Pharmaceutical Research, Vol. 13, No. 4, 1996

Use of Near-Infrared Spectroscopy to Evaluate an Active in a Film Coated Tablet

**Bruce R. Buchanan,[1,2] Mark A. Baxter,[1]
T.-Show Chen,[1] Xue-Zhi Qin,[1] and
Priscilla A. Robinson[1]**

Received October 25, 1995; accepted January 23, 1996

Purpose. To provide a method to rapidly screen tablets in the development of new coating technology.

Methods. Near-Infrared (NIR) reflectance spectroscopy was used to quantitatively analyze tablets which were composed of a drug active encasing an active drug core. Diffuse reflectance NIR scans of 240 individual tablets over the range of 1100–2500 nm were obtained. High Performance Liquid Chromatography (HPLC) was used as the reference method.

Results. Both qualitative, Principal Component Analysis, and quantitative results showed a strong agreement between the NIR and HPLC methods. The NIR analysis was non-invasive and allowed subsequent testing of the tablets. The contents of the drug active contained in a drug coating was determined to ±4% of the target value using NIR analysis. Over 400 samples were analyzed in less than a month utilizing this technique which allowed the optimization of a new coating technology.

Conclusions. NIR analysis allowed the evaluation of the efficiency of a new drug film coating manufacturing process more quickly and inexpensively. Because the Near-Infrared method was non-invasive the tablets were available for further analysis unlike the chromatography method.

KEY WORDS: near-infrared spectroscopy; chemometrics; quantitation; process monitoring; coating technology.

INTRODUCTION

The application of Near-Infrared (NIR) Reflectance spectroscopy to pharmaceutical samples is proving to be a valuable means to provide fast and reliable results. For example, researchers have applied NIR reflectance spectroscopy to quantitate moisture in granulations, distinguish placebo tablets from those containing a drug active, and classify granulations according to physical characteristics (1–3). Application of Principle Component Analysis (PCA) to classify tablets according to type was accomplished rapidly and non-invasive using NIR reflectance spectroscopy (2–3). While the main goal of the studies was to separate placebo tablets from those with the drug active, the authors found that the method would also distinguish the different types of placebos, or differing drug active concentrations. Also, the method was able to classify the granulation according to physical characteristics (1). It was estimated that a 2-hour per sample pack Thin-Layer Chromatography (TLC) test could be effectively be replaced by a 7 minute per sample pack NIR test (3). Additionally the NIR method did not generate any waste compared to the TLC method.

A rapid and non-invasive assay method for quantifying drug concentration in the surface coating of a tablet was required to aid in the development of precision film coating processes. The precision film coating is part of a new manufacturing process to produce a combination product, consisting of an immediate-release drug active coating encasing an extended-release drug active core matrix. The analytical test method of choice for evaluation of the coating was High Performance Liquid Chromatography (HPLC). The selection of an HPLC method for the process development and scale-up of the coating process would have required a substantial amount of analytical support at a high cost of both time and money. Additionally, HPLC is invasive and the lengthy laboratory processing time required for analysis limits the number of samples which can be evaluated. Because film coating processes represent a relatively new application in coating technology, Near Infrared (NIR) spectroscopy was evaluated as an alternate analytical tool for rapid analysis of the drug active in tablet coatings. NIR analysis is non-invasive which allows for subsequent evaluation of the same tablets via an alternate analytical method, which in this study was HPLC.

When converting a film coating process from a pilot plant to production scale, it is important to evaluate the uniformity and precision of the drug coating after each trial. Implementation of NIR spectroscopy not only allowed rapid evaluation of the tablets, but also assisted in identifying "dead zones" in the Wurster column. This allowed immediate correction and revision of the sampling plan.

Prior to film coating, core tablets were compressed using standard concave tooling. First, a sub-coat was applied to the core tablets in a rotary pan. Next, a subsequent layer was applied over the sub-coat in a Wuster column. The compression punches were plain on one side and logo embossed on the other side.

MATERIALS AND METHODS

Data Collection and Analysis

NIR reflectance data were collected using a Model 6500 Near-Infrared Spectrophotometer (NIRSystems, Inc., Silver Springs, MD) equipped with a scanning grating monochromator and a Rapid Content Sampler (RCS) module. The RCS is a diffuse reflectance module that uses a 6-element PbS detector for signal collection. A short anhydroguide fiber was placed between the exit slit of the monochromator and the sample in order to randomize the output light. Data were collected using an IBM PS-2 model 90 XP computer and Near-Infrared Spectral Analysis Software (NSAS, NIRSystems, Inc.). To reduce noise, 32 scans per sample were obtained and averaged. A background scan, R_o, of a Coors ceramic disk was obtained and used as the reference. The spectra obtained consisted of $-\mathrm{Log}(R_s/R_0)$ over the wavelength range of 1100 to 2500nm, where R_s is the light reflected from the sample.

The tablets were positioned in such a way that the plain side, which contained no embossing, was being sampled. Proper positioning of the tablets over the anhydroguide fiber was achieved using specially machined centering plates. The centering plate was removed prior to obtaining spectra to avoid spec-

---

[1] Analytical Methods & Support, Merck and Co., Inc., P.O. Box 4 WP38-3, West Point, Pennsylvania 19486.

[2] To whom correspondence should be addressed.

0724-8741/96/0400-0616$09.50/0

tral reflections from the plate. Care was taken not to move or rotate the tablet because changes in the position of the tablet are known to affect the resulting spectrum. A total of two hundred and forty tablets were scanned, labeled, and stored. A sub-set of the stored tablets were analyzed later by HPLC. The results of these analyses were used in the chemometric models as the laboratory values. The results from a PCA analysis allowed us to select both adequate and inadequately coated tablets for inclusion in the 41 sample sub-set selected for HPLC analysis. The ideal drug film coating contains 100% of the target amount of active drug. Acceptable film coatings for developmental work range from 85% to 105% of the target amount of active drug.

Referee Method-HPLC Analysis

HPLC analysis of the selected tablets was performed by an in-house support laboratory using established validated methods. Results were reported as a percentage of the expected target concentration of the active drug coating. Because HPLC required the destruction of a tablet, NIR scans of the tablet were obtained prior to analysis by HPLC. The standards of active drug were manufactured by Merck Research Laboratories (Rahway, New Jersey, USA).

Calculations

Chemometrics is a branch of chemistry which uses statistical methods to interpret chemical or spectroscopic data. The goal of any chemometric calibration method is to reduce general and non-specific data, such as NIR spectra, to concise and useful information, such as a concentration. Three chemometric methods were investigated in this study, Multiple Linear Regression (MLR), Partial Least Squares (PLS), and Principal Component Analysis (PCA). Only an abbreviated description of each method will be presented in this publication. A more in-depth discussion can be found in references 4 through 7.

For spectral data, a set of wavelengths which are correlated to the parameter of interest is selected for the MLR method. Equation 1 shows the general calibration equation:

$$[C] = a[A] + b + [E] \tag{1}$$

where **[C]** is the concentration matrix, **[A]** is the matrix of selected vectors, and **[E]** is the residual error. The MLR method determines a vector of coefficients, **a**, and the scalar offset, **b**, which minimize the error **[E]**. The coefficients in **a** can then be used later to determine an unknown concentration using the measured absorbances **[A]**. Only a few wavelengths are chosen, typically less then 5, for an MLR calibration.

Partial Least Squares (PLS), a canonical method (i.e., it uses all wavelengths or a sub-set thereof), is another calibration technique. When performing a PLS calibration, the spectral data is reduced to a set of orthogonal vectors which are related to the parameter of interest. Each sample spectrum can then be mapped onto the set of orthogonal vectors. The positions on this map are designated as "scores" and can be used in Equation 1 instead of the absorbance. The main advantage of this method is that the majority of spectral noise will be relegated to higher order vectors, which are not used in the calibration. Also, the PLS calibration equation is usually more robust than an MLR calibration, and allows the detection of spectral outliers. A PLS regression is useful with small populations of samples that contain some experimental noise in the NIR spectra and/or the chemical data. In a multi-component system, overlap and convolution of individual species often preclude the use of a linear correlation between absorbance and analyte concentration. It is for this reason that the PLS method is more commonly used as a means of quantifying a single component in pharmaceutical mixtures.

In performing the data reduction, several statistical calculations are performed to evaluate the calibration. Care must be taken not to underfit the calibration curve, which yields an inadequate calibration, or overfit the calibration, which introduces noise into the calibration, when building a calibration equation. To assure that the model is not underfit, the Standard Error of Calibration (SEC), given in Equation 2, is calculated when building a model for each wavelength or when a PLS factor is added.

$$\mathrm{SEC} = \sqrt{[\Sigma(Y - \mu)^2/(n - m - 1)]} \tag{2}$$

In this equation the **Y** values are the laboratory values, μ are the NIR calculated values, n is the number of samples in the data set, and m is the number of factors or wavelengths used. The minimum SEC shows the number of wavelengths or factors which optimally fit the data. To avoid overfitting the data, the Standard Error of Prediction (SEP) is also calculated using Equation 2, except that the μ are now the predicted values obtained using Equation 1. The main difference between the SEC and the SEP is that the samples used in calculating the SEP are not included in the calibration set when developing the model. Ideally, the SEC and the SEP should be similar if the equation is to be considered robust and both data sets are representative of the sample population. Because the HPLC values are expressed in terms of percentage, all of the parameters in Equation 2 will have units of percentage.

To determine the equivalency of the HPLC and NIR methods, a single sample would have to be measured repeatably using each method and the intrasample variance respective to the "true" value obtained for each method should be compared. Because these data are not available for this study and the "true" value is not known, the sample-to-sample variance of each method was compared. If the intrasample variance is significantly greater for one method compared to the other, it should translate into the sample-to-sample variance. The variances of the NIR verification data are listed in Table I. The NIR variance

Table I. Statistical Summary for Verification Set with Sample 1-12B Omitted

Parameter	MLR	PLS
Number of samples	13	13
Bias	0.58	0.16
Root Mean Square(RMS)	3.36	1.90
SEP	3.30	1.88
Slope Adjustment	0.994	0.996
Intercept Adjustment	1.00	0.14
Simple Correlation	0.996	0.999
Achievable SEP	3.39	2.18
Degrees of Freedom	10	8
σ^2 off of mean[a]	105	103

[a] σ^2 off of mean for HPLC = 104.

is not significantly different compared to the HPLC variance, which is presented in a footnote of Table I. The F-statistic can be used to compare two or more variances, see Equation 3 (4, 5).

$$F_{\text{calculated}} = \sigma_1^2/\sigma_2^2 \tag{3}$$

The statistic is used to reject the hypothesis that $\sigma^2{}_1 = \sigma^2{}_2$, where σ^2 denote the variance of the data set. If the calculated F value is less than $F_{\alpha/2}$ or greater than $F_{(1-\alpha/2)}$, then the hypothesis can be rejected, where α is the area under the probability density function. The $F_{\text{calculated}}$ is 0.99 for the PLS model and 1.00 for the MLR model. The tabulated F values are 2.69 for ($\alpha/2$) and 0.37 for ($1 - \alpha/2$) for $\alpha = 0.05$. Both $F_{\text{calculated}}$ values are less than $F_{\alpha/2}$ and greater than $F_{(1-\alpha/2)}$ for each respective measurement, which does not allow the rejection of the null hypothesis, $\sigma^2{}_{\text{NIR}} = \sigma^2{}_{\text{HPLC}}$. Therefore, we have no basis to suspect that the NIR method is not equivalent to the HPLC method.

Principal Component Analysis

Principal Component Analysis (PCA) analysis is similar to PLS analysis in that a set of orthogonal vectors is generated to fit the data. The distinguishing characteristic of PLS analysis is the use of the parameters of interest to adjust the vectors. In PCA analysis, the vectors are generated to describe the successive orthogonal axes of variance. The parameter of interest, in this case % of claim, is not needed to determine these vectors. Typically, PCA is used to determine groupings of samples or to identify outliers. Figure 1 shows how the sample scores show the trend of samples away from 100 percent of target. The spectrum of an unknown sample is decomposed using PCA in order to find the unknown sample's position on the scores plot. The unknown samples scores will indicate if the unknown sample is adequately or inadequately coated. This method was used to determine which samples were to be analyzed by HPLC in order to reduce the number of samples to be analyzed. By using the scores, the widest possible range of coatings could be evaluated.

Several algorithms, such as Non-Linear Iterative Partial Least Squares (NIPLS) or Singular Value Decomposition (SVD) are employed to generate orthogonal vectors. SVD was the method of preference in this study because the algorithm was more robust. The data were imported into Matlab (The Mathworks, Nantucket, MA) and the SVD function provided in the software package was applied to the Multiplicative Scatter Corrected corrected data (see the next section).

Multiplicative Scatter Correction

The Multiplicative Scatter Correction (MSC) is a means whereby the effects of inhomogeneous particle scatter can be reduced in a NIR spectrum (8). The method is performed by determining the slope and y-intercept from the mean of the calibration data set. Let $x_1 \ldots x_k$ be the spectral values of a sample and $m_1 \ldots m_k$ be the mean of the calibration data set. The index k denotes the number of channels over which the data were collected. The slope and y-intercept are determined by performing a least squares fit of Equation 4:

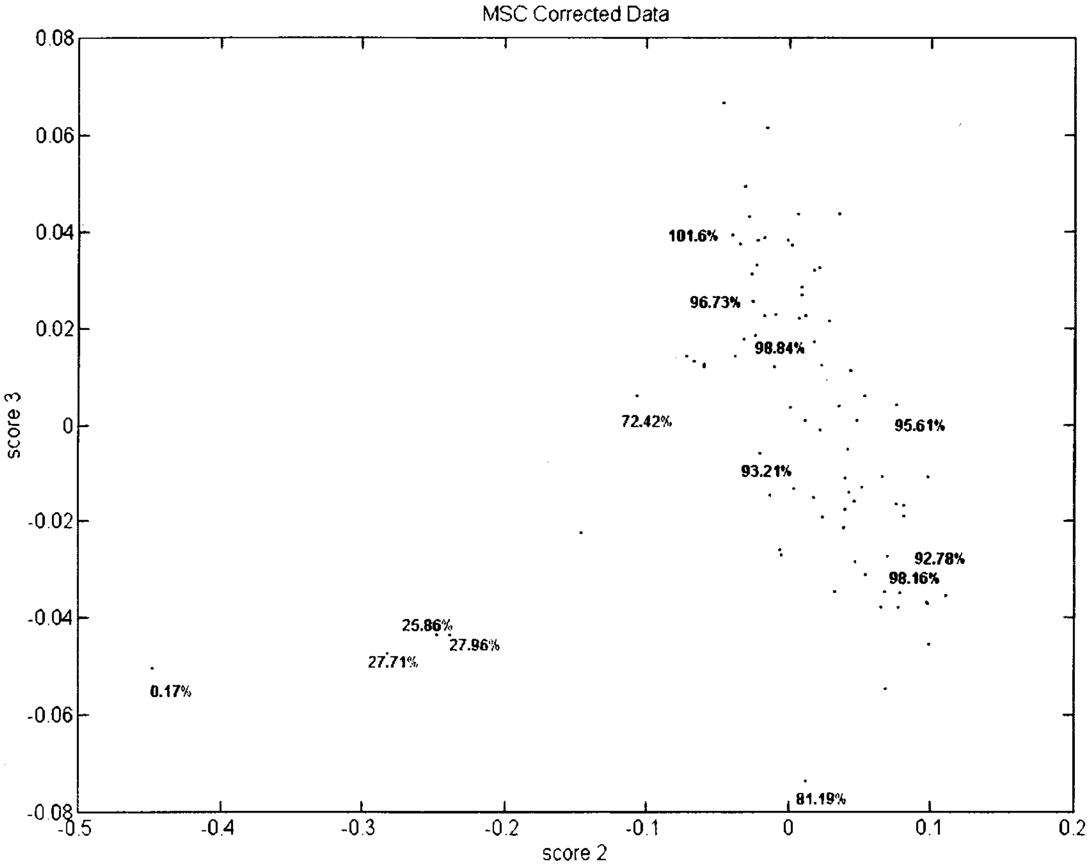

Fig. 1. Scores plot showing data trends. Multiplicative scatter correction (MSC) performed to minimize particle size effects.

$$x_k = b + a * m_k + e_k \qquad (4)$$

where b is the y-intercept, a is the slope, and e_k is the error. The corrected spectrum is obtained by subtracting the y-intercept and dividing by the slope, see Equation 5.

$$x_{k,\text{corrected}} = (x_{k,\text{original}} - b)/a \qquad (5)$$

The mean values of the calibration set are stored and used later to find b and a for the unknown spectra.

RESULTS AND DISCUSSIONS

Creation of the Calibration Equation

Forty-one samples were selected based on the NIR scores analysis, and were subsequently analyzed by HPLC. Next, they were randomally split into two data sets. We calculated the second derivative of each spectrum to correct for particle size effects and any specular component which might be present in our data. The second derivative spectra were generated using a segment of 20 nm and a gap size of 0. The segment width was chosen to be on the order of the bandwidth of the instrument. We created a quantitative regression curve based on the first data set (calibration set) which contained 27 samples (Table II). The second data set (verification set, Table III) contained 14 samples. Two regression curves were generated: an MLR equation and a 4-factor PLS model, which used the entire spectrum. The values reported, % of target, were calculated as follows:

$$100 \times \frac{(\text{Actual Weight of Active})}{(\text{Expected Weight of Active})} \qquad (6)$$

Ideal film coating is indicated by 100% of target.

Table II. Comparison of NIR and HPLC Results. Calibration Set

Tablet ID	NIR Result (PLS)	NIR Result (MLR)	HPLC Result (%)	Residual (PLS)	Residual (MLR)
3-2	91.58	90.40	91.14	0.44	−0.74
3-6	25.76	27.40	25.86	−0.10	1.54
3-10	92.80	93.88	87.35	5.45	6.53
3-13	98.17	95.02	95.85	2.32	−0.83
3-16	103.53	100.61	99.47	4.06	1.14
3-20	93.52	89.57	92.78	0.74	−3.21
3-24	26.07	27.77	27.71	−1.64	0.06
3-26	78.48	76.52	81.19	−2.71	−4.67
3-30	97.18	98.74	96.73	0.45	2.01
3-31	90.05	90.51	93.21	−3.16	−2.70
3-39	94.00	92.65	94.30	−0.30	−1.65
3-46	100.26	102.28	100.40	−0.14	1.88
3-47	95.16	91.99	95.61	−0.45	−3.63
3-55	1.01	1.63	0.17	0.84	1.56
3-62	98.72	101.14	101.60	−2.89	−0.46
3-72	73.97	71.60	72.42	1.55	−0.82
3-73	26.09	28.00	27.96	−1.87	0.04
3-77	97.84	101.43	98.84	−1.00	2.59
3-80	98.78	101.74	98.16	0.62	3.58
3-3C	39.04	36.16	37.86	1.18	−1.70
3-9C	14.39	14.22	14.70	−0.31	−0.48
3-13C	101.01	101.56	98.32	2.69	3.24
3-50C	28.74	28.37	29.14	−0.40	−0.77
1-23B	63.82	62.35	60.58	3.24	1.77
2-18B	94.74	92.85	100.26	−5.52	−7.40
3-18C	66.37	65.08	61.12	5.25	3.96
2-5B	89.36	96.95	97.84	−8.48	−0.89

Table III. Comparison of NIR and HPLC Results. Verification Set

Tablet ID	NIR Result (PLS)	NIR Result (MLR)	HPLC Result (%)	Residual (PLS)	Residual (MLR)
3-3	45.47	44.51	46.46	−0.99	−1.95
3-6C	0.88	−0.56	0.00	−0.88	−0.56
3-10C	97.74	100.97	97.34	0.40	3.63
3-31C	91.58	94.16	93.20	−1.62	0.96
3-47C	101.55	103.04	103.40	−1.85	−0.36
3-49C	21.39	22.63	21.52	−0.13	1.11
1-8B	78.73	78.90	75.36	3.37	3.54
1-9B	17.54	14.83	16.14	1.40	−1.31
1-12B	97.70	102.34	109.46	−11.76	−7.12
2-1B	95.25	95.93	95.76	−0.51	0.17
2-9B	92.26	86.17	94.46	−2.20	−8.29
2-11B	88.73	86.46	88.04	0.69	−1.59
2-19B	105.86	104.03	101.84	4.02	2.19
2-23B	93.25	87.83	92.92	0.33	−5.09

Multiple Linear Regression

The MLR model was determined via an automatic search of all available channels, i.e. wavelengths, in the spectra. The computer program identified 1680nm as the optimum wavelength. The resulting regression curve yielded a correlation coefficient (R) of −.9951 and an SEC value of 3.29. The negative R value indicates correlation to the drug active and not to an excipient or other component within the tablet matrix. The correlation coefficient and SEC were further improved by dividing through by another wavelength, 1552nm, which is an isosbestic point. The resultant calibration equation is given below in Equation 7:

$$\%\ \text{Drug Active} = -59.51 * (A_{\lambda=1680\text{nm}}/A_{\lambda=1552\text{nm}}) + 13.35 \qquad (7)$$

where A is the absorbance at the designated wavelengths, −59.51 is the coefficient in Equation 1, and 13.35 is the scalar offset in Equation 1. The SEC for this equation was determined to be 2.98 with a correlation coefficient of −.996 (see Figure 2).

Partial Least Squares

Determination to use four factors in the PLS equation was based upon the correlation coefficient value of .9956 and a minimum SEC value of 3.30 (see Figure 3 and Table IV). Visual inspection of plots of the loadings showed that all four factors were valid. The 4-term equation was determined not to be an overfit for the data being evaluated because the 6th eigenvector was observed to still contain meaningful information. The MLR and PLS equations predicted similar values for the verification set. The SEP of the MLR and PLS equations were 3.62 and 3.63 respectively, (see Table IV). These values are comparable to the SEC values obtained for the two regression curves (i.e., PLS = 3.30, MLR = 2.98). This indicates that the regression models are accurate and representative of the samples being analyzed.

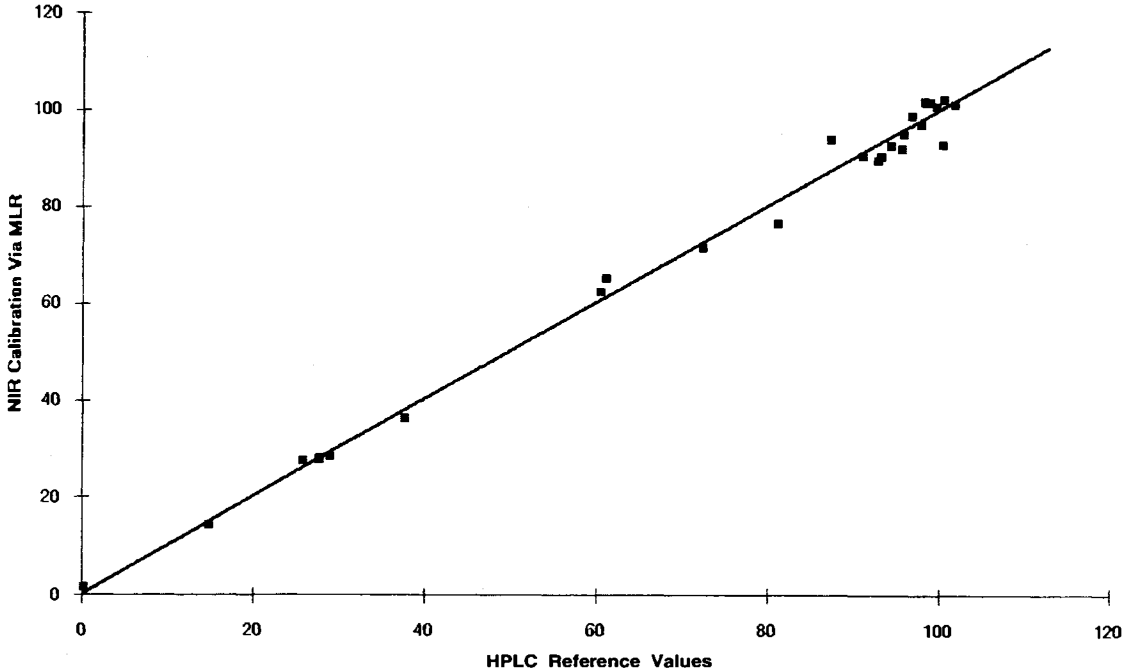

Fig. 2. MLR Calibration of coated tablets. NIR versus HPLC.

The SEP for the PLS equation is reduced to 1.88 if sample number 1–12B is removed from the verification set (see Table I). The SEP for the MLR equation also decreases, but only to a value of 3.30. The notable decrease observed for the PLS model implies that it is a better model for the data set being evaluated. This result supports the conclusion found in reference 1. The fact that the SEP value of 1.88 is significantly lower than the 2.98 SEC value, suggests that some outliers may exist in the calibration set, or that the data may not be adequately spread over the range of the calibration set. Ideally, the calibration set will include values that are evenly distributed over the expected theoretical range. The goal in tablet manufacturing is to obtain tablets as close to the 100% of target level as possible, while tablet film coatings in the 85–105% range are acceptable for this study. Most engineers resist adjusting parameters to purposefully manufacture unacceptable tablets because it is

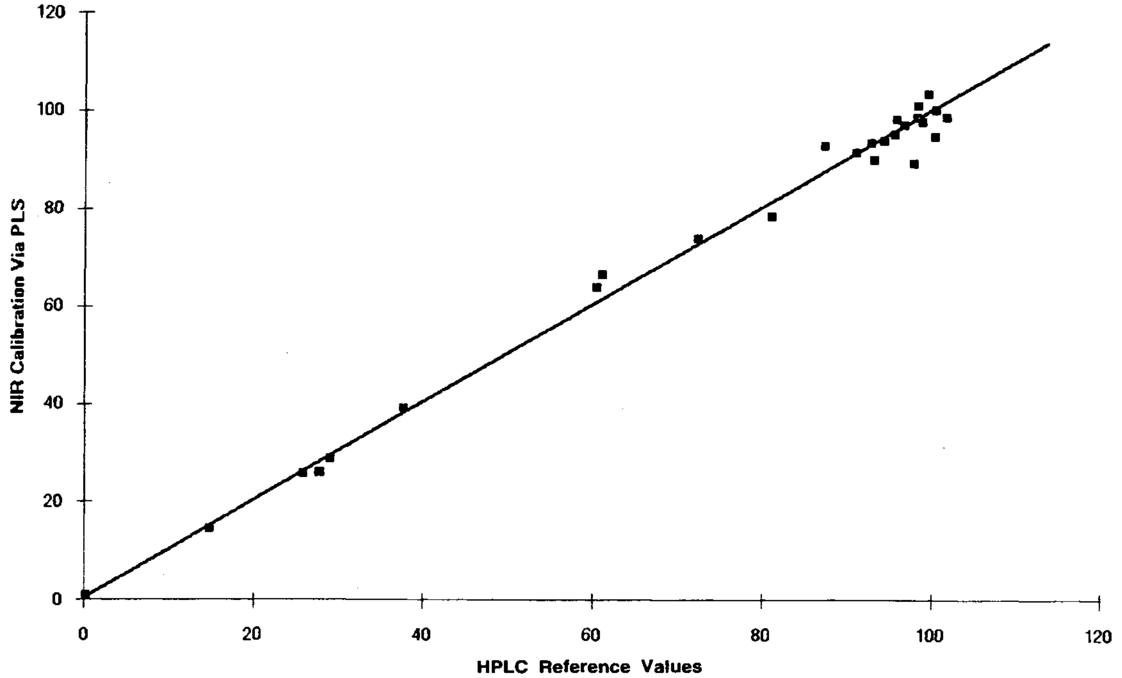

Fig. 3. PLS Calibration of coated tablets. NIR versus HPLC.

Table IV. Statistical Summary for Verification Set (All Data Used)

Parameter	MLR	PLS
Number of Samples	14	14
Bias	−1.05	−0.70
Root Mean Square(RMS)	3.79	3.74
SEP	3.62	3.63
Slope Adjustment	1.01	1.01
Intercept Adjustment	0.65	−0.27
Simple Correlation	0.995	0.995
Achievable SEP	3.06	3.19
Degrees of Freedom	11	9

costly and impractical. The constraint of not having undercoated or overcoated tablets, makes it difficult to develop as robust a model as is desired.

Justification for the omission of sample 1-12B from the verification set is based upon the HPLC reference value of 109%, which lies outside the range of the calibration set. Even though it is desirable, it is not reasonable to expect a model to predict values which are outside the calibration range. It is a caveat of regression analysis that the calibration equation or model is applicable only over the range of the calibration set (5). Therefore, it is quite reasonable to exclude a sample which falls outside the calibration set range from the verification set when calculating the verification statistics.

CONCLUSION

The accuracy and reliability of the 4-factor PLS model coupled with the non-invasive nature of NIR analysis made NIR the method of choice for the analysis of a drug active film-coated tablets. Models demonstrating greater ruggedness and accuracy might have been developed by creating several regression curves over smaller ranges, and/or by assaying more tablets and adding them to the equation. However, the level of accuracy needed for this experiment, ±4%, did not mandate further refinement of our equations. In all, over 400 samples were successfully analyzed within ±4% using NIR analysis. This allowed Merck's research personnel to evaluate the efficiency of a new drug film coating manufacturing process more quickly and inexpensively.

ACKNOWLEDGMENTS

The authors thank L. L. Fegley, M. C. Pignoli, and R. L. Clarin for performing HPLC assay, J. D. Kohutka and Bo Huang for assisting the NIR tablet scanning, and J. Timmermans and G. Catena for comments on the manuscript.

REFERENCES

1. E. Dreassi, G. Ceramelli, L. Savini, P. Corti, P. L. Perruccio, and S. Lonardi. Application of Near-infrared Reflectance Analysis to the Integrated Control of Antibiotic Tablet Production. *Analyst* **120**:319–323 1995.
2. M. A. Dempster, J. A. Jones, I. R. Last, B. F. McDonald, and K. A. Prebble. Near-infrared methods for the identification of tablets in clinical trail supplies. *J. Of Pharmaceutical & Biomedical Analysis* **11**:1087–1092 1993.
3. P. K. Aldridge, R. F. Mushinsky, M. M. Andino, and C. L. Evans. Identification of Tablet Formulations Inside Blister Packages by Near-Infrared Spectroscopy. *Applied Spectroscopy* **48**:1272–1276, 1994.
4. N. R. Draper and H. Smith. *Applied Regression Analysis,* John Wiley and Sons, New York, 1966.
5. J. B. Kennedy and A. M. Neville. *Basic Statistical Methods for Engineers and Scientists,* Harper and Row, New York, 1986.
6. M. A. Sharaf, D. Illman, and B. R. Kowalski. *Chemometrics,* John Wiley and Sons, New York 1986.
7. D. A. Burns and E. W. Ciurczak (eds.). *Handbook of Near-Infrared Analysis,* Marcel dekker, New York, vol. 13, 1992.
8. T. Isaksson and T. Næs. The Effect of Multiplicative Scatter Correction (MSC) and Linearity Improvement in NIR Spectroscopy. *Applied Spectroscopy* **42**:1273–1277 (1988).