4.40	N/A
In-spec development samples	141	2.29	7.17	0.71	36.8	49.4	100	0.966	4.18	0	0.788	3.99	0
Out-of-spec development samples	254	3.73	20.6	77.2	30.5	60.0	100	0.875	4.91	0	0.735	6.17	1.18

*CI indicates conformity index; MD, Mahalanobis distances; and PCA, principal component analysis.

(1) Perform PCA on the original spectra, and compute MD on those untransformed data.

(2) Perform PCA on the data as transformed according to the already-validated transformation used for those data, ie, MSC followed by computation of the first derivative ($dA/d\lambda$).

Each of these data transformations was applied, and computations of the MD matrices were performed on the resulting values; then the MD matrices were applied to the production samples to compute the threshold, and the same computations for the non-production samples were performed that were used for the wavelength-based MD calculations.

RESULTS AND DISCUSSION

After performing the CI and MD computations, the resulting criteria were then used to assess both the in-spec and out-of-spec development samples. **Table 2** summarizes the results found. Plugge's computation of the CI is a semi-empirical calculation, probably selected after trials of other candidate algorithms. The distribution of CI values is unknown, as far as the authors this study are aware. This distribution is probably unknown to Plugge and Van Der Vlies, as well as to the authors of this study. The threshold value of 5 set by Plugge and Van Der Vlies for the Ampicillin analysis was also selected empirically. From the results of applying the CI calculation to many samples, all of Plugge and Van Der Vlies' samples of "standard activity" were found to have CI values of less than 5. For the tablet analyses, therefore, an empirical threshold was set for the CI calculation for the tablets. In this case, **Table 2** shows that all tablets from the set of calibration samples had CI values of less than 7, thus this became the threshold for CI.

Whitfield et al have calculated theoretical values for confidence limits for MD.[15] The values from Whitfield's tables are in good agreement with the empirical values for the MD values from all calibration samples. Based on these combined criteria, the threshold for MD was set at 4. Based on the maximum value of each statistic as shown in **Table 2**, therefore, the threshold (limit) value for CI was set at 7 and for MD at roughly 4. The percentage of the samples found to be beyond the limit was also computed for each sample set and is listed in **Table 2**.

The CI and the MD calculations reacted differently to the differences in the sample sets. As **Table 2** shows, the CI did not pass all the in-spec samples nor did it fail all the out-of-spec samples. Nevertheless, it did a creditable job in flagging and assigning the majority of samples in both sets. It may well be that the optimum parameters depend on the sample type and are likely to be different for different samples. We noted that the criteria developed for Ampicillin may not be optimum for other types of samples, such as the ones tested here.

MD, on the other hand, did an excellent job of flagging all the samples that were not made by the production process, ie, both the in-spec and out-of-spec synthetic laboratory (development) samples. This performance demonstrates its ability to flag nonstandard and potentially counterfeit samples.

Comparison of the Methods with the Guidelines

Based on the summary of the study discussed above, the following question is addressed: Would the same

Table 3. Comparison of NIR Results From Plugge et al With the <1119> Guidelines*

NIR Method Validation Parameter	Proposed <1119>	NIR of Ampicillin Trihydrate Proposed
Specificity	Challenges from similar compounds. The bands of the analyte of interest should be free from interference from other competing absorption bands. Wavelengths, loadings or factors can be inspected for corresponding analyte information. Coefficients can be plotted, and the regions of large coefficients compared with the spectrum of the analyte matrix variation can be used to assess effect on corresponding analyte.	Challenges from anhydrous ampicillin, amoxicillin, and other similar compounds were performed. Identification based on SMV
Linearity	A line plot for linearity, slope, intercept, and correlation coefficient are all reported	For water determination only, a plot of KF versus NIR was presented (for the n = 17 samples used therein). The intercept, slope, SD of the slope, SD of the intercept, and the correlation coefficient were reported. Visual examination of plot of NIR versus KF water analyses shows possible marginal nonlinearity.
Range	The range of analyte reference values in the validation set defines the range of the NIR method. The range of analyte reference values also effectively defines the quantification limits for an NIR method. A limited sample set does not preclude the use of an NIR method.	The lower and upper bounds of the linear plot define the range for water determination. The use of the newly proposed CI test is based on the premise of a limited sample set for the assessment of the hydroxylamine assay. The range was set at ±5 SD around the mean CI value.
Accuracy	Accuracy can be indicated by how close the Standard Error of Prediction (SEP) is to the standard error of the reference method used for validation. The error of the reference method may be known based on historical data or a measurement of the Standard Error of the Laboratory (SEL) may be carried out. Several statistical comparison methods can be applied to the NIR predicted values and reference values from the validation set samples to determine if there is any statistical difference between the results from the 2 methods, at a specified confidence limit (eg, paired *t* test, bias evaluation).	See range statement above.

Table 3 (Continued)

NIR Method Validation Parameter	Proposed <1119>	NIR of Ampicillin Trihydrate Proposed
Precision	Repeatability Statistical evaluation of a number of replicate measurements of the same sample without variation in sample position. Statistical evaluation of multiple sample positioning or aliquots as appropriate.	For water determination only, the precision was established by filling 10 individual cells with powder from a well homogenized batch sample and measuring. One cell was filled 10 times with the same powder and measured each time. One cell was filled once and measured 10 times. SD and RSD values were calculated for each precision measurement described above.
Robustness	Analyst should demonstrate effects of environmental variation, sample presentation, and instrument variability.	Not applicable
Ruggedness	N/A	For water determination only, the authors recommended the use of a bias correction to adjust the model if it is to be used by more than one instrument. Slope adjustments between instruments are not usually performed. Operator variability was present in the cell preparation; therefore, this source of variability should be built into the model in the form the spectra selected to be used for the reference spectra.
Ongoing model evaluation	Ongoing monitoring of method accuracy, precision, or other suitable parameters.	Skip lot testing by KF and the hydroxylamine test should be performed on every 10th batch to verify that the nominal 13.4% water value and the standard activity value previously established still hold.
Model transfer	Transfer of electronic models to a second instrument requires that procedures and criteria must be applied to demonstrate that the model remains valid on their second instrument.	N/A

* CI indicates conformity index; KF, Karl Fisher; NIR, near-infrared analysis; Relative Standard Deviation; and SMV, spectral match value.

data be accepted today based on the newly proposed <1119> method validation parameters for NIR?[14]

Table 3 lists the <1119> method validation parameters and the corresponding NIR parameters for ampicillin trihydrate presented above.

The study by Plugge and Van Der Vlies[1] deals with the analysis of 2 constituents in the powder they are concerned with. One is the active ingredient (ampicillin); the second is moisture. The moisture is treated as a constituent to be quantified, and the validation of the NIR analytical methodology for moisture addresses the several characteristics of a quantitative method that the guidelines require: repeatability, intermediate precision, accuracy, linearity, and range. The evidence shown by Plugge and Van Der Vlies for specificity of their NIR measurement for moisture analysis is weak. However, the results achieved by computing the spectral match values (SMV) of both anhydrous ampicillin and production ampicillin containing moisture outside

the allowed limits can be taken as indirect evidence of the specificity of the NIR response to moisture.

The validation of the analysis for ampicillin is less straightforward. Indeed, the report begins with a discussion of a fairly extensive series of tests showing that ampicillin trihydrate can be identified and that it can be distinguished from other similar antibiotics and from other forms of ampicillin. Qualitatively, this procedure might suffice to identify the product, but identification is not the goal of the study, nor of the regulations governing pharmaceutical production; it is only 1 step on the road to validating the analytical method. Validation of a method of qualitative analysis is also not the primary goal of either Plugge and Van Der Vlies' study[1] or this one.

While Plugge and Van Der Vlies indeed address some of the validation requirements for identification of a qualitative method of NIR analysis, in that they challenged their method with other, similar materials, their work should not be compared with the guidelines for validating methods of identifying pharmaceutical materials.

Plugge and Van Der Vlies need to verify that the result of their production process conforms to the specified requirements of the product, which is the percentage of pharmaceutically active material in the output of the process. This need is a problem in quantitative NIR analysis because of the limited range of concentration of ampicillin in the samples that were available for the calibration development. This limited concentration range is a common problem when seeking to perform NIR analyses, and recent work has recommended increasing the range with pilot plant and/or laboratory development samples.[7-14] At that time, however, Plugge and Van Der Vlies felt that it would suffice to use the CI to show that materials whose spectra conformed to the optical behavior of known good samples would themselves constitute good samples. Reproducing Plugge and Van Der Vlies' method and simultaneously testing the application of MD to the same problem is in itself not an unreasonable approach (this point is considered below). However, in the light of modern guidelines, both approaches are wanting in the area of meeting current validation requirements.

At the most basic level, Plugge and Van Der Vlies do not provide any indication as to either the precision or the accuracy of the method. While Plugge and Van Der Vlies show the precision, and even the distribution, of the CI values they calculate, no indication is given of the precision of the corresponding concentration values for the ampicillin content. In addition, they are unable to show any connection between the CI precision and the precision of ampicillin measurement, except that both are bounded. MD also does not address these validation requirements.

Most of the other common validation requirements for a quantitative analytical method are also ignored, in addition to the accuracy and precision: linearity and intermediate precision. CI did not show any correspondence of the wavelengths used to known absorbance bands of the analyte. MD, on the other hand, did use the wavelengths found in the previous study, and therefore does address this issue. In addition, neither the CI nor the MD was tested in the face of environmental variations, such as temperature and humidity, or for the effect of using different instruments or having the test performed by different operators.

The only one of the standard validation parameters given for the ampicillin analysis is the range of concentrations of ampicillin in the samples, which was reproduced in the current tests also.

On the other hand, Plugge and Van Der Vlies' contention that samples with conforming spectra would constitute good material has merit. Since values of the CI or MD outside the allowable range indicate departure of the samples from the behavior of known good material, lack of such departure indicates that, in some respects, the samples do conform to the behavior of good material. It is not clear, however, which aspects of the material are included in this conformity. Nevertheless, such behavior on the part of an analytical method is not unknown in pharmaceutical practice.

By way of analogy, the US Pharmacopeia includes cases in which samples are subjected to several different tests, none of which stand alone as a definitive indication of the quality of the pharmaceutical product. An example of this is the compendial test for lactose monohydrate.[18] In order to verify that a given quantity of lactose is suitable for use in pharmaceutical products, the following tests are performed:

(1) Clarity of solution

(2) Color of solution

(3) Absorbance at 400 nm

(4) Thin Layer Chromatography

(5) Heating with ammonium hydroxide

(6) Specific (optical) rotation

(7) Acidity/alkalinity

(8) Loss on drying

(9) Residue on ignition

(10) <presence of> Heavy metals

(11) <presence of> Light-absorbing impurities

None of these tests (eg, test for reducing sugar) individually verifies either that the sample is lactose monohydrate or that it is sufficiently pure for pharmaceutical use. Even the chromatographic test for identification ascertains only that the sample at hand is not any of a group of other sugars (dextrose, sucrose, or fructose). The set of tests together, however, provide strong confirmation that the material at hand is lactose and is suitable for use.

A similar situation exists with the concept of using CI or MD. Alone, neither is sufficient to unequivocally identify or quantify the amount of material in a product. However, by limiting the range of possible substances and quantities, either CI or MD could be a valuable component in a battery of tests, which collectively would verify the identity and usability of a given batch of ampicillin. As in the lactose example, the CI, MD, and SMV guard against the possibility of a material being any of several similar materials or out-of-spec in other respects; although they do not identify the material uniquely or quantitatively determine the amount or concentration of ampicillin in the product. Therefore, supporting tests are needed.

CONCLUSION

In the course of reviewing the tests performed by Plugge and Van Der Vlies, they were compared with the list of requirements stated in the Introduction; the NIR measurement of moisture was found to address all the requirements, although their method of addressing specificity was weak. From this finding, the conclusion is that the protocols for NIR measurement of moisture presented by Plugge and Van Der Vlies would meet the current validation criteria and would be acceptable even under modern stringent guidelines.

The use of the CI and /or the MD (calculated for 3 different sets of computational parameters) as a replacement for an assay of the active ingredient was also compared with the list of analytical characteristics required by the guidelines. Since the MD computation had not been previously optimized, its performance was computed for a very small set of different computational conditions. Widely varying performance characteristics were found under the different conditions, from all prediction samples being flagged as out-of-range (for MD computations based on individually selected wavelengths) to none being flagged (using the PCA transformation for the otherwise untransformed spectral data), with the 1 trial pretreatment falling in between. This result confirms a suspicion that a good statistic requires a considerable amount of effort to optimize and that an extensive research effort would have to be undertaken to create a computational method intended to replace an actual assay, as was likely done with the CI itself.

Nevertheless, when compared with the requirements for validation of a method, the only characteristics addressed by the CI were the range and identity (the weak surrogate for specificity). The MD computation also addressed the question of specificity (somewhat more strongly) when the MDs were computed from individually selected wavelengths.[14] Even so, both of these methods of assessing the "goodness" of the data fail to address the majority of the requirements specified in the official guidelines. For this reason, the conclusion drawn here is that the proposal for use of CI or MD as a replacement for quantitative assay testing does not meet modern standards. Nevertheless, they are valuable weapons in an arsenal of test methods and could be part of a series of tests, which could be used together to ascertain the usability of a pharmaceutical substance. A series of tests such as these, used in conjunction with a method of rapid measurement such as NIR, can be part of a synergistic set of measurements that together create a reliable and efficient process measurement technology.

ACKNOWLEDGEMENTS

The authors thank Purdue Pharma LP (International Research and Development) Ardsley, NY, for the support received.

REFERENCES

1. Plugge W, Van Der Vlies C. Near-infrared spectroscopy as an alternative to assess compliance of ampicillin trihydrate with compendial specifications. J Pharm Biomed Anal. 1993;11(6):435-442.

2. Mark HL, Tunnell D. Qualitative near infrared reflectance analysis using Mahalanobis distances. Anal Chem. 1985;57(7):1449-1456.

3. Burns DA, Ciurczak EW. Handbook of Near-Infrared Analysis. 2nd ed. New York, NY: Marcel Dekker; 1992: 385-386.

4. Food and Drug Administration. Guideline for Industry: Text on Validation of Analytical Procedures. Rockville, MD: Food and Drug Administration; March 1, 1995. ICH-Q2A.

5. Food and Drug Administration. Guidance for Industry: Validation of Analytical Procedures: Methodology. Rockville, MD: Food and Drug Administration; May 19, 1997. ICH-Q2B.

6. US Pharmacopeia. <1225> Validation of Compendial Methods. Vol 24. Rockville, MD: US Pharmacopeia; 1999:2149-2152.

7. <1119> Near Infrared Spectrophotometry. Pharmacopeial Forum. Nov-Dec 2002;28(6):1119.

8. <1119> Near Infrared Spectroscopy [New] (2nd Supp). Pharmacopeial Forum. Sept-Oct 2002;27(5):1119.

9. European Pharmacopeia. European Pharmacopeia (EP): Physical and Physicochemical Methods, Near Infrared Spectrometry. Strasbourg, France: European Pharmacopeia; 1997.

10. Broad N, Graham P, Hailey P, et al. Guidelines for the Development and Validation of Near Infrared Spectroscopic Methods in the Pharmaceutical Industry. In: Chalmers JH, Griffiths PR, eds. Handbook of Vibrational Spectroscopy. Chichester, UK: John Wiley & Sons Ltd; 2002:1-22.

11. The European Agency for the Evaluation of Medicinal Products (EMEA). Note for Guidance on the Use of Near Infrared Spectroscopy by the Pharmaceutical Industry and the Data to be Forwarded in Part II of the Dossier for a Marketing Authorization. EMEA; November 14, 2001:1-13.

12. Moffat AC, Trafford AD, Jee RD, Graham P. Meeting the International Conference on Harmonization's Guidelines on Validation of Analytical Procedures: Quantification as Exemplified by a Near-Infrared Reflectance Assay of Paracetamol in Intact Tablets. Analyst. 2000;125:1341-1351.

13. Mark H, Ritchie GE, Roller RW, Ciurczak EW, Tso C, MacDonald SA. Validation of a Near-Infrared Transmission Spectroscopic Procedure, Part A: Validation Protocols. J Pharm Biomed Anal. 2002;28(2):251-260.

14. Ritchie GE, Roller RW, Ciurczak EW, Mark H, Tso C, MacDonald SA. Validation of a Near-Infrared Transmission Spectroscopic Procedure, Part B: Application as an Alternate Content Uniformity and Release Assay Method for Pharmaceutical Solid Dosage Forms. J Pharm Biomed Anal. 2002;29(1-2):159-171.

15. Whitfield RG, Gerger ME, Sharp RL. Near-Infrared Spectrum Qualification via Mahalanobis Distance Determination. Appl Spectrosc. 1987;41(7):1204-1213.

16. Burns D, Ciurczak E. Handbook of Near-Infrared Analysis. In: Practical Spectroscopy Series, 2nd ed. New York, NY: Marcel Dekker; 2001:4, 363-393, 569.

17. Kamat MS, Lodder RA, DeLuca PP. Near-Infrared Spectroscopic Determination of Residual Moisture in Lyophilized Sucrose Through Intact Glass Vials. Pharm Res. 1989;6(11):961-965.

18. US Pharmacopeia. Lactose Monohydrate (USP Method). Rockville, MD: US Pharmacopeia; 2002:2570. NF 20.