

Near-infrared spectroscopic characterization of pharmaceutical powder blends

David J. Wargo, James K. Drennen*

Graduate School of Pharmaceutical Sciences, Division of Pharmaceutics, Duquesne University, Pittsburgh, PA 15282, USA

Received for review 19 June 1995; revised manuscript received 7 December 1995

Abstract

Near-infrared (near-IR) spectroscopy was used to qualitatively assess the homogeneity of a typical direct compression pharmaceutical powder blend consisting of hydrochlorothiazide, fast-flo lactose, croscarmellose sodium, and magnesium stearate. Near-IR diffuse reflectance spectra were collected from thieved powder samples using a grating-based spectrometer. A second-derivative calculation and principal component analysis were performed on the spectra prior to qualitative evaluation. Blend homogeneity was determined using single- and multiple-sample bootstrap algorithms and traditional chi-square analysis. The results suggested that bootstrap techniques provided greater sensitivity for assessing blend homogeneity than chi-square calculations and that near-IR has great potential as an analytical tool in powder blend analysis.

Keywords: Content uniformity; Homogeneity; Mixing; Near-infrared spectroscopy; Powder blend; Qualitative pharmaceutical analysis

1. Introduction

In recent years, the potential of near-infrared (near-IR) spectroscopy as a rapid non-destructive analytical technique for the pharmaceutical industry has been reported [1–3]. Qualitative applications of near-IR spectroscopy depend on pattern recognition methods to analyze multivariate data generated by the spectrometer. Pharma-

ceutical applications of qualitative near-IR spectroscopy include identity and quality testing of raw materials [4,5], detection of tablet and capsule tampering [6,7], detection of tablet degradation [8], analysis of parenteral products [9], and determination of ointment homogeneity [10]. The potential of near-IR in validating powder mixing processes has also been described [11]. Because near-IR allows the analysis of complex matrices to be performed rapidly, non-destructively, and without the use of organic solvents, it offers substantial advantages over traditional wet chemical techniques.

* Corresponding author. Tel.: (+1) 412-396-5662; Fax: (+1) 412-396-5130.

Preparation of a uniform powder blend prior to tableting or encapsulation is a vital step in the production of solid pharmaceutical dosage forms. The determination of powder blend homogeneity is typically a labor-intensive process involving the removal of unit-dose samples from defined mixer locations using a sample thief, extraction of the active drug from the sample matrix, and drug content analysis by either high performance liquid chromatography or UV spectroscopy. The distribution of individual excipients is typically assumed to be homogeneous if the active ingredient is uniformly distributed.

Because most pharmaceutical active ingredients and excipients absorb near-IR radiation, studies utilizing near-IR may complement the assay for the active ingredient by providing homogeneity information regarding all mixture components. Additionally, a direct non-destructive method for assessing powder blend homogeneity could be of great value in minimizing the sample preparation and assay time associated with traditional blend analysis procedures.

This paper examines the application of near-IR in assessing the homogeneity of typical direct compression pharmaceutical powder blends consisting of hydrochlorothiazide, fast-flo lactose, croscarmellose sodium, and magnesium stearate. Blend homogeneity and optimal mixing times have been qualitatively determined using single- and multiple-sample bootstrap algorithms and traditional chi-square analysis.

2. Experimental

2.1. Materials

Hydrochlorothiazide (Gyma Laboratories, Westbury, NY), Fastflo lactose (Foremost, Baraboo, WI), croscarmellose sodium (Ac-Di-Sol, FMC, Philadelphia, PA), and magnesium stearate (Whitaker, Clark, and Daniels, South Plainfield, NJ) were donated and used as received. All bulk powders were passed through a #30 mesh screen prior to use. Methanol (VWR Scientific, West Chester, PA) was spectra grade.

2.2. Mixing

All mixing studies were carried out in a stainless-steel 8-quart twin-shell blender (Patterson-Kelly Co., Inc., Stroudsburg, PA). The multi-component powder system consisted of fast-flo lactose (80%), hydrochlorothiazide (15%), croscarmellose sodium (4%), and magnesium stearate (1%). Screened powders were loaded into the blender in the order listed above. The blender was filled to 90% of working capacity (v/v) and rotated at 25 rev min⁻¹ throughout the study.

Initially, five blends of identical composition were subjected to different mixing times of 1, 5, 10, 15, and 20 min. At the specified time, the blender was stopped and samples equivalent to 1–3 dosage units, 200–600 mg, were removed from ten different powder bed locations using a sample thief. In the second part of this study, a single formulation was mixed for 30 min during which time six unit-dose samples were thieved from the blender at 2 min time intervals. In both experiments, powder samples were transferred to tared borosilicate sample vials and the sample weights were recorded for reference determination of drug concentration.

2.3. Near-IR methodology

Near-IR spectra for each powder sample were collected in triplicate using a Quantum 1200 Plus grating-based spectrometer (LT Industries, Rockville, MD) by scanning directly through the base of the sample vials. Spectral processing and chemometric analysis were performed using proprietary software written in SPEAKEASY IV EPSILON+® (Speakeasy Computing Corp., Chicago, IL).

Near-IR spectra were collected by scanning from 1200 to 2400 nm in the reflectance mode. Reflectance values were linearized with a log (1/R) transformation. Individual sample vials were rotated 120° between triplicate scans to ensure representative spectra. Triplicate scans were averaged to obtain one spectrum for each powder sample. To remove shifts in spectral baselines caused by particle size and packing differences, a second derivative calculation was used.

A principal component analysis (PCA) [12] of the second derivative spectra was performed. PCA was useful as a preprocessing treatment before application of the bootstrap algorithms because it simplified the spectra by removing the wavelengths that added only noise information. Spectra originally recorded as 1201 wavelengths were expressed through PCA as points in four-dimensional space.

2.3.1. Bootstrap error-adjusted single-sample technique (BEST)

The BEST represents a type of analytical procedure designed to operate in the high-speed parallel or vector mode required of pattern-recognition tests involving thousands of samples. Lodder and Hieftje [13] have discussed this technique, derived from Efron's bootstrap calculation [14], in detail and have provided examples of its application. The BEST can be used to provide both quantitative and qualitative analyses of intact products. The BEST begins by treating each wavelength in a spectrum as a single point in multidimensional space ("hyperspace"). Each point is translated from the origin along each axis by an amount that corresponds to the magnitude of the signal observed at each wavelength. Samples with similar spectra map into clusters of points in similar regions of hyperspace, with larger cluster size corresponding to samples with greater intrinsic variability.

The BEST develops an estimate of the total sample population using a small set of known samples. A point estimate of the center of this known population is also calculated. When a new sample is analyzed, its spectrum is projected into the same hyperspace as the known samples. A vector is then formed in hyperspace to connect the center of the population estimate to the new sample spectral point. A hypercylinder is formed about this vector to contain a number of estimated-population spectral points. The density of these points in both directions along the central axis of the hypercylinder is used to construct an asymmetric nonparametric confidence interval. The use of a central 68% confidence interval produces BEST distances analogous to standard deviations.

BEST distances are used to identify sample constituents. Uncorrected BEST distances (suitable for unskewed training sets) are calculated as follows:

$$\left(\sum_{j=1}^d (c_j - x_j)^2 \right)^{1/2} / \sigma$$

where c_j is the center of the bootstrap distribution, x_j is the test sample spectrum and σ is a BEST standard deviation. When a sample spectrum projects to a point within three standard deviations of the center of a cluster of spectral points from a known substance or product, the sample is considered to be a sample of the known material. The known product is either a pure substance or a mixture of components. When the new sample contains different substances or components in concentrations that differ from the known product, the new sample spectral point is displaced from the known spectral cluster. The magnitude of this displacement increases as the difference between the new sample and the set of known samples increases. Furthermore, the direction of the displacement of the new sample point corresponds to the spectra of the constituents responsible for the displacement.

2.3.2. Modified bootstrap technique (modified BEST)

Where the single-sample BEST algorithm provides for the qualitative analysis of a single test sample, the modified BEST algorithm provides a test that uses multiple test spectra to detect false samples (as subclusters) well within the three standard deviation (SD) limit of a training set. The accurate detection of subclusters allows the determination of very small changes in component concentration. The modified BEST technique is described in detail in Ref. [15].

The first steps of the modified BEST test are the same as in the single-sample BEST. A training set is constructed from known samples and a Monte-Carlo approximation to the bootstrap distribution is calculated to estimate the population from which the training set is drawn. The center of the cluster represents the best estimate of the spectrum of the compound. The

modified BEST varies from the single sample BEST by next projecting the bootstrap estimated test population into the same multidimensional space as the training set and calculating cumulative distribution functions (CDFs) for the estimated populations (training set and test set). As the number of observations (bootstrap replicates) increases, the rough empirical cumulative distribution function (ECDF) approaches the smooth theoretical cumulative distribution function (TCDF).

Plotting the ECDF vs. the TCDF, for a given probability, generates the linear version of a cumulative distribution function, a quantile–quantile (QQ) plot (see Figs. 9 and 10). Each cumulative probability value yields a pair of order statistics (one from each CDF) that form a point in the QQ plot. QQ plots are valuable tools for distinguishing differences in shape, size, and location between spectral clusters. Two similar clusters of spectra will demonstrate a linear QQ plot. Breaks and/or curves in the QQ plot indicate that differences exist between the groups.

Spectral similarity is based on the linearity of the QQ plots generated. A 98% confidence limit for the correlation coefficient of the training set is calculated. QQ plots with correlation coefficients less than the calculated confidence limit are considered to be spectrally different from the training group, while QQ plots with correlation coefficients greater than the confidence limit are considered to be spectrally similar.

2.3.3. Chi-squared analysis

The final method utilized for determination of optimal blending times employed traditional chi-square analysis [16] to assess near-IR spectral variability. For each time point, the pooled variance of the near-IR absorbance values at individual wavelengths is calculated as the weighted average of the variances, where the weights are the degrees of freedom. A chi-square statistic is then calculated and compared to a tabulated value for significance at the 5% level. A significant value for the chi-square statistic indicates that the variances are not equal.

2.4. Hydrochlorothiazide assay

The powder samples were dissolved and extracted in methanol–water (50:50, v/v). Residual excipients were removed by filtration through a 0.45 μm pore diameter filter (Millipore, Marlborough, MA). The first 10 ml of the filtered solution was discarded to account for any adsorption of the solute to the membrane filter. Samples were appropriately diluted prior to analysis. Hydrochlorothiazide concentrations were determined on a Lambda 2S UV–Vis spectrometer (Perkin-Elmer Corp., Norwalk, CT) at 271 nm.

Calibration curves were prepared for the hydrochlorothiazide solutions. A linear relationship was obtained over the range 5–20 $\mu\text{g ml}^{-1}$. Method reproducibility was evaluated by analyzing 25 samples from the same stock solution prepared from accurately weighed amounts of hydrochlorothiazide and excipients in the same proportion as used in the mixing studies.

3. Results and discussion

3.1. Multiple blend study

Initially, five blends of identical composition were each subjected to different mixing times of

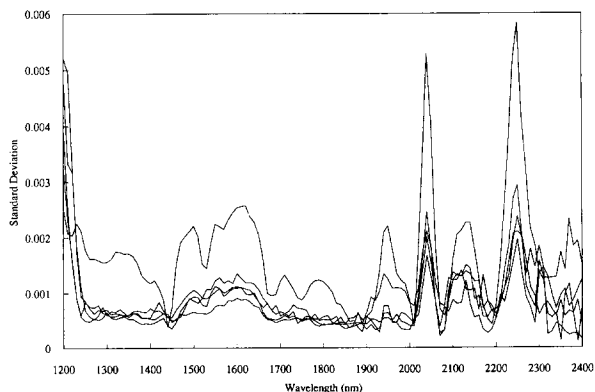


Fig. 1. Near-IR absorbance standard deviations ($n = 10$) for 1, 5, 10, 15, and 20 min powder blends. Top to bottom at 2030 nm (hydrochlorothiazide) and 2240 nm (lactose): 1, 5, 10, 15, 20 min.

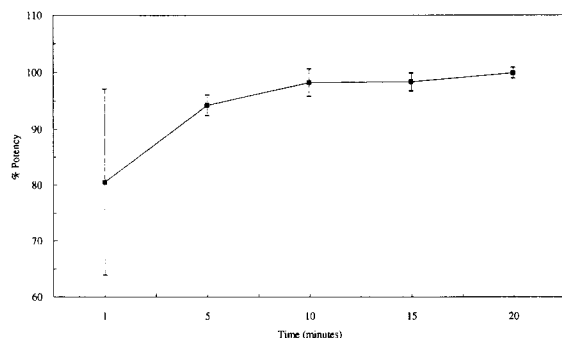


Fig. 2. UV reference method potency results ($n = 10$) for 1, 5, 10, 15, and 20 min powder blends. Bars indicate percent relative standard deviations.

1, 5, 10, 15, and 20 min. Characteristic hydrochlorothiazide absorbance occurs primarily at 2030 nm while lactose demonstrates maximal absorbance at 2240 nm. The effect of blend time on homogeneity is evident in Fig. 1 which depicts the sample absorbance standard deviations versus near-IR wavelength for each batch. Again, the most notable regions of spectral variation occur at wavelengths characteristic of hydrochlorothiazide and lactose, 2030 nm and 2240 nm respectively. Qualitative bootstrap single- and multiple-sample algorithms were employed to systematically determine the time at which the blend had reached a desired degree of mixing.

Fig. 2 represents the UV reference method potency results for the 1, 5, 10, 15, and 20 min blend samples. Because the 20 min blend demonstrated the most desirable potency and uniformity results (99.9% potency, 0.99% RSD) according to the reference method, the samples from this time point were used as the training group for the bootstrap calculations. Using the BEST algorithm, single-sample spectra from the 1, 5, 10, and 15 min blends were tested against the 20 min training group. To validate the model, individual samples at the 20 min time point were tested against remaining samples via a leave-one-out cross-validation method.

Fig. 3 illustrates the results obtained from the BEST calculation. The method identified the fact that only samples from the cross-validation study (20 min time point) were similar to the

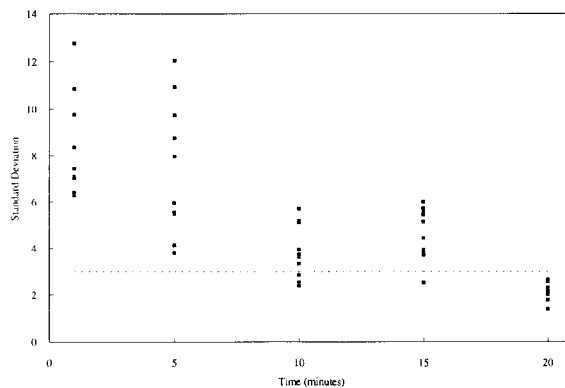


Fig. 3. Bootstrap standard deviations calculated from the BEST calculation using the 20 min samples as the training set. The broken line represents a 3SD limit for spectral similarity.

training samples. This was attributed to the fact that although the 10 and 15 min samples had similar %RSD when compared to the 20 min training group, hydrochlorothiazide assay values for both test groups differed from the training group by approximately 2% (98.2% and 98.4% of theoretical concentration). To evaluate this hypothesis, bootstrap calculations were then performed using the 15 min samples as the training group with the results of this study depicted in Fig. 4. The method distinguished a majority of the samples from the 1, 5, and 20 min blends from the training population and identified 70% of the 10 min blend samples as being spectrally similar to the 15 min training population, thus

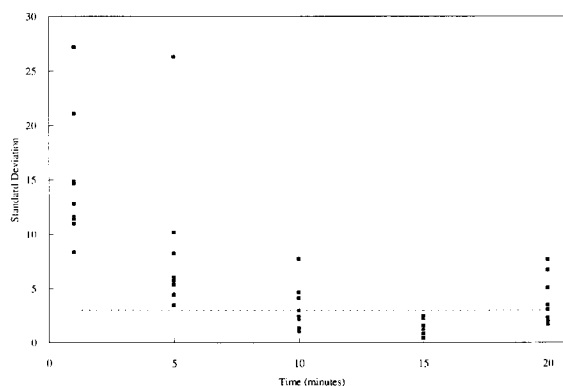


Fig. 4. Bootstrap standard deviations from the BEST calculation using the 15 min samples as the training set. The broken line represents a 3SD limit for spectral similarity.

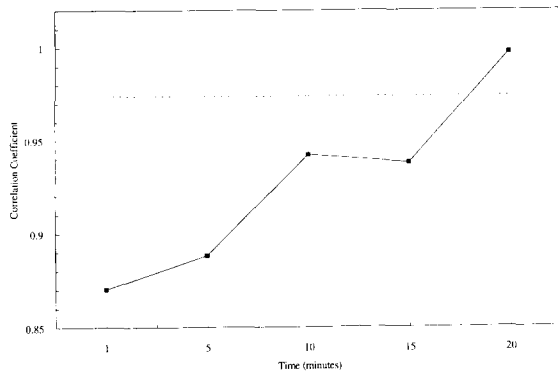


Fig. 5. The effect of blend time on the correlation coefficient calculated from a modified BEST QQ plot. The broken line represents a 98% confidence limit on the 20 min training set.

confirming the sensitivity of this method in detecting minor changes in sample component concentration.

Where the single-sample BEST provides for the qualitative analysis of a single test sample, the modified BEST algorithm uses multiple test samples to detect spectral variations well within the three standard deviation limit of a training set. The detection of subclusters by this method allows for the determination of very small changes in component concentrations.

Fig. 5 represents the results of the modified BEST calculations. The 1, 5, 10, and 15 min spectral groups were tested against the spectra from the 20 min training set. Spectral similarity was based on the linearity of the QQ plots generated. A 98% confidence limit for the correlation coefficient, represented by the broken line in Fig. 5, was calculated using the 20 min training set. QQ plots with correlation coefficients less than the calculated confidence limit were considered to be spectrally different from the training group, while QQ plots with correlation coefficients greater than the confidence limit were considered spectrally similar. The modified BEST algorithm recognized all spectral test groups to be significantly different from the 20 min training group. The results can be attributed to differences in the hydrochlorothiazide concentration of the samples. As a pattern recognition method, the modified BEST algorithm is sensitive to variations in cluster size, shape and location.

3.2. Evaluation of optimal mixing time for a single blend

Having determined that minor differences in powder blend component concentration and content uniformity for different blends could be detected by near-IR spectroscopy, the second part of the preliminary experiments focused on evaluating an optimal mixing time for a single powder blend formulation. A single formulation of the identical composition described earlier was blended for a total of 30 min, during which time six unit dose samples were thieved from the blender at 2 min intervals. The samples were analyzed in the manner previously described. The single-sample BEST and the multiple-sample modified BEST algorithms as well as traditional chi-square analysis for sample variability were employed to determine when the blend had reached homogeneity.

Fig. 6 represents the UV reference method potency results ($n = 6$) for the 30 min blend study. UV results demonstrate that at the 10 min time point the blend has achieved acceptable potency (99.1%) although only marginal uniformity (RSD 2.4%). At the 16 and 18 min time points, both potency (100.0%) and uniformity (0.9% RSD) reach optimal values.

Studies conducted during this experiment indicated that a minimum number of ten sample spectra are necessary to achieve acceptable bootstrap calculation results. Because of the sample number requirement and the optimal reference potency and uniformity results demonstrated by the 16 and 18 min samples, these two sample

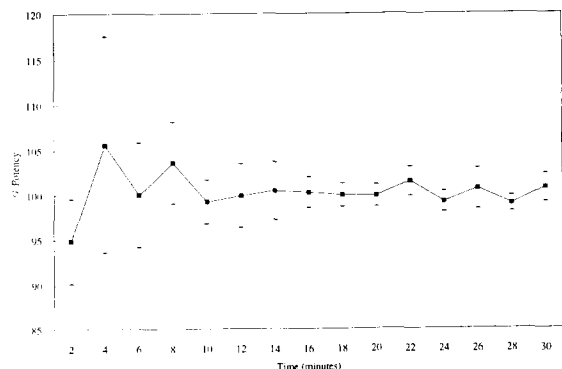


Fig. 6. UV reference method potency results ($n = 6$) for the single blend study. Bars indicate percent relative standard deviations.

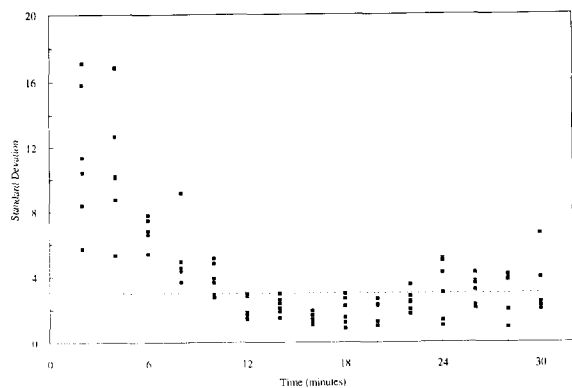


Fig. 7. Bootstrap standard deviations from the BEST calculation using the 16 and 18 min samples as the training set. The broken line represents a 3SD limit for spectral similarity.

groups were combined ($n = 12$) and used as the training set for the bootstrap calculations.

Figs. 7 and 8 depict the BEST and modified BEST results respectively. The BEST identifies that the blend reaches homogeneity at the 12 min time point, whereas the modified BEST concludes that blend uniformity is achieved at the 10 min time point.

Examples of linear and nonlinear QQ plots generated from the modified BEST are portrayed in Figs. 9 and 10. Fig. 9 depicts the QQ plot of the 8 and 10 min samples and Fig. 10 represents the QQ plot of the 14 and 16 min samples tested against the 16 and 18 min training group. It becomes apparent that as sample groups become more similar to the

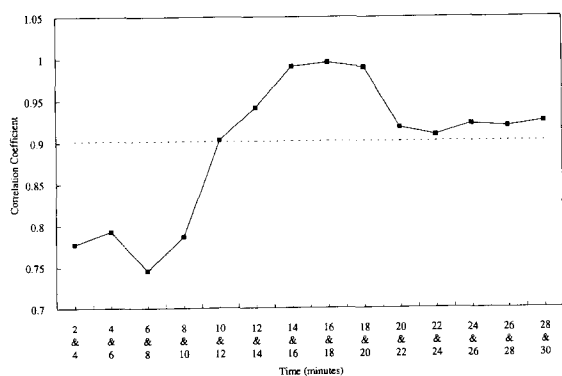


Fig. 8. The effect of blend time on the correlation coefficient calculated from a modified BEST QQ plot. The broken line represents a 98% confidence limit on the 16 and 18 min training set.

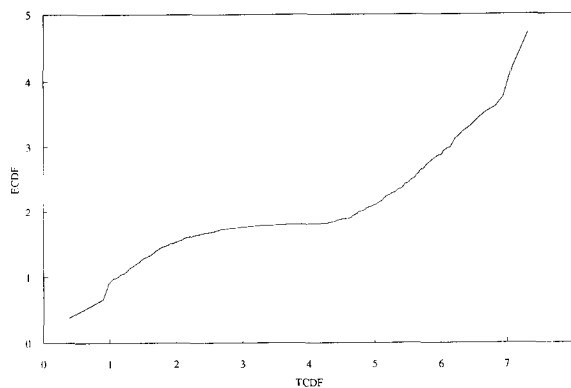


Fig. 9. A QQ plot from the modified BEST calculation demonstrating spectral dissimilarity between the training and test groups. The 8 and 10 min samples were tested against the 16 and 18 min training set.

training set, the QQ plots become more linear.

Although both reference and near-IR methods indicate that the blend becomes homogeneous at approximately 10–12 min, Figs. 6–8 reveal an apparent trend towards increased variability in blend uniformity as mixing continues beyond 20 min. After 20 min of blending, assay values appear to become variable, single-sample BEST distances increase, and the correlation values for the modified BEST calculation level off just above the confidence limit. A principal component (PC) plot of the spectral data (PC 2 vs. PC 4) indicates that a consistent shift occurs along the fourth principal

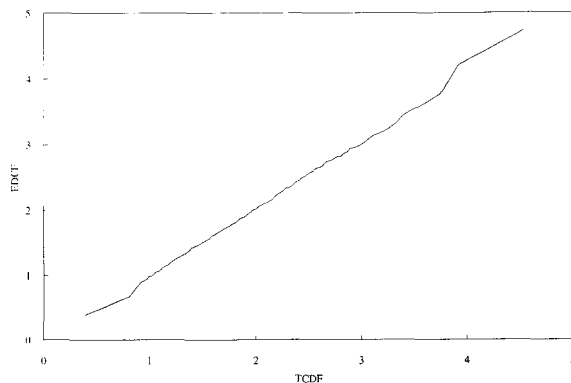


Fig. 10. A QQ plot from the modified BEST calculation demonstrating spectral similarity between the training and test groups. The 14 and 16 min samples were tested against the 16 and 18 min training set.

component axis during the entire mixing process. However, a one-way analysis of variance of the reference data (Fig. 6) combined with a Scheffe's multiple range comparison test ($\alpha = 5\%$) indicates that there are no significant differences between assay values from 10 min to 30 min.

In an attempt to identify the physical or chemical phenomenon responsible for spectral changes beyond the 20 min time point, loadings spectra for each of the first four PCs used in this model were plotted. The loadings spectra display the coefficient or weighting given to each wavelength in a particular PC array. Often, a loadings spectrum will exhibit similarity to the spectrum of an individual component, indicating that the PC in question will describe changes in the concentration or nature of that component. In this study, the loadings spectrum of PC 4 demonstrated features similar to that of a spectrum from pure magnesium stearate. The potential for decreased powder blend uniformity and related near-IR spectral variability following extended blending of formulations containing magnesium stearate appears to be possible in the light of the results reported by the following investigators.

Shah and Mlodozeniec [17], in their study of surface lubrication phenomena, suggested that during the mixing process, lubricant particles such as magnesium stearate first adsorb onto the surface of individual powder particles or granules, then, as mixing continues, distribute more uniformly upon the granule surface following delamination or deagglomeration mechanisms. By affecting the surface characteristics of the powder particles, the magnesium stearate may alter the flow properties of the material and affect the apparent bulk volume of the blended material. Furthermore, Murphy and Samyn [18] observed that the drug dissolution profiles for lactose–magnesium stearate compacts were related to the degree of shear applied during the mixing process. Longer periods of shear resulted in extended dissolution times. The authors also noted that powder blend bulk and tapped densities increased as lubricant blend times increased. The increased density was attributed to the improved flow properties of the blend. Although evidence suggests that the spectral variations apparent beyond 20 min in this

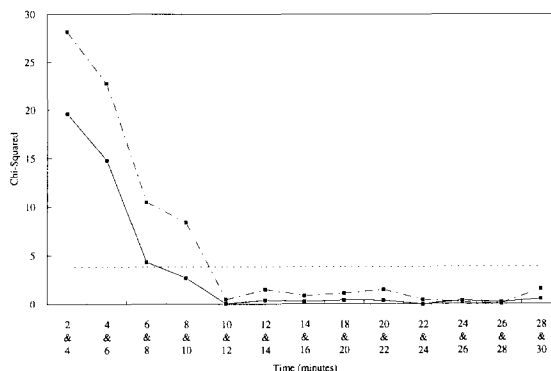


Fig. 11. Chi-square analysis results ($n = 12$) from the single blend study using the 16 and 18 min samples as the reference spectra: solid line = 2030 nm, corresponding to hydrochlorothiazide; chain line = 2240 nm, corresponding to lactose; dashed line = 5% significance level for the chi-squared statistic. Points below the dashed line demonstrate variance similar to the reference set.

study are related to the effects of magnesium stearate, further studies are necessary before this can be established with greater certainty.

Figs. 11 and 12 depict the results of the chi-square analysis. Absorbance values at 2030 nm and 2240 nm were chosen for the study because they correspond to wavelengths characteristic of hydrochlorothiazide and lactose respectively. In the 30 min study, the 16 and 18 min samples were chosen as the reference spectra. The chi-square test

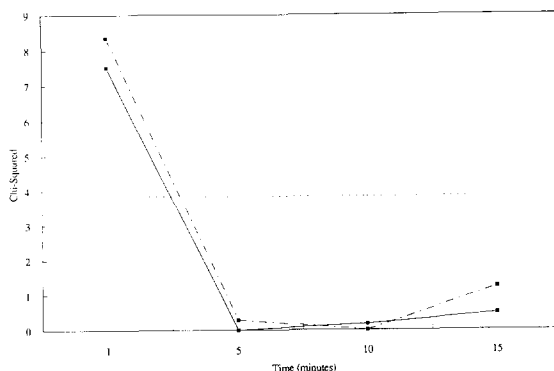


Fig. 12. Chi-square analysis results ($n = 10$) from the multiple blend study using the 20 min samples as the reference spectra: solid line = 2030 nm, corresponding to hydrochlorothiazide; chain line = 2240 nm, corresponding to lactose; dashed line = 5% level for the chi-squared statistic. Points below the dashed line demonstrate variance similar to the reference set.

demonstrates that at the 10 min time point and beyond, the variance of the thieved samples did not significantly differ from later samples. These results corresponded well to the results obtained from the bootstrap calculations.

Interestingly, when the initial 20 min blend study was analyzed using the chi-square technique, the method indicated that an homogeneous blend had been reached at the 5 min time point. An evaluation of the reference method results demonstrates that although the %RSDs for the samples at each time point were similar (1.81, 2.39, 1.58, and 0.944%) the potencies of the 5, 10, 15, and 20 min samples (94.2, 98.2, 98.4, and 99.9% respectively) were variable. From these results it is evident that one disadvantage of using the chi-square method to determine blend homogeneity is that it assesses only the variability between sample absorbances, and is not sensitive to differences in component concentrations.

4. Conclusions

The experiments conducted in this study indicate that near-IR spectroscopy has great potential as an analytical tool for blend uniformity analysis. Qualitative near-IR analysis can be employed to assess the uniformity of a single production blend or to define optimal mixing times during the development process. The near-IR spectrum of a powder sample can be compared with spectra from different blender locations, different blend times or from a spectral library.

These experiments have demonstrated that both the bootstrap calculations and the chi-square test for equality of variance are effective methods for qualitative evaluation of powder blend homogeneity. The chi-square technique is limited to the analysis of only individual wavelengths and assesses sample variability without regard to component concentration. However, the bootstrap techniques provide additional sensitivity because they utilize the entire near-IR spectrum and allow

recognition of variations in both component concentration and content uniformity. With regard to the bootstrap techniques, the modified BEST algorithm was more sensitive than the single-sample BEST in recognizing the minor spectral variations that occur in well-mixed systems.

Acknowledgements

This work has been supported by a predoctoral fellowship from Merck and Co.

References

- [1] E.W. Ciurczak, *App. Spectrosc. Rev.* 23 (1987) 147–163.
- [2] W. Plugge and C. Van der Vlies, *J. Pharm. Biomed. Anal.*, 10 (1992) 797–803.
- [3] B.F. MacDonald and K.A. Prebble, *J. Pharm. Biomed. Anal.*, 11 (1993) 1077–1085.
- [4] W. Plugge and C. Van der Vlies, *J. Pharm. Biomed. Anal.*, 11 (1993) 435–442.
- [5] P. Corti, E. Dreassi, G. Ceramelli, S. Lonardi, R. Viviani and S. Gravina, *Analisis*, 19 (1991) 198–204.
- [6] R.A. Lodder and G.M. Hieftje, *App. Spectrosc.*, 42 (1988) 556–558.
- [7] R.A. Lodder, M. Selby and G.M. Hieftje, *Anal. Chem.*, 59 (1987) 1921–1930.
- [8] J.K. Drennen and R.A. Lodder, *J. Pharm. Sci.*, 79 (1990) 622–627.
- [9] L. Galante, M. Brinkley, J. Drennen and R. Lodder, *Anal. Chem.*, 62 (1990) 2514–2521.
- [10] J.K. Drennen and R.A. Lodder, in G. Patony (Ed.), *Advances in Near-Infrared Measurements*, Vol. 1 JAI Press, Greenwich, CT, 1993, pp. 93–112.
- [11] E.W. Ciurczak, *Pharm. Tech.*, 15(9) (1991) 140–145.
- [12] B. Manly, *Multivariate Statistics: A Primer*, Chapman and Hall, London, 1986, p. 59.
- [13] R.A. Lodder and G.M. Hieftje, *Appl. Spectrosc.*, 42 (1988) 1351–1365.
- [14] B. Efron and R. Tibshirani, *Science*, 253 (1991) 390–295.
- [15] R.A. Lodder and G.M. Hieftje, *Appl. Spectrosc.*, 42 (1988) 1500–1512.
- [16] S. Bolton, *Pharmaceutical Statistics*, M. Dekker, New York, 1990, pp. 179–180.
- [17] A.C. Shah and A.R. Mlodozieniec, *J. Pharm. Sci.*, 66 (1977) 1377–1382.
- [18] K.S. Murthy and J.C. Samyn, *J. Pharm. Sci.*, 66 (1977) 1215–1219.