2006

CHEMOMETRICS, SPECTROMETRY, AND SENSORS FOR INTEGRATED SENSING AND PROCESSING: ADVANCING PROCESS ANALYTICAL TECHNOLOGY

Joseph Peter Medendorp
University of Kentucky, jpmede2@uky.edu

ABSTRACT OF DISSERTATION

Joseph Peter Medendorp

The Graduate School

University of Kentucky

2006

CHEMOMETRICS, SPECTROMETRY, AND SENSORS FOR INTEGRATED SENSING AND PROCESSING: ADVANCING PROCESS ANALYTICAL TECHNOLOGY

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By
Joseph Peter Medendorp

Lexington, Kentucky

Director: Dr. Robert A. Lodder, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2006

ABSTRACT OF DISSERTATION

CHEMOMETRICS, SPECTROMETRY, AND SENSORS FOR INTEGRATED SENSING AND PROCESSING: ADVANCING PROCESS ANALYTICAL TECHNOLOGY

The research contained in the following dissertation spans a diverse range of scientific scholarship, including; chemometrics for integrated sensing and processing (ISP), near infrared and acoustic resonance spectrometry for analyte quantification and classification, and an ISP acoustic sensor as an alternative to conventional acoustic spectrometry. These topics may at first seem disjointed; however, closer inspection reveals that chemometrics, spectrometry, and sensors taken together form the umbrella under which applied spectrometry and analytical chemistry fall. The inclusion of each of these three serves to paint the complete portrait of the role of applied spectrometry for the advancement of process analytical technology. To illustrate the totality of this portrait, this research seeks to introduce and substantiate three key claims. (1) When applicable, optical spectrometry and acoustic spectrometry are preferred alternatives to slower and more invasive methods of analysis. (2) Chemometrics can be implemented directly into the physical design of spectrometers, thus sparing the need for computationally demanding post-collection multivariate analyses. (3) Using this principle, ISP sensors can be developed specifically for use in highly applied situations, making possible automatic analyte quantification or classification without the computational burden and extensive data analysis typically associated with conventional spectrometry. More concisely, these three claims can be stated as follows: spectrometry has a broad range of uses, chemometrics for ISP makes spectrometry more efficient, and for all analytical problems with a spectrometric solution, an ISP sensor, specifically tailored to the needs of the experiment, can more effectively solve the same analytical problem.

KEYWORDS: chemometrics, analytical, sensors, near-infrared, acoustics

Multimedia Elements Used: JPEG (.jpg); Bitmap (.bmp); TIF (.tif); GIF (.gif).

Joseph Medendorp

August 1, 2006

CHEMOMETRICS, SPECTROMETRY, AND SENSORS FOR INTEGRATED SENSING AND PROCESSING: ADVANCING PROCESS ANALYTICAL TECHNOLOGY

By

Joseph Peter Medendorp

Robert A Lodder
Director of Dissertation

James Pauly
Director of Graduate Studies

August 1, 2006

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

DISSERTATION

Joseph Peter Medendorp

The Graduate School

University of Kentucky

2006

CHEMOMETRICS, SPECTROMETRY, AND SENSORS FOR INTEGRATED SENSING AND PROCESSING: ADVANCING PROCESS ANALYTICAL TECHNOLOGY

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By
Joseph Peter Medendorp

Lexington, Kentucky

Director: Dr. Robert A. Lodder, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2006

For Clare

Acknowledgements

I would like to express my warmest and highest gratitude to all those who have generously offered their support, guidance, and patience. To my family, I appreciate you more than words can express. Special thanks to my advisor, from whom I have learned that science knows no bounds but those imposed by lack of imagination.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

API	Active pharmaceutical ingredient
ARS	Acoustic resonance spectrometry
ATR	Attenuated total reflectance
BEST	Bootstrap error-adjusted single sample technique
CCD	Charge coupled device
CO	Colloidal oatmeal
CV	Cross validation
C10	Sodium caprate
4-CP	4-cyanophenol
DARPA	Defense Advance Research Projects Agency
EN	Econazole nitrate
EST	Estradiol
FDA	Food and Drug Administration
FG	Function generator
FFT	Fast Fourier transform
FM	Frequency modulation
FT	Fourier transform
FT-IR	Fourier transform infrared
GMP	Good manufacturing practice
HICI	Hyperspectral integrated computational imaging
HPLC	High performance liquid chromatography
IR	Infrared
ISP	Integrated sensing and processing
KPCA	Kernel principal component analysis
LOO	Leave one out
MANOVA	Multiple analysis of variance
MCC	Microcrystalline cellulose
MFC	Molecular factor computing
MOE	Multivariate optical element
MPD	MatTek permeation device
MSD	Multivariate standard deviation
NAS	Net analyte signal
NIR	Near infrared
OSP	Optical signal processing
PAT	Process analytical technology
PCA	Principal component analysis
PCR	Principal component regression
PZT	Piezoelectric transducer
RMSECV	Root mean square error of cross validation
RMSEP	Root mean square error of performance
RSD	Relative standard deviation
SD	Standard deviation
SEC	Standard error of calibration
SEE	Standard error of estimation

SEP	Standard error of performance
SN	Signal to noise
SRB	Solid rocket booster
SVM	Support vector machine
UVE	Uninformative variable elimination

List of Files

Figure #	File name	Size
Figure 1.1,	Figure1.1.bmp	1,573,254 bytes
Figure 1.2,	Figure1.2.jpg	26,697 bytes
Figure 1.3,	Figure1.3.bmp	464,466 bytes
Figure 1.4,	Figure1.4.bmp	751,158 bytes
Figure 1.5,	Figure1.5.bmp	624,694 bytes
Figure 1.6,	Figure1.6.bmp	1,387,962 bytes
Figure 1.7,	Figure1.7.bmp	402,898 bytes
Figure 2.1,	Figure2.1.bmp	1,198,958 bytes
Figure 2.2,	Figure2.2.gif	157,073 bytes
Figure 2.3,	Figure2.3.bmp	894,518 bytes
Figure 2.4,	Figure2.4.bmp	2,251,962 bytes
Figure 2.5,	Figure2.5.bmp	891,958 bytes
Figure 2.6,	Figure2.6.bmp	894,518 bytes
Figure 3.1,	Figure3.1.bmp	1,271,954 bytes
Figure 3.2,	Figure3.2.tif	8,482,940 bytes
Figure 3.3,	Figure3.3.tif	493,994 bytes
Figure 3.4,	Figure3.4.tif	955,766 bytes
Figure 4.1,	Figure4.1.bmp	894,518 bytes
Figure 4.2,	Figure4.2.bmp	891,958 bytes
Figure 4.3,	Figure4.3.bmp	891,958 bytes
Figure 4.4,	Figure4.4.bmp	2,660,710 bytes
Figure 5.1,	Figure5.1.bmp	336,886 bytes
Figure 5.2,	Figure5.2.bmp	3,272,246 bytes
Figure 5.3,	Figure5.3.bmp	26,448,054 bytes
Figure 5.4,	Figure5.4.bmp	11,539,234 bytes
Figure 5.5,	Figure5.5.bmp	26,448,054 bytes
Figure 5.6,	Figure5.7.bmp	9,378,870 bytes
Figure 5.7,	Figure5.8.bmp	26,100,054 bytes
Figure 6.1,	Figure6.1.bmp	2,687,942 bytes
Figure 6.2,	Figure6.2.bmp	2,644,542 bytes
Figure 6.3,	Figure6.3.bmp	2,944,206 bytes
Figure 6.4,	Figure6.4.bmp	2,912,166 bytes
Figure 6.5,	Figure6.5.bmp	4,477,026 bytes
Figure 6.6,	Figure6.6.bmp	2,538,194 bytes
Figure 6.7,	Figure6.7.bmp	2,323,254 bytes
Figure 6.8,	Figure6.8.bmp	2,536,434 bytes
Figure 6.9,	Figure6.9.bmp	2,531,766 bytes
Figure 7.1,	Figure7.1.JPG	198,221 bytes
Figure 7.2,	Figure7.2.JPG	599,728 bytes
Figure 7.3,	Figure7.3.tif	848,552 bytes
Figure 7.4,	Figure7.4.tif	1,277,918 bytes
Figure 7.5,	Figure7.5.TIF	452,604 bytes
Figure 8.1,	Figure8.1.bmp	1,728,598 bytes

Figure 8.2,	Figure8.2.bmp	21,233,718 bytes
Figure 8.3,	Figure8.3.bmp	903,478 bytes
Figure 8.4,	Figure8.4.bmp	903,478 bytes
Figure A1,	FigureA1.gif	21,039 bytes
Figure B1,	FigureB1.bmp	120,054 bytes
Figure B2,	FigureB2.bmp	120,054 bytes
Figure B3,	FigureB3.bmp	120,054 bytes
Figure B4,	FigureB4.bmp	120,054 bytes
Figure B5,	FigureB5.bmp	120,054 bytes
Figure B6,	FigureB6.bmp	120,054 bytes
Figure B7,	FigureB7.bmp	120,054 bytes
Figure B8,	FigureB8.bmp	120,054 bytes
Figure B9,	FigureB9.bmp	1,610,454 bytes
Figure B10,	FigureB10.bmp	94,134 bytes
Figure B11,	FigureB11.bmp	35,910 bytes
Figure B12,	FigureB12.bmp	120,054 bytes
Figure B13,	FigureB13.bmp	22,734 bytes
Figure B14,	FigureB14.bmp	359,438 bytes
Figure B15,	FigureB15.bmp	120,054 bytes

Preface

This research seeks to substantiate three claims. (1) When applicable, optical spectrometry and acoustic spectrometry are preferred alternatives to slower and more invasive methods of analysis for sample quantification and classification. The prospect of noninvasive and nondestructive analysis is of interest to analysts across a number of different fields. For this reason, near infrared and acoustic resonance spectrometry are two rapidly growing analytical techniques in the pharmaceutical, food, and chemical industries. NIR spectrometry involves irradiation of the target analyte, excitation of vibrational and rotational motion of the functional groups, and the measurement of the absorbance pattern across the frequency range of interest. Despite vibrational and rotational motion, the energy provided with NIR light is insufficient to excite electrons from their ground state electronic configuration. The degree to which a molecule is free to vibrate and rotate can be visualized when one imagines two molecules attached to each other by a spring, where the spring represents the chemical bond. Effects such as molecule size and mass, force of attraction, position of neighboring molecules, and the length and type of the chemical bond all exert influences on the patterns of motion. The NIR region of the electromagnetic spectrum is a result of combination and overtone bands from the mid-IR fundamental vibration modes. Due to this fact, nearly every molecule absorbs in multiple locations of the NIR spectrum. As such, NIR analysis usually requires chemometric deconvolution and multivariate statistics to extract chemical information from the spectra.

Acoustic resonance spectrometry is based on the acoustic excitation of a molecular system while observing a given frequency range for resonance structures. The coupling between the acoustic excitation device and the chemical system creates a system of standing resonance waves as a

function of properties including: frequency and amplitude of the excitation signal, sample compressibility, density, mass, size, viscosity, and temperature. There are two primary effects seen in acoustic spectra, a shift in resonance frequency to a higher or lower frequency, and an increase/decrease in amplitude at a particular frequency. Often a frequency spectrum is sufficiently complex where multivariate statistics are needed to extract the exact relationship between physical properties and the accompanying spectra. As optical and acoustic data sets become more and more descriptive with each emerging generation of spectrometer, the computational burden associated with data processing also increases.

(2) Rather than digitally calculating multivariate relationships between physical properties and analytical spectra, chemometrics can be implemented directly into the physical design of spectrometers, thus sparing the need for computationally demanding post-collection multivariate analyses. This approach, termed integrated sensing and processing, is quickly emerging as the ideal approach for optical and acoustic imaging systems. Optical systems involve construction of optical transmission filters or masks, while acoustic systems involve construction of tailored excitation signals. Identification of the distinguishing spectral features from which to design the chemometrics is still accomplished digitally. Once the spectral features best used for classification or quantification are known, a weight function can be physically implemented into the design of the analytical system.

(3) Using this principle, ISP sensors can be developed specifically for use in highly applied situations, making possible automatic analyte quantification or classification without the computational burden and extensive data analysis typically associated with conventional

spectrometry. Ideally, ISP sensors are small, portable, inexpensive devices that can be designed to improve analysis in virtually any application. For example, small ISP sensors placed in-line in a pharmaceutical manufacturing process can offer insight into process variability. This type of information could be used as a feedback system to monitor and control manufacturing steps, in order to guarantee a predefined quality of pharmaceutical product. Contained within this dissertation are a number of near infrared applications, acoustic resonance applications, and examples of integrated sensing and processing sensors. The overall purpose of this dissertation is to clearly outline the advantages that ISP offers over both slower and more invasive methods of analysis, and conventional spectrometry for pharmaceutical and chemical analysis.

Section I: Chemometrics for ISP

Chapter One – Applications of integrated sensing and processing in spectroscopic imaging and sensing

Introduction

Integrated sensing and processing (ISP), a term coined by the Defense Advanced Research Projects Agency (DARPA), is an initiative calling for the implementation of new developments in chemometrics to the design of sensors and spectroscopic instrumentation.[1] Though not yet termed ISP, the idea was actually conceived in the 1970's and 1980's as a method to optically perform simple mathematical operations on spectroscopic data.[2] It has more recently been extended to more complex situations such as military imaging systems (in tanks and unmanned aerial vehicles) and the pharmaceutical industry for use in process analytical technology (PAT). PAT calls for the design and development of processes to guarantee a predefined quality of pharmaceutical materials at the end of the manufacturing process as warranted by risk analysis.[3] The goal of PAT is to encourage pharmaceutical companies to monitor each step of a manufacturing process, thus making ISP a perfect solution to the PAT initiative. ISP sensors are capable of measuring physical and chemical attributes in real-time, thus giving in-process and end-point knowledge of the quality and integrity of pharmaceutical materials.[3] Pharmaceutical PAT requires the development of rapid and accurate sensors so that existing full spectrum analytical instruments can be replaced with smaller, less expensive online alternatives. ISP is currently being studied as an alternative to traditional Fourier transform infrared (FT-IR), near-infrared (NIR), IR, UV-visible, fluorescence, Raman, and acoustic-resonance spectrometry (ARS).

Acoustic and optical spectrometry coupled with multivariate mathematics is routinely used for analyte quantification and identification, such as in NIR imaging for the quantification of organics in the presence of interferents,[4,5] scanning monochromator instruments and

multichannel array spectrometers,[6] and acoustic resonance spectrometry for the classification of tablets and quantification of active pharmaceutical ingredient in semi-solids.[7,8] As the complexity and dimensionality of data sets increases with each emerging generation of hyperspectral imaging, the amount of computing power necessary for processing increases significantly. For example, each image collected with a digital micromirror array or an acoustic-resonance spectrometer may contain 500,000 discrete wavelengths or frequencies. When experimenters measure the correlation between spectral features and physical properties (e.g. concentration, temperature, and/or hydration), high order relationships must be extracted. Often the relationship is determined with a method like principal component analysis (PCA). Linear PCA is limited by the fact that the maximum number of principal components that can be extracted is *I-1* where *I* is the number of spectra.[9] Because the number of wavelength or frequency observations often far exceeds the number of samples, these *I-1* components can fail to identify the spectral features that correlate to the measured property of interest. Higher order algorithms such as support vector machines (SVM) and kernel PCA (KPCA) can be employed in such situations, such as multidimensional image and multivariate data analysis. SVMs have been used for analysis of mid-infrared spectra,[10] nonlinear parametric models called multilayer perceptrons,[11] NIR spectra affected by temperature induced spectral variation,[12] NIR spectra for acidity prediction in grapes,[13] and neural networks.[14] KPCA has been used for optical character recognition and analysis of DNA,[15] images,[16] and NIR spectra.[17]

While these approaches have proven successful, SVM and KPCA are computationally demanding. In ISP, rather than performing the calculations explicitly in a computer, some portion of the instrumentation (typically the detector) is designed to complement the spectral

features of the incoming data set (e.g. frequencies and their respective amplitudes). The instrument itself extracts the distinguishing data structures without the need for further data processing. As new generations of spectrometry and imaging arrive with higher dimensionality, detectors are built to share the computational burden, obviating the need for increases in post-collection computing power. Currently, there are many different research groups approaching this goal in a number of different ways.

ISP with Optical Signal Processing

Theory. The most widely used ISP alternative to conventional optical spectrometry involves the construction of optical interference filters. One of the first published methods, termed optical signal processing (OSP), is based on the four components illustrated in Figure 1.1: an element to disperse wavelength-dependent information into space-dependent information, a spatially variant optical transmission mask to filter incoming spectroscopic information, optical or analog elements to integrate the transmitted intensity, and a processor to perform simple algebraic computations.[18] The dispersion element directs wavelengths to different spatial locations on the transmission mask where the mathematical product is calculated between the spectroscopic information and the orthogonal weight function in the mask. The vector dot product is calculated as the sum or the integrated total signal by focusing the light from the transmission mask on the photodetector such that only one voltage, proportional to the concentration of the analyte, is recorded at the detector.

To optically process an emission signal composed of several component spectra, a vector function is formulated so that it is orthogonal to the emission spectra of interfering species. The

product of this vector and the interfering species effectively eliminates light emission from the interferents. The vector function is also formulated so that it is not orthogonal to the emission signal of the analyte of interest. In this manner the integral product of the vector and dispersed emission leaves only the signal from the analyte of interest, making the signal directly proportional to the analyte concentration, suggesting that the overall purpose of an OSP mask is to maximize the signal-to-background ratio. Clearly, this means that if there is significant overlap between the interferents and the analyte emission spectrum, or in the presence of many interferents, the throughput of the instrument will decrease significantly.

Mathematical construction of the optical filters can be illustrated in the following example. The total emission spectrum, free from interferences and directly related to the analyte property of interest is known as the Net Analyte Signal (NAS).[19,20,21,22] Free from instrument variation (noise, baseline shift) and chemical properties (temperature, viscosity, etc.), the NAS can be represented by the wavelength-dependent emission spectrum, $\mathbf{a}_\mathbf{a}$, and a total emission magnitude $\mathbf{x}_\mathbf{a}$. The interference components $\mathbf{a}_\mathbf{i}$ and $\mathbf{x}_\mathbf{i}$ are arranged in matrix **I**. Gram-Schmidt orthonormalization of **I** results in a matrix of orthogonal vectors, **Q**, related to the original matrix by the factorization matrix, **F**.

$$\mathbf{I} = \mathbf{QF} \qquad \mathbf{1.1}$$

For an emission spectrum **a**, a least-squares solution is used to determine the individual emission magnitudes, $\mathbf{x}_\mathbf{n}$, which are components of the vector **x**, such that the 2-norm is a minimum:

$$\min\left\|\mathbf{Ix} - \mathbf{a}\right\|_2 \qquad \mathbf{1.2}$$

Premultiplication of Equation 1.2 gives:

$$\min\left\|(\mathbf{Q}^{\mathrm{T}}\mathbf{I})\mathbf{x} - (\mathbf{Q}^{\mathrm{T}}\mathbf{a})\right\|_2 \qquad \mathbf{1.3}$$

where the superscript T indicates the matrix transpose. With Equation 1.1 and the fact that vector components of **Q** are orthonormal, the solution is:

$$\mathbf{Fx} = \mathbf{Q}^{\mathrm{T}}\mathbf{a} \qquad \mathbf{1.4}$$

$\mathbf{Q}^{\mathrm{T}}\mathbf{a}$ is a vector with elements equal to the dot products of **a** with the individual component vectors of $\mathbf{Q}^{\mathrm{T}}$. The dot product of $\mathbf{q_n}$, the last vector of $\mathbf{Q}^{\mathrm{T}}$, with the emission spectrum, **a**, is the optimal estimate, $\mathbf{x_n}$, which is used as the optical transmission filter.

Many recently published ISP applications stem directly from OSP research and the construction of optical interference filters. A discussion of the results appears later with the presentation of more recent OSP material.

ISP with Optical Regression

Theory. ISP with optical regression[23] is designed as an improvement over scanning spectrometers. In response to the assumption that in any data set there are many uninformative or useless variables, this method uses analog variable selection as opposed to post-collection processing, saving valuable time by collecting only the important variables. The overall purpose of optical regression is to improve S/N. Some authors contend that a digital regression, while powerful and flexible, does not necessarily give the optimal precision. In a typical scanning spectrometer, equal time is dedicated to the collection of each wavelength, thus valuable integration time is sometimes dedicated to the collection of useless data. The regression vector in optical regression serves as template for the amount of integration time spent at each wavelength. Thus, by varying the amount of time spent collecting light at each wavelength, the

spectrum collected at the detector automatically represents the dot product of the transmission spectrum and the regression vector, saving valuable post-collection processing time. Because the transmission spectrum cannot be negative, two spectra are collected and differenced; one for the positive lobes of the regression vector, and one for the negative lobes. This difference is directly proportional to the concentration of the analyte of interest.

The integration times are determined by the inverse calibration model calculated by:

$$\mathbf{c} = \mathbf{A}\mathbf{b} + \mathbf{e} \qquad \textbf{1.5}$$

where **b** is the regression vector, **c** is a set of properties associated with an *I x J* matrix **A** composed of *I* independent samples and *J* observations, where **e** is a *J x 1* vector of model errors. The dot product is calculated between the regression vector and the unknown spectrum, **a**:

$$\hat{c}_{un} = \mathbf{a}^{\mathrm{T}}\mathbf{b} = \sum_{j=1}^{J} a_j \mathbf{b}_j \qquad \textbf{1.6}$$

where $\hat{c}_{\mathrm{un}}$ is the estimation of the property of interest. If the regression vector is free from random errors, the errors in $\hat{c}_{\mathrm{un}}$ are the dot product of the errors, **e,** in unknown spectrum **a** with the regression vector **b**:

$$\mathbf{e}_{\hat{c}} = \mathbf{e}^{T}\mathbf{b} = \sum_{j=1}^{J} \mathbf{e}_j \mathbf{b}_j \qquad \textbf{1.7}$$

The expected standard deviation of $\hat{c}_{\mathrm{un}}$ for normally distributed errors with mean of zero and standard deviation, σ, is given by:

$$\sigma_{\hat{c}} = \left(\sum_{j=1}^{J} \sigma_j^2 \mathbf{b}_j^2 \right)^{1/2} = \sigma \left(\sum_{j=1}^{J} \mathbf{b}_j^2 \right)^{1/2} = \sigma \left\| \mathbf{b} \right\|_2 \qquad \textbf{1.8}$$

In a spectrum with J wavelengths, there are J time units for the integration of the entire spectrum. The time dedicated for collection at each individual wavelength is determined by normalizing with the 1-norm, the sum of absolute values, on the regression vector to J:

$$\mathbf{b^t} = \mathbf{b}(J/\|\mathbf{b}\|_1) \qquad \textbf{1.9}$$

where $\mathbf{b^t}$ is the sum of $\mathbf{b}^{\mathbf{t}}_{j+}$, the time dedicated to each of the wavelength with a positive regression value and $\mathbf{b}^{\mathbf{t}}_{j-}$, the time dedicated to wavelengths with a negative regression value. The difference between the two measurements is equal to $J/\|\mathbf{b}\|_1$ times the calibrated property.

$$\hat{c}_{un} = \left(\sum_{j=1}^{J} \mathbf{b}^{t}_{j+} - \sum_{\backslash j=1}^{J} \mathbf{b}^{t}_{j-} \right) \|\mathbf{b}\|_1 / J \qquad \textbf{1.10}$$

If the measurement errors are equivalent to the errors in Equation 1.8, the standard deviation of $\hat{c}_{un}$ is:

$$\sigma_{\hat{c}} = (\sigma_+^2 + \sigma_-^2)^{1/2} \|\mathbf{b}\|_1 / J = 2^{1/2} \sigma \|\mathbf{b}\|_1 / J \qquad \textbf{1.11}$$

In digital regression, the detector collects K measurements per unit time over the J digitized channels, thus there are KJ total measurements. The regression vector must therefore be scaled such that the 1-norm equals KJ:

$$\mathbf{b^t} = \mathbf{b}(KJ\|\mathbf{b}\|_1 \qquad \textbf{1.12}$$

where the J elements of $\mathbf{b^t}$ designate the number of signals integrated at each channel. The standard deviation of estimation here is expressed as:

$$\sigma_{\hat{c}} = (KJ)^{1/2} \sigma \|\mathbf{b}\|_1 / J = \sigma K^{1/2} \|\mathbf{b}\|_1 / J^{1/2} \qquad \textbf{1.13}$$

The relative magnitudes of the standard deviations from analog (Equation 1.8) and digital (Equation 1.11) regression can now be compared. Analog optical regression predicts a lower standard deviation of estimation than digital regression by:

$$2^{1/2}\sigma\|\mathbf{b}\|_1 / J < \sigma\|\mathbf{b}\|_2 \qquad \textbf{1.14}$$

since by definition of multivariate analysis J is 2 or greater.

Results from Optical Regression. To demonstrate that optical regression outperforms digital regression in practice, Prakash et al. provide a short description of its application to the concentration prediction of 4 mixtures of rhodamine B, sulforhodamine, and rhodamine 590 in 20:80 methanol:water using fluorescence emission spectra. The regression vector was calculated by principal component regression (PCR), and visual inspection was used to select the 20 wavelengths with the highest predictive ability. The regression vector determined how much time was spent integrating the signal at each of the 20 wavelengths. With 20 wavelengths, digital regression allocated 30 ms to integrate each channel, whereas optical regression integrated each channel according to its weight in the regression vector. Elimination or reduction of the contribution of less important wavelengths in optical regression effectively raised the signal-to-noise ratio. Of the 12 different mixtures, on average, optical regression predicted analyte concentration 28.0% more precisely than digital regression. Table 1.1 summarizes the results from optical regression.

Multivariate Optical Elements

Theory. The overall purpose of these optical interference filters, termed multivariate optical elements (MOE),[4,5,6,24,25,26,27] is to minimize the standard error of prediction (SEP), or the root-mean-squared error of calibration (RMSEC) of chemical species. MOE methods begin with the collection of spectra of target analytes. For example, if the objective is the NIR quantification of ethanol in water, spectra from different concentrations of ethanol in water are collected with a

broadband source. Calculation of the principal components of the raw spectra is usually accomplished by a singular value decomposition of matrix **A** according to Equation 1.15:

$$\mathbf{A} = \mathbf{USL} \qquad \textbf{1.15}$$

where **A** is the matrix of original spectra, the eigenvectors of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ make up the columns of **L**, the eigenvectors of $\mathbf{AA}^{\mathbf{T}}$ make up the columns of **U**, and **S** is a diagonal matrix of singular values.[9]

As shown in Equation 1.16, a regression using **U** reveals which has the strongest correlation to a change in ethanol concentration **c**, where y is the y-intercept, **b** is a vector of regression coefficients, and **e**, is the residual.

$$\mathbf{c} = y + \mathbf{bU} + \mathbf{e} \qquad \textbf{1.16}$$

Equation 1.17 demonstrates how a leave-one-out cross validation can be used in predicting the reliability of the concentrations of ethanol from NIR spectra, where σ^2 is the variance, $\hat{c}_{\text{un}}$ is the prediction of the model for the *i-th* pattern *I* in the training set, after it has been trained on the *I-1* other patterns.

$$\sigma^2_{LOO} = \frac{1}{I}\sum_{i=1}^{i=I}(y_i - \hat{c}_{un})^2 \qquad \textbf{1.17}$$

In the case of a minimal two-component system using ethanol and water, it is a simple matter to observe a linear change in ethanol concentration. With just two chemical constituents, only one principal component is needed to accurately describe the concentration changes that occur in mixtures. The loading corresponding to this principal component reflects the contribution of each wavelength to the overall classification. The concentration of unknown samples can be quantified by multiplying the loading and spectral intensity at the same wavelength, then summing over all wavelengths according to Equation 1.18 below:

$$y_i = \sum_{j=1}^{j=J} l_j \mathbf{A}_{ij} \qquad \textbf{1.18}$$

where l_j is the loading at wavelength j, $\mathbf{A_{ij}}$ is the intensity of the raw spectrum of the *i-th* mixture at wavelength j, y_i is the chemical analysis result, and j is an index over J wavelengths in the spectrum.

Figure 1.2 illustrates a generic version of the MOE selection process. The spectral variables demonstrate that the two peaks needed for sample differentiation describe the majority of the variation. The corresponding principal component loading identifies these regions and can be used as a template for the design of the optical filters. Because transmission spectra cannot be negative, negative loadings must be inverted as shown in MOE-2. Figures 1-3 and 1-4 illustrate the basic configurations of this instrument.

The MOEs in these configurations act as optical beamsplitters. The MOEs are designed so the transmittance to detector D_T at wavelength j is expressed by Equation 1.19, while the reflectance to detector D_R is expressed by Equation 1.20:

$$\mathbf{T}_j = 0.5 + kl_j \qquad \textbf{1.19}$$

$$\mathbf{R}_j = 0.5 - kl_j \qquad \textbf{1.20}$$

where k is a proportionality constant. These equations assume that the MOE does not absorb light or scatter stray light, thus the sum of **T** and **R** is unity at each wavelength. When light from all J wavelengths strikes the MOE, the total reflectance is given by Equation 1.21 and the total transmittance is given by Equation 1.22:

$$\mathbf{R}_{t,i} = \sum_{j=1}^{j=J}(0.5 - kl_j)\mathbf{A}_{ij} = 0.5\sum_{j=1}^{j=J}\mathbf{A}_{ij} - k\sum_{j=1}^{j=J} l_j \mathbf{A}_{ij} \qquad \textbf{1.21}$$

$$\mathbf{T}_{t,i} = 0.5\sum_{j=1}^{j=J}\mathbf{A}_{ij} + k\sum_{j=1}^{j=J} l_j \mathbf{A}_{ij} \qquad \textbf{1.22}$$

The difference between transmittance and reflectance is shown in Equation 1.23.

$$(\mathbf{T}_{t,i} - \mathbf{R}_{t,i}) = 2k\sum_{j=1}^{j=J} l_j \mathbf{A}_{ij} = 2ky_i \qquad \textbf{1.23}$$

This equation demonstrates that the difference between transmittance and reflectance is directly proportional to analyte concentration.

The actual construction of the MOE has been approached in two different ways. The objective of both methods, however, is consistent; to design a MOE that has a transmission spectrum resembling the desired loading vector. In the first approach, a library of transmission spectra from organic compounds is screened to identify the optimum combination of organic materials whose transmission spectra will sum to the desired regression vector. The identified MOEs can be placed in a glass cuvette or in a rotating filter wheel directly before or after the analyte. These specific MOEs have been termed molecular optical filters. The second approach involves the deposition of alternating layers of two metal oxide films, Nb_2O_5 and SiO_2, on a 2.54 cm BK-7 glass substrate with reactive magnetron sputtering.[28] The interference effects of the layers are monitored as the layers are deposited to be sure the end-product resembles the target transmission spectrum.

The utility of optical transmission filter spectrometers lies in the speed and accuracy at which they operate. Rather than using complex, bulky machinery with a dispersion element such as a grating or a stepping monochromator, one photodetector is capable of extracting the same information with no post-collection computation (the ISP advantage). Collection of all wavelengths simultaneously provides the multiplex advantage. The absence of a slit to achieve wavelength resolution provides the throughput advantage. An instrument simultaneously measuring a signal over a range of wavelengths/frequencies obtains a $t^{1/2}$ advantage in the time t required to obtain a given signal-to-noise ratio compared to that which would be necessary using dispersive methods.[29]

Select Results from Multivariate Optical Elements. Soyemi et al. describe an example of the use of a single MOE filter for the integrated sensing and prediction of analyte concentration. Transmission spectra from 400-650 nm were collected from mixtures of Bismarck Brown (BB) and Crystal Violet (CV), where BB was the analyte and CV was the random interferent. Digital principal component regression (PCR) using a 4- component model gave a standard error of calibration (SEC) of 2.9% error relative to the mean of the calibration set, and a standard error of performance (SEP) of 4.3% relative to the mean of the calibration set. A plot of the root mean squared error of cross validation vs. number of principal components indicated how many components were required for the best predictive ability. Ten MOEs, with layers ranging from 1 to 31, were constructed to resemble the regression vector as calculated from the PCR loadings. The best filter had 15 layers and gave SEC = 0.7% and SEP = 0.8% relative to the mean of the calibration set. (See Table 1.2 for the results of all the filters in comparison to the 4- component digital regression model.) MOEs demonstrated an improvement of two orders of magnitude in

both SEP and SEC over digital PCR for the prediction of analyte concentration.[6] The authors speculated that this is partly due to the inclusion of more model error in the PCR model, whereas the MOE avoids model error because it passes a smooth transmission waveform to the detector.

ISP in Hyperspectral Integrated Computational Imaging

Background. Hyperspectral integrated computational imaging is the imaging of a target analyte at a number of discrete wavelengths simultaneously.[30] As illustrated in Figure 1.5, HICI forms a data cube or hypercube. Data cube axes have different dimensions depending on the application. For example, the three dimensions in fluorescence excitation-emission matrix spectrometry are excitation wavelength, emission wavelength, and sample number, which all serve as indices for the recorded property, spectral intensity.[31] However, for NIR imaging, spectral intensity is indexed by two spatial dimensions (length and width), excitation wavelength, and sample number.[30,31] Typical fluorometry and NIR spectrometry are conducted by varying the excitation wavelengths and raster scanning across an analyte surface until an entire spectrum is collected at the photodetector. In HICI, data across all wavelengths are collected simultaneously with a digital camera such as a charge-coupled device (CCD) camera. In either case, the resulting image is often presented as a contour plot as illustrated in Figure 1.5. The collection of all data in one measurement speeds up data collection times and standardizes the results. For example, environmental conditions, temperature fluctuations, ambient light differences, and detector drift for a given set of scans may all be subject to change when collecting multiple spectra over time. On the other hand, these effects are negligible when collecting only one measurement. Thus, a main advantage of HICI over conventional imaging spectrometry is reduced data collection time.

Theory. A number of chemometric techniques have been applied to the processing of the resulting data cubes from HICI images, including parallel factor analysis (PARAFAC)[31] and a nonparametric cluster analysis algorithm, the bootstrap error-adjusted single sample technique (BEST).[32] This review will concentrate on the BEST, which can be implemented in simple hardware and is based on the premise that spectra from similar samples tend to cluster together in the same region of hyperspace. The BEST draws probability-density contour plots around analyte populations based on asymmetric standard deviations (SDs).[32] For a more detailed description of the BEST, an in-depth discussion is provided by Hamilton et. al. In the BEST, a population **P** is created as an *I x J* matrix in hyperspace H whose rows are the individual samples and the columns are the wavelengths. **P*** is a discrete realization of **P** based on a calibration set **A** of the same dimensions as **P***. This realization is chosen one time from **P** to approximate all possible sample variations present in **P**. **P*** has parameters $\mathbf{M_c}$ and $\overline{\mathbf{P}}$, where $\overline{\mathbf{P}} = E(\mathbf{P})$ and $\mathbf{M_c}$ is the Monte Carlo approximation to the bootstrap distribution. The expectation value, E(**P**), is the center of **P**, and $\overline{\mathbf{P}}$ is a row vector with the same number of rows as there are columns in vector **P**. New test spectra $\mathbf{A}_T$ are projected into H containing $\mathbf{M_c}$ rows of $\mathbf{M_c}$ are mapped onto a vector connecting $\overline{\mathbf{P}}$ and $\mathbf{A}_T$. $\overline{\mathbf{P}}$ and $\mathbf{A}_T$ have the same dimensions. The integral over H is calculated from the center of **P** in all directions. A skew-adjusted SD is based on the comparison of the expectation value $\overline{\mathbf{P}} = E(\mathbf{P})$ and $\overline{\mathbf{P}} = med(\mathbf{A})$, the median of **A** in hyperspace is projected on the hyperline connecting $\overline{\mathbf{P}}$ and $\mathbf{A}_T$. The result is an asymmetric SD that provides two measures of the SD along the hyperline connecting $\overline{\mathbf{P}}$ and $\mathbf{A}_T$. Skew adjusted SDs can be used to calculate mean distances between spectra of different samples.

Select Results from Hyperspectral Integrated Computational Imaging. Groups of aspirin tablets exposed to different levels of moisture have been used to test the assumption that the BEST approach offered an advantage over conventional NIR spectrometry.[30] Tablet groups were exposed to water vapor or a pH 9.0 ammonium hydroxide solution. Water uptake was determined by gravimetric analysis and NIR spectrometry. To verify that NIR could be used to differentiate between different tablets, calibration lines were constructed for both water uptake and for the decomposition of aspirin into salicylic acid. The standard error of estimate (SEE) for the prediction of water content was 0.05% and the standard error of performance (SEP) was 0.06%. The SEE and the SEP for the prediction of salicylic acid were each 0.06%. NIR images of 1286 control tablets were collected simultaneously with a CCD camera, and the BEST algorithm was used to calculate multidimensional standard deviations. Figure 1.6 is a contour plot with lines drawn in BEST standard deviations. It is apparent from this plot that NIR imaging can readily indicate water exposure time. In this example, NIR imaging resulted in a 30,000 fold increase in speed over HPLC and a 1,000 fold increase in speed over conventional NIR spectrometry, suggesting a high degree of utility in process analytical technology, as thousands of tablets can be scanned simultaneously and in real-time as they move down a manufacturing line.

Acoustic-Resonance Spectrometry with an ISP Excitation Signal

Theory. Typical acoustic-resonance spectrometric measurements involve very large amounts of data. For example, 15 seconds of data collection at a sample rate of 44.1 KHz results in 661,500 data points per spectrum. Normally, there is a large disparity between the number of sample spectra and the number of frequencies at which data are collected on the samples.[33] Due to

instrumental (noise, baseline shift, etc.) and chemical inconsistencies (temperature, viscosity, etc.), spectral features of interest may be severely overlapped by interferences. Additionally, not all frequencies respond linearly or exclusively to the desired change in analyte concentration, therefore, linear principal component regression may be inadequate to describe the relationship between AR spectra and physical properties. For PAT applications the data burden mandates that an ISP method be developed to reduce the dimensionality, retaining only those frequencies directly proportional to the concentration of the physical property of interest.

Figure 1.7 illustrates the ISP encoding and processing scheme. Full spectra are collected in the time domain, and transformed to the frequency domain with an FFT. Fisher weights are calculated for each discrete frequency according to Equation 1.24.[7,8] Frequencies with the top weights are selected, and an inverse Fourier transform is performed to put the frequencies back into the time domain. This file, which serves as the new excitation signal, can be saved as an MP3 file and encoded on a CD/MP3 player.

$$\mathbf{F} = \frac{\left|\overline{\mathbf{P}}_1 - \overline{\mathbf{P}}_2\right|}{\left(\frac{1}{I-1}\sum_{i=1}^{i=I}(\mathbf{P}_{1i} - \overline{\mathbf{P}}_1)\right) - \left(\frac{1}{I-1}\sum_{i=1}^{i=I}(\mathbf{P}_{2i} - \overline{\mathbf{P}}_2)\right)} \qquad \textbf{1.24}$$

Because the excitation signal was created using only those frequencies directly proportional to the concentration of the physical property of interest, once the signal passes through the sample the signal intensity is directly proportional to the analyte concentration. No FFT or spectral sweeping is required to generate the result. In fact, a simple rectifier circuit and volt meter will suffice.[8]

In generating the Fisher weights, data populations $\mathbf{P_1}$ and $\mathbf{P_2}$ contain *I x J* points where *I* rows contain the samples and *J* columns contain the frequencies. The algorithm uses an *Jx1* classification matrix, **c**, to identify the number of unique classes, $\mathbf{c}_u$, in a given data set. The number of possible combinations of classes, and thus the number of rows in the Fisher weights matrix **F,** I_f**,** is calculated by Equation 1.25.

$$I_f = \frac{1}{2}\mathbf{c}_u(\mathbf{c}_u - 1) \qquad \textbf{1.25}$$

Each row in **F** is composed of short sound segments or "bites" that are designed for the differentiation of specific analytes. When trying to differentiate between groups of similar samples, the most distinguishing frequencies for one analyte are usually the most distinguishing frequencies for all analytes in that group. It is this property that allows the use of only one row in **F** for the final integrated sensing and processing sound segment. The row is selected based upon some other method of multivariate analysis, such as partial least squares, PCR, or the BEST nonparametric clustering algorithm. Whichever two analytes exhibit the largest separation according to these methods indicates which of the rows in **F** are used as the excitation source.

Despite the potential introduction of a few uninformative variables for each individual calibration by using only one row in **F**, the overall reduction in insignificant frequencies boosts the S/N much higher than seen in conventional ARS. In addition, the ISP approach offers a large advantage in terms of speed, which outweighs the small loss in performance due to the possible inclusion of useless variables. Data sets collected with the new excitation source are presently still collected in full, however, the computational burden is greatly reduced by filtering the incoming data according to the pre-selection of the distinguishing frequencies. Future AR

spectrometers will include analog frequency selection elements and an analog integrator so that fewer data points can be collected.

Results from ISP Acoustic-Resonance Spectrometry. ISP ARS has been applied to the differentiation of similarly sized and shaped tablets.[7,8,33] The success metric for the differentiation of tablets was intercluster distance in BEST standard deviations. Five groups of tablets, aspirin, ibuprofen, acetaminophen, vitamin B12, and vitamin C were selected for analysis. For initial data collection, tablets were scanned in random order for 15 seconds at a sample rate of 8 KHz. BEST SDs were calculated, and the excitation signal was selected from the tablets with the largest intergroup separation. All tablets were rescanned with the new excitation source for the ISP approach. The average multidimensional SD between tablet groups for the non-ISP approach was 13.52, and for the ISP approach was 10.8. Table 1.3 lists all results. While the average SD was higher for the non-ISP approach, tablet populations separated by 3 or more SDs are considered separable. In addition, not all tablet groups separated in the non-ISP approach. This suggests that the ISP offers an advantage in reliability of separation as well as in data reduction and simplicity of detection, presumably by the increase in S/N due to the elimination of uninformative variables.

Conclusion

Integrated sensing and processing is a viable alternative to conventional spectrometry and imaging, offering substantially faster analysis times while maintaining equivalent performance with its digital counterparts. Without ISP, it is necessary for analytical techniques to make a trade-off between longer data acquisition (scan) times or lower predictive ability. With the

emergence of ISP, however, experimenters can reduce the data collection times and data dimensionality without forfeiting high predictive performance. ISP has shown promising results in spectrometry and imaging, and has demonstrated a high degree of utility for military applications and pharmaceutical process analytical technology.

Chapter One Figures

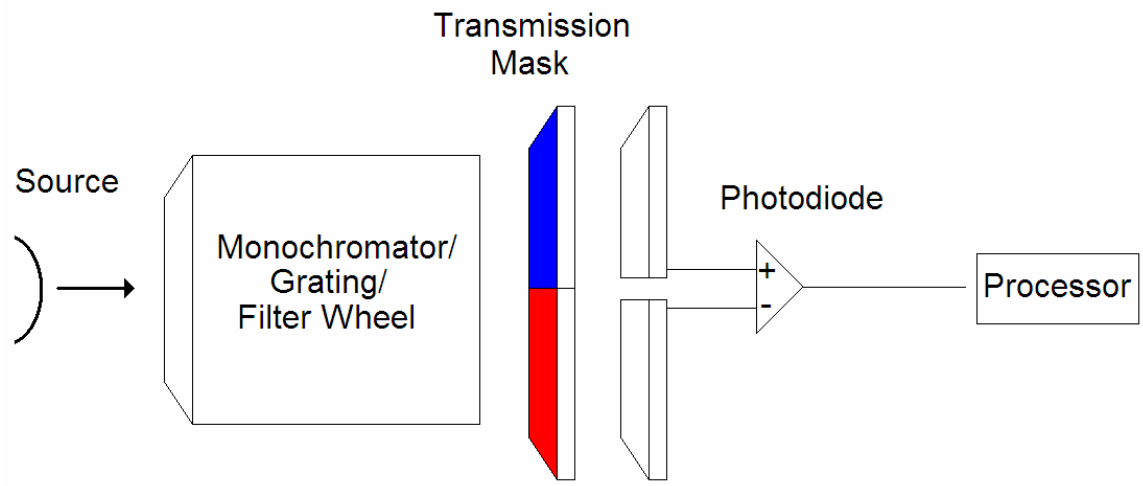

Figure 1.1: Optical signal processing (OSP) is based on the four components illustrated in this figure. The dispersion element (monochromator/grating/filter wheel) wavelength-dependent information is coupled to a spatially variant optical transmission mask to filter incoming spectroscopic information, optical or analog elements integrate the transmitted intensity, and the analog processor subtracts the signal from the negative lobe of the regression from the positive lobe signal.

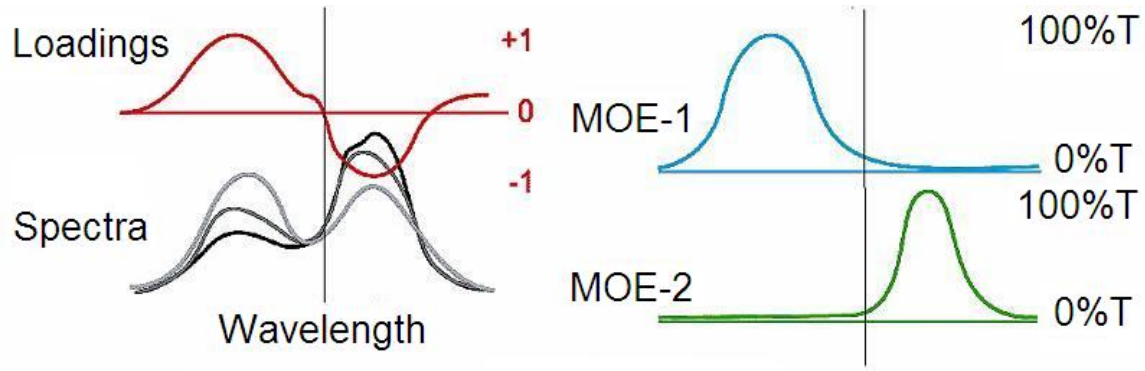

Figure 1.2: In this simple example, the spectral variables demonstrate that the two peaks needed for sample differentiation describe the majority of the variation in the spectra. The principal component loading corresponding to this change can be used as a template for the design of the optical filters. Molecular filters are designed so that their transmission spectra match the desired principal component loading vector. Transmission spectra cannot be negative; therefore, negative loadings must be inverted separately as MOE-2.

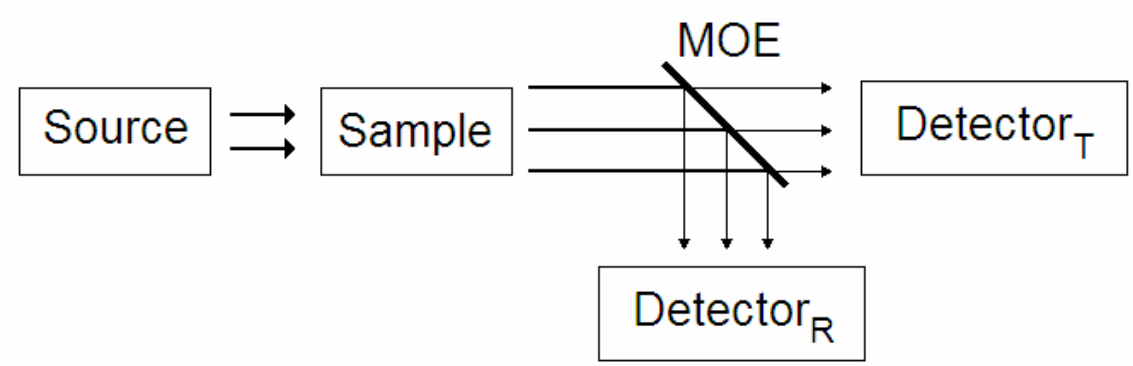

Figure 1.3: A simple T-format spectrometer with the multivariate optical element (MOE) as the beamsplitter. The transmitted signal passes through the MOE where it is collected at detector T, and the reflected signal bounces off the MOE for collection at detector R. The difference between transmission, **T**, and reflection, **R**, is directly proportional to the concentration of the analyte.

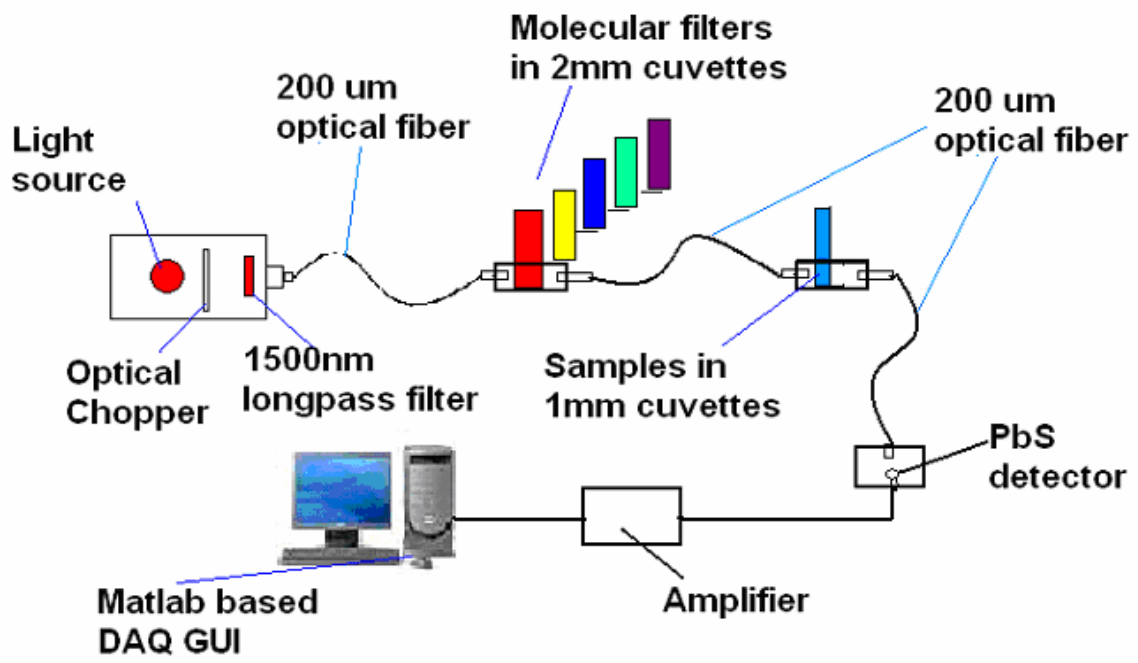

Figure 1.4: Diagram of a high-throughput molecular filter spectrometer. The molecular filters (MF), which are designed to resemble the principal component loading vectors of interest, are placed in line with the sample. This implicitly calculates the dot product between the MFs and the sample transmission spectra, automatically weighting the most distinguishing variables for rapid and accurate analyte quantification.

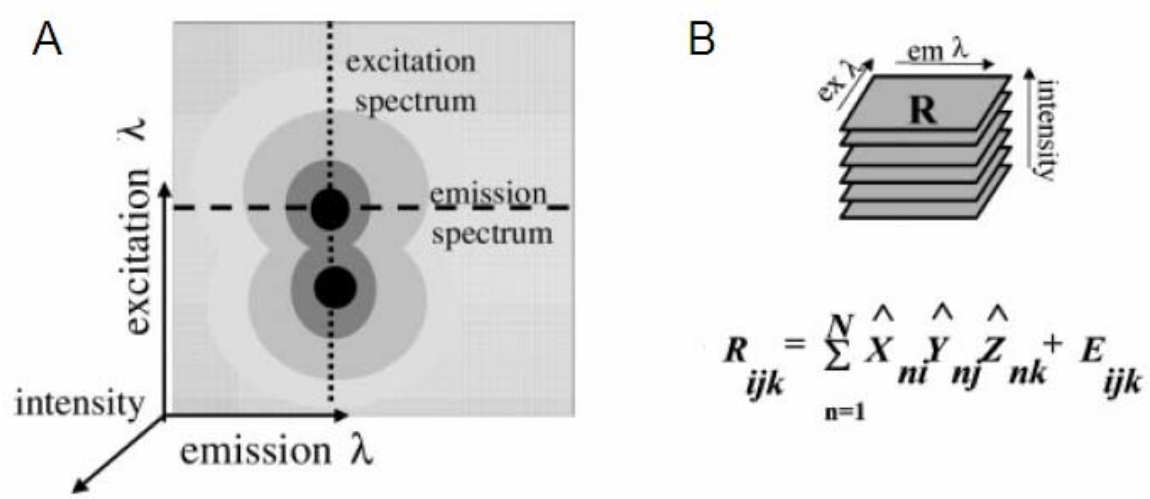

Figure 1.5: (a) A contour plot of an excitation-emission fluorescence spectrum. Rows represent emission spectra at a particular excitation wavelength, and columns represent excitation spectra at a particular emission wavelength. **(b)** A three-way data cube can be constructed by stacking correlated EEM spectra. The cube can be mathematically modeled by a set of trilinear components, where $\mathbf{R}_{ijk}$ is the fluorescence intensity of sample *k* at excitation wavelength *i* and emission wavelength *j*, *N* is the number of unique spectral profiles found in the data cube, the columns of X, Y and Z are the estimations of the pure excitation, emission and concentration profiles respectively and $\mathbf{E}_{ijk}$ is the error matrix.

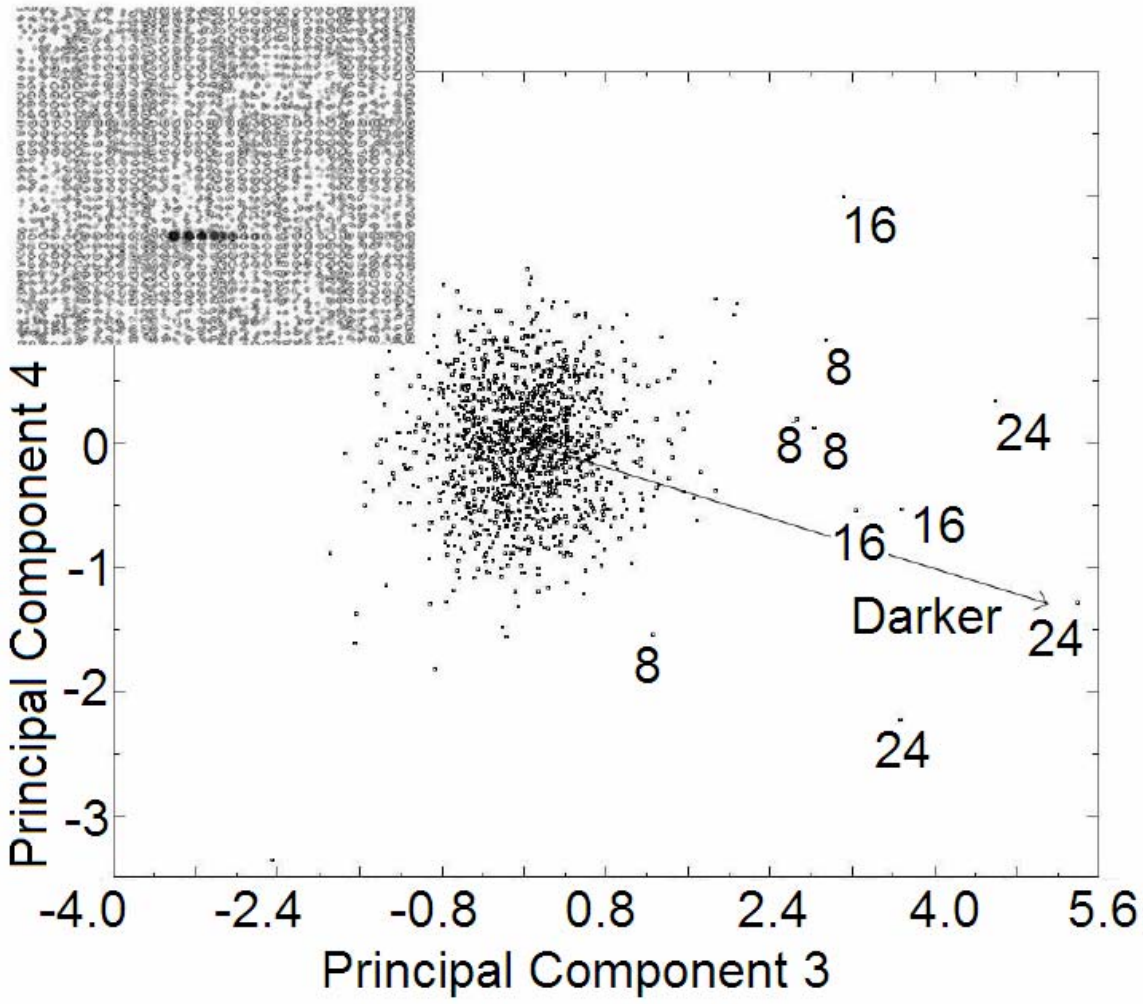

Figure 1.6: Bootstrap error-adjusted single sample technique contour plot of 1286 standard aspirin tablets scanned through the blister packaging with a NIR camera. Distances calculated in BEST standard deviations from the center of the population distribution indicate the moisture absorption of individual tablets.

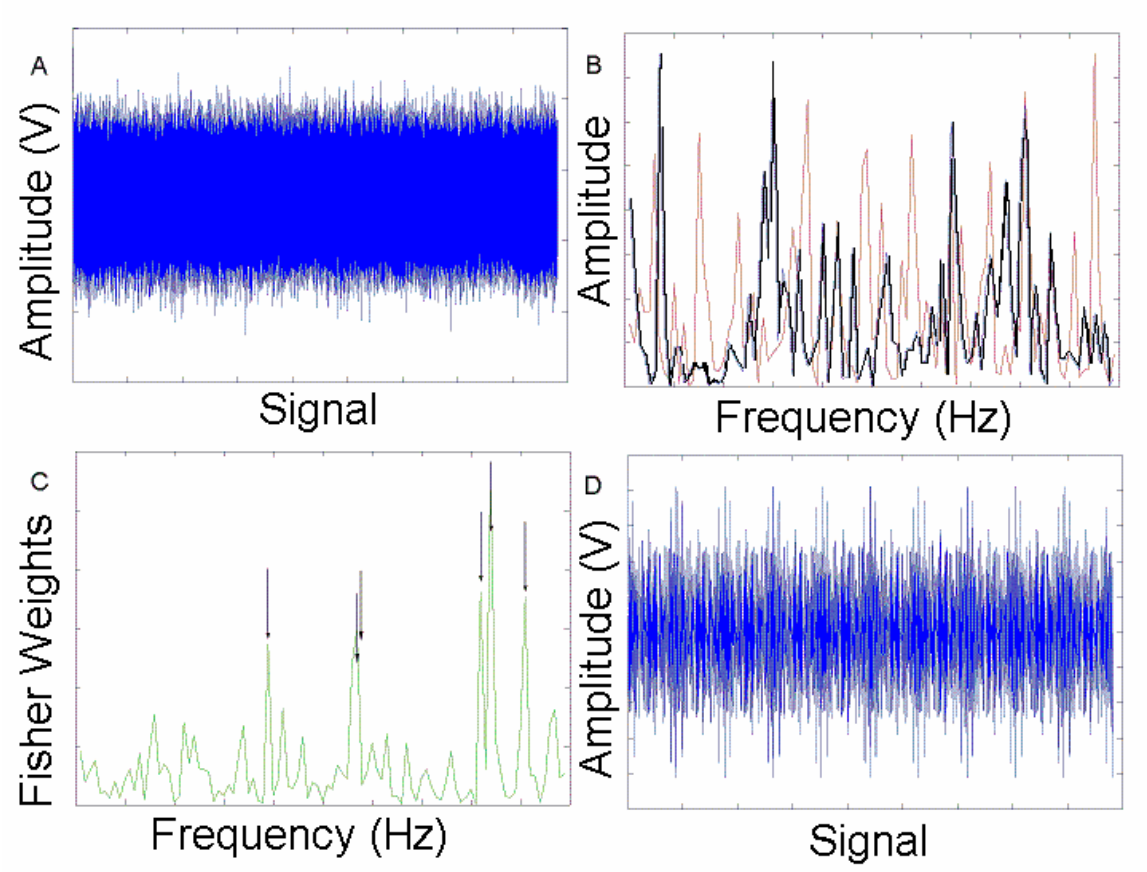

Figure 1.7: (a) Voltage signal collected for each analyte [sample rate of 44.1 kHz, duration 15 seconds, thus 661,500 data points] **(b)** Fourier transform calculated to put data into frequency domain, allows for identification of distinguishing frequencies, thin line = analyte 1, thick line = analyte 2, **(c)** calculation of Fisher weights identifies most distinguishing frequencies, and **(d)** those frequencies are used to construct the sound file, (for simplicity of visualization, the sum of only 6 frequencies is used in this example), using a sample rate of 44.1 kHz for 15 second duration.

Chapter One Tables

Table 1.1: Optical regression demonstrated an average improvement over digital regression by 28.0% for the prediction of fluorescent dye mixtures.

Analyte	Exp. #	Conc. (ppm)	Mean Bias (ppm)	$\sigma_{\hat{c}}$ Digital	$\sigma_{\hat{c}}$ Optical	Improvement %
Rhodamine B	1	5	-0.0074	0.0944	0.0700	25.8
	2	5	-0.0126	0.0828	0.0526	36.5
	3	0	0.0085	0.0880	0.0527	40.1
	4	5	0.0624	0.0815	0.0552	32.2
Rhodamine 590	1	1	-0.0248	0.0561	0.0527	6.1
	2	1	-0.0039	0.0438	0.0329	24.9
	3	1	0.0378	0.0570	0.0358	37.2
	4	0	-0.0502	0.0469	0.0364	22.4
Sulforhodamine	1	0	-0.0012	0.0620	0.0509	17.9
	2	10	0.0007	0.0617	0.0468	24.1
	3	10	-0.0526	0.0747	0.0458	38.7
	4	10	0.1089	0.0640	0.0450	29.7

Table 1.2: Multivariate optical elements (MOE) with different numbers of layers constructed for the quantification of Bismarck Brown and Crystal Violet dyes. Optical regression with the use of MOEs outperformed digital principal component regression by an order of magnitude for standard errors of calibration, performance, and the root mean squared errors. Statistics are reported relative to the mean of the calibration set, and as such, are unitless.

No. layers	SEC	SEP	Average RMS
29	**0.00950**	**0.00792**	**0.00875**
28	**0.00924**	**0.00802**	**0.00865**
19	**0.00923**	**0.00796**	**0.00862**
29	**0.00841**	**0.00875**	**0.00858**
20	**0.00885**	**0.00822**	**0.00854**
30	**0.00919**	**0.00782**	**0.00853**
21	**0.00909**	**0.00784**	**0.00849**
20	**0.00879**	**0.00794**	**0.00838**
25	**0.00728**	**0.00846**	**0.00789**
15	**0.00699**	**0.00794**	**0.00748**
PCR Model 4 components	**0.029**	**0.043**	**0.045**

Table 1.3: Bootstrap error-adjusted single sample technique (BEST) standard deviations for tablet clusters as scanned by acoustic resonance spectrometry (ARS). The non-integrated sensing and processing (ISP) approach resulted in an average intertablet standard deviation of 13.52 and the ISP approach had an average intertablet standard deviation 10.8. Aspirin and vitamin B12 failed to separate in the non-ISP approach, whereas all tablets separated in the ISP approach. Therefore, ISP offers both an advantage in separation as well as faster processing times.

Non-ISP	**Aspirin**	**Ibuprofen**	**APAP**	**Vitamin B12**	**Vitamin C**
Aspirin	**X**	**19.17**	**5.32**	**1.84**	**6.13**
Ibuprofen		**X**	**12.30**	**29.13**	**30.68**
APAP			**X**	**10.09**	**16.10**
Vitamin B12				**X**	**4.47**
Vitamin C					**X**
ISP					
Aspirin	**X**	**3.41**	**4.01**	**5.09**	**15.03**
Ibuprofen		**X**	**6.98**	**14.07**	**21.17**
APAP			**X**	**10.26**	**16.13**
Vitamin B12				**X**	**12.94**
Vitamin C					**X**

Copyright Statement

Joseph Medendorp. Robert A. Lodder. *J. Chemometrics*. **2005**; 19: 533–542.

Section II: Acoustic Resonance Spectrometry

Chapter Two - Acoustic-resonance spectrometry as a process analytical technology for rapid and accurate tablet identification

Introduction

On December 4, 2003, the US Food and Drug Administration recalled 504 bottles of mislabeled Magno-Humphries, Inc, Dixon's, Acetaminophen, 325 mg Analgesic Tablets.[34] A rapid, nondestructive online method of analysis and comparison with labeling could have prevented the recall and disposal of the batch of mislabeled acetaminophen, thus sparing the company considerable cost and a major safety risk. Process analytical technology (PAT) encompasses the design and development of processes to guarantee a predefined quality of pharmaceutical materials at the end of the manufacturing process, as warranted by risk analysis.[3] These processes come in various forms, including multivariate data acquisition and analysis tools, and in-process and endpoint monitoring tools. With the PAT system in effect, it becomes possible to implement real-time product release and increase pharmaceutical automation, thus providing the environment for significant reduction in manufacturing and labeling accidents.

Acoustic-resonance spectrometry (ARS) is an underutilized PAT that could become an analytical method of choice for the physical characterization of some analytes in pharmaceutical manufacturing. The wide-ranging measurements that can be made by ARS include sample compaction and axial strain, deformation, hydration and drying endpoint, elasticity, molecular stacking, and homogeneity, making ARS a very descriptive method of sample analysis.[35] In addition, ARS provides a rapid and efficient way to nondestructively identify and quantify an analyte with no sample preparation.[36] It has thus far successfully analyzed tablets,[7,36,37] powders,[8,38-40] and semisolids and liquids,[8,41-42] and it has quantified bulk moisture levels in otherwise identical samples.[36] Unlike most methods of analysis using photons, in ARS acoustic waves penetrate centimeters or even meters of material. Currently, the industry standard for

tablet characterization and identification is high performance liquid chromatography (HPLC), which, in addition to requiring extensive preparation (tablet grinding, dissolution, and extraction), destroys the tablet during the analytical process.[43-44] Only a few samples are collected from each batch for HPLC analysis, but entire batches must be disposed of if there are a few failed samples. Through chemometrics, near-infrared (NIR) and acoustic spectrometry are now making their way into the pharmaceutical industry as alternatives to HPLC testing.[45-46] Though the identification of specific functional groups from AR spectra is beyond the scope of this research, ARS can easily be applied to the quantification of active pharmaceutical ingredient (API) or moisture in tablets because of the high correlation between AR spectral features and chemical composition. Chemical changes lead to changes in physical characteristics such as density, compressibility, and acoustic velocity. For example, the density of acetaminophen is 1.2083 g/mL, while the density of aspirin is 1.3571 g/mL. The velocity of sound, therefore, is much higher in aspirin than it is in acetaminophen, resulting in a different set of acoustic interactions between each tablet and the quartz rod. These different interactions permit the extraction of bulk physical properties from the AR spectra and the characterization of even subtle correlated chemical differences between tablet groups. The chemical/physical relationship linkage can also be exploited to differentiate between unique formulations of the same API, as numerous constituents contribute to the final formulation, including excipients, compression aids, binders, coatings, disintegrants, and emulsifying agents.

Vibrational techniques such as NIR and infrared spectrometry are very sensitive to the presence of water. In fact, water's absorption is so strong that it often overwhelms other signals of interest.[47] Much computation can be required to correct for or deconvolve spectra with a large

water absorption to extract other useful features.[47-53] Acoustic waves are not necessarily as sensitive to the presence of water and do not suffer from the same limited penetration depth in water that causes problems in the NIR. ARS can be used along with NIR spectrometry to decrease interferences from water and increase the range of analytes that can be determined with good sensitivity.[54] The last few years have seen the advent of commercial instruments for acoustic analysis of liquid and semisolid samples.[55] Thus, this technology is rapidly gaining popularity for the physical characterization of analytes.

Sound is a longitudinal wave—one whose compressions and rarefactions oscillate parallel to the direction of propagation.[56] When an acoustic wave is applied to a sample, the medium responds by locally expanding and contracting, with particles in the medium drawing closer together and moving farther apart. The degree to which a particular medium responds is a product of its bulk modulus, or simply, its incompressibility. For dense materials with very little compressibility, a sound wave propagates very rapidly, while for less dense samples, sound travels more slowly.

Figure 2.1 shows a schematic of the AR spectrometer used for the following discussion. In the absence of a tablet at the vertex of the quartz rod, the applied acoustic signal received at the detector is a standing wave that is characteristic of the quartz wave guide. When a tablet comes in contact with the rod, the acoustic waves propagate through the tablet/quartz interface and pass to the tablet holder (which may contain a second transmitting transducer). Waves are transmitted through the tablet or reflected back through the tablet, where they reenter the quartz rod. The two sound paths lead to a pattern of in-phase and out-of-phase interferences (between the standing wave traveling through the quartz rod from piezoelectric transducer (PZT) to PZT with no tablet

interaction, and the wave propagating through the tablet) that is characteristic of the tablet.[57] This phenomenon is illustrated in an animation in Figure 2.2. The quartz rod was chosen as the means of propagation because the speed of sound is significantly larger in the rod than in the tablet; thus, the difference in delay time between the reference channel and the tablet channel is maximized. A rod composed of quartz also helps keep acoustic impedance to a minimum, according to Zac = P/(vA), where Zac is acoustic impedance, P is pressure (which is proportional to the wave amplitude), v is the velocity of the sound wave, and A is a unit of area.[56]

The next-generation AR spectrometer, which is already in construction and testing, will use integrated sensing and processing (ISP), encoding the excitation signal source so the detector output is directly proportional to the analyte concentration or classification (such an excitation signal is also called an ISP waveform). When the most distinguishing frequencies are preselected for excitation, the ISP excitation source can be an encoded weighted frequency spectrum, eliminating the need for postcollection computation and simplifying the AR spectrometer for use as a PAT sensor. Thus, an ISP excitation waveform is one in which each frequency in the frequency-domain representation of the waveform is weighted (typically between +1 and –1) so that the sum of the signals over the frequency and the integration time is proportional to the property of interest. ISP has demonstrated significantly reduced analysis times compared with the full-spectrum approach and appears to perform consistently with full-spectrum ARS for analyte separation and API quantification.[7,8] The utility in the ISP approach lies in its ease of application. For example, to differentiate between two visually indistinguishable groups of tablets, one can download from an online database an ISP waveform that is a prerecorded MP3

file written to distinguish between the required tablets. The final PAT sensor can be very simple when the waveform is recorded to a CD or an MP3 player and "played" through the tablets.

The current work represents a step in the direction of ISP-ARS. This research was performed to test the hypothesis that an AR spectrometer using ISP-compatible waveforms is able to rapidly and accurately differentiate tablets of similar size and shape. An ISP-compatible waveform is an excitation signal in which each frequency in the frequency-domain representation of the waveform is weighted +1. In an ISP calibration process, the calibration samples are usually tested first with an ISP-compatible waveform to determine the distinguishing frequencies for the actual ISP encoded waveform. In this research, an FM radio receiver supplies one source of broadband random excitation for ARS. Such a waveform could easily be recorded in MP3 format or on a CD for use in a high-throughput PAT sensor.

The ultrasound literature indicates that buffer rods (steel, aluminum, quartz, etc) coupled to PZTs are commonly used as probes to monitor various online processes, including fabrication of microfluidic devices,[58] polymer processing,[59] and inline die casting.[60] A typical industrial tablet press is capable of pressing slightly more than one tablet per punch station per second.[61] Each punch station can easily be fitted with a PZT, much like the buffer rod probe. With one second per punch, there is ample time to collect an acoustic signal from the tablet as it is being punched. Such technology would be capable of testing 100% of the tablets passing down the manufacturing line; therefore, no scale-up would be required to reach production rates. Thus, ARS is a high-throughput assay suitable for 100% product testing.

Theory

ARS spectra collected at *n* frequencies can be expressed as points in an *n*-dimensional hyperspace. These spectra usually form complex patterns in that hyperspace as analyte concentration changes because of the nonlinear nature of acoustic interactions.[62] Nevertheless, a plot of canonical correlation coefficients calculated from the principal component scores shown in Figure 2.3 illustrates that spectra from similar tablets tend to cluster in the same region of hyperspace.[9,63] However, in order to prove this assertion more quantitatively, the BEST nonparametric cluster analysis algorithm can be employed, and multidimensional standard deviations (MSDs) between clusters and spectral data points can be calculated.[30] To calculate the distances in MSDs, a population **P** is created as an *m x n* matrix in hyperspace **R** whose rows are the individual samples and the columns are the frequencies. The BEST method considers each frequency from a spectrum of *n* frequencies to be taken as a separate dimension[64] **P*** is a discrete realization of **P** based on a calibration set **T** of the same dimensions as **P***. This realization is chosen one time from **P** to approximate all possible sample variations present in **P**. **P*** has parameters **B** and **C**, where **C** = E**(P)** and **B** is the Monte Carlo approximation to the bootstrap distribution. The expectation value, E**(P)**, is the center of **P**, and **C** is a row vector with the same number of rows as there are columns in vector **P**. New test spectra **X** are projected into **R** containing **B**; rows of **B** are mapped onto a vector connecting **C** and **X**. **C** and **X** have the same dimensions. The integral over **R** is calculated from the center of **P** in all directions. A skew-adjusted MSD is based on the comparison of the expectation value **C**=E**(P)** and **C**=med**(T)**, the median of **T** in hyperspace projected on the hyperline connecting **C** and **X**. The result is an asymmetric MSD that provides two measures of the MSD along the hyperline connecting **C** and **X**. Equation 2.1 defines the MSD in the direction of **X**, and Equation 2.2 defines the MSD in the

opposite direction. Skew adjusted MSDs can be used to calculate mean distances between spectra of different samples.

$$+\vec{\sigma}\left|\frac{\int_{0}^{+\sigma}\left(\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)\right.}{\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)}\right| = 0.34 \qquad \textbf{2.1}$$

$$-\vec{\sigma}\left|\frac{\int_{0}^{-\sigma}\left(\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)\right.}{\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)}\right| = 0.34 \qquad \textbf{2.2}$$

Materials and Methods

A simple schematic of the instrument is shown in Figure 2.1 and an animated illustration of its operation is shown in Figure 2.2.[36,65-66.] The acoustic range was limited to 0 - 4 KHz in the computational model for experiments to assure compatibility with the simple ISP strategy. Three different broadband acoustic noise sources were tested; a radio receiver (Model D-3150 radio receiver, Consumer Electronics Corp., Knoxville, TN, USA) tuned to a local unused frequency (94.9 MHz FM), a white-noise generating Zener diode amplifier circuit constructed for these experiments,[67] and a function generator (FG) (Model DS335, Stanford Research Systems, Sunnyvale, CA, USA). The PZTs (CUI Inc., Beaverton, OR, USA) were connected directly to the audio-amplifier leads inside the receiver, and due to the magnitude of the receiver signal, no pre- or post-amplification was necessary. Ideally, a white noise generator creates equal excitation at each frequency in its spectral range. However, even in the case of the best possible applied white noise, the PZTs responded much better at some frequencies than others. The

receiver and Zener white noise generators were compared with a FG programmed to output a swept 10 V signal from 10 Hz to 4 kHz in steps of 10 Hz. The following discussion concentrates on a brief comparison of the two most effective methods of generating the noise excitation, the FG and the FM receiver. Typical Fourier transform spectra of the received signal from both methods are shown in Figure 2.4. Spectra appear different between the two methods because the FG sweeps a signal from 10 Hz to 4 kHz in discrete steps of 10 Hz, while the FM receiver delivers essentially random noise excitation across the entire range from 0 – 4 kHz. Because the overall magnitudes of the signals per unit time at the receiving PZT were comparable between the two methods, the FG had to deliver a much more intense excitation per unit time at each of its frequencies, which increases the acoustic nonlinearities caused by mechanical interface discontinuities in the sampling system. To more effectively visualize the distinctions between AR tablet spectra, Figure 2.5 illustrates second derivatives of the logarithms calculated from the Fourier transform of the radio data.

The vertex of the quartz rod was in mechanical contact with the sample, causing it to resonate with the rod. Tablet samples were mounted on a scale (Model 3120, Health O Meter, Bridgeview, IL, USA) to ensure that the pressure on each sample was consistent and reproducible. The PZT detector output signal was collected via a 16-bit soundcard (Realtek AC97, Realtek Semiconductor Japan Corp., Yokohama, Kanagawa, Japan) for processing and analysis in Matlab 7.0.1 (The Mathworks Company, Natick, MA, USA). For each sample, 15 seconds of data were collected at a sample rate of 8 kHz and a Fourier transform was calculated. Using the spectra of the blank (quartz rod), a signal-to-noise ratio of 50/1 was calculated by dividing the magnitude of the largest peak by the standard deviation of its replicates. With the

sample rate of 8 kHz, the maximum frequency collected was 4 kHz, though there was very little signal above 2.5 kHz.

The objective of the initial experiment was to prove that the selected tablets of similar size and shape could be differentiated unequivocally. Tablets used were ibuprofen (Perrigo, Allegan, MI, USA), acetaminophen (Leiner Health Products, Carson, CA USA), aspirin (Wal-Mart Stores, Inc, Bentonville, AR, USA), vitamin C (Leiner Health Products, Carson, CA, USA), and vitamin B-12 (Weider Nutrition Group, Salt Lake City, UT, USA). Please see Appendix B for chemical structures of named compounds. In order to effectively capture small variations and inconsistencies inherent to different tablets, ten tablets were scanned from each group. The instrument was configured so the quartz rod applied 100 grams of pressure to each tablet for the duration of the scan. To eliminate any effects of instrument drift on the analysis, tablets were scanned in random order. Each tablet was removed from the sample holder prior to replicate scans, and the quartz rod was raised and repositioned between each subsequent scan.

Tablet mass was determined with a digital mass balance (Mettler BB244, Mettler Instrument Corp., Hightstown, NJ, USA), thickness was measured to ± 0.1mm with a Vernier caliper, and volume for density measurements was determined by water displacement. Figure 2.6 demonstrates the correlation between acoustic spectra and these physical properties. For initial data analysis, principal components and canonical correlation coefficients were calculated from the spectral data. Principal components were calculated by a singular value decomposition of matrix **A** according to Equation 2.3:

$$\mathbf{A} = \mathbf{USV} \qquad \textbf{2.3}$$

where **A** is the matrix of original spectra, **U** is the matrix of eigenvalues (scores), **S** is a diagonal matrix of singular values, and **V** is the matrix of eigenvectors (loadings).

In Equation 2.4, a multiple linear regression of **U** indicates which of the components have the strongest correlation to tablet thickness, mass, or density, **y**, where **a** is the y-intercept, **b** is a vector of regression coefficients, and **c** is the residual.

$$\mathbf{y} = \mathbf{a} + \mathbf{bU} + \mathbf{c} \qquad \textbf{2.4}$$

These components were used in a leave-one-out cross validation to determine how effectively the ARS predicted these physical properties from their acoustic spectra according to Equation 2.5, where σ^2 is the variance, $f_i(U_i)$ is the prediction of the model for the *i-th* pattern *m* in the training set, after it has been trained on the *m-1* other patterns.

$$\sigma^2_{LOO} = \frac{1}{m}\sum_{i=1}^{i=m}(y_i - f_i(\mathbf{U}_i))^2 \qquad \textbf{2.5}$$

These physical properties have a tightly knit relationship, thus their individual effects on acoustic spectra need not be isolated for the purpose of the following discussion. Using a multidimensional translation of tablet populations, the dynamic range of the ARS instrument for each of the properties listed above was established.[30]

To estimate detection limits for dynamic range calculations, ARS spectra between tablet groups (intertablet spectra) from the two tablet populations ($\mathbf{P_1}$ and $\mathbf{P_2}$) with the smallest MSD separation (aspirin and ibuprofen) were used as *m x n* matrices.[30] The columns of the matrices were averaged by Equations 2.6 and 2.7, giving two *1 x n* vectors.

$$\mathbf{P_{1ave}} = \frac{1}{m}\sum_{i=1}^{i=m}\mathbf{P}_{1i} \qquad \textbf{2.6}$$

$$\mathbf{P_{2ave}} = \frac{1}{m}\sum_{i=1}^{i=m}\mathbf{P}_{2i} \qquad \textbf{2.7}$$

A difference spectrum **X** was calculated from $\mathbf{P_{2ave}} - \mathbf{P_{1ave}}$. One population was spatially translated toward the other, $\mathbf{P_{Adjusted}} = \mathbf{y*X+P_2}$, where **y** (defined on the interval $\{0<y\leq 1\}$) started at zero, and increased in increments of 0.01 until $\mathbf{P_1}$ and $\mathbf{P_{Adjusted}}$ were inseparable. To estimate the smallest possible tablet thickness, mass, and density differences that could still be separated, the fraction of the final spatial translation was multiplied by the difference in physical properties.[30] The procedure was conducted separately using the two tablet groups with the smallest MSD separation (ibuprofen and aspirin) and the largest MSD separation (acetaminophen and vitamin C).

NIR data were collected from tablets concurrently with ARS for comparison. Spectra from 1100-2500 nm were collected with a NIR spectrometer (Technicon InfraAlyzer 500, Tarrytown, NY) interfaced to a computer (OptiPlex GXM 5166, Dell, Round Rock, TX, USA) running SESAME 3.1 (Bran+Luebbe, Norderstedt, Germany). The data set consisted of fifty tablets, ten of each type as listed above. To reduce room noise, external interferences, and thermal detector drift, the tablets were scanned inside the instrument drawer and scanned in random order. All NIR data were also exported to Matlab 7.0.1 for processing and analysis.

Results

The BEST MSDs for the AR spectra are shown in Table 2.1. The average intertablet MSD for the AR spectra using receiver-generated noise was 65.44, and the average for the FG was 38.65. Leave-one-out cross-validations were calculated to quantify the intratablet variation (variation in AR spectra for multiple scans of the same type of tablet). A distance in MSDs between a tablet and a spectral data cluster that is less than 3 is defined as inseparable (a 99.8% confidence limit). The average cross-validation MSDs were 1.92 for the receiver-generated noise and 2.2 for the FG, suggesting that the receiver-generated noise was better for maximizing intertablet variation and minimizing intratablet variation. While the FG and the radio provided different excitation signals, the magnitudes of the total integrated signals at the receiving PZT at any given time were comparable between the 2 methods. Therefore, the FG had to deliver a much larger excitation at each of its individual frequencies as the frequency was swept. Because of small mechanical interface discontinuities between the PZTs and the quartz rod, and between the rod and the samples, the larger excitation signal with the FG produced more acoustic nonlinearity effects in the spectra. This suggests that a CD/MP3 player with an excitation signal specifically tailored to match the frequencies and amplitudes of the most distinguishing spectral features will also be more effective than an FG. NIR spectrometry (with a 120-second sample scan integration time) provided a median MSD intertablet separation of 363.55 and a median intratablet separation of 1.96. These results suggest that NIR spectra are superior to AR spectra in differentiating among the tablet classes. However, it should be noted that the NIR integration time was almost 10× greater than the ARS integration time, so the distances in MSDs should be almost 3× greater for NIR spectrometry than for ARS based on the integration time alone. In many process sensing applications, the fact that NIR spectrometry produces slightly better results

will not justify the much higher cost of using it. It should be noted that both NIR spectrometry and ARS were perfectly accurate in all of their tablet classifications.

The strong correlation between AR spectra and tablet thickness, mass, and density is illustrated in Figure 2.6. ARS predicted tablet thickness with $r^2 = 0.977$, a standard error of estimate (SEE) of 5.23%, and a standard error of performance (SEP) of 6.17%. ARS predicted tablet mass with $r^2 = 0.977$, SEE = 5.73%, and SEP = 6.16%, and ARS predicted tablet density with $r^2 = 0.900$, SEE = 11.0%, and SEP = 12.2%. NIR spectrometry only slightly outperformed ARS for these measurements, with average leave-one-out cross-validation $r^2 = 0.998$, SEE = 1.30%, and SEP = 1.40% (data not shown).

Discussion

Versatility and Flexibility of ARS. ARS provides an extremely versatile instrument for process analytical technology. The same instrument and chemometrics can be applied to multiple sample types including powders,[8,38-40] semi-solids and liquids,[8,41-42]. Due to the large difference in acoustic velocities between a quartz rod and a sample, there is no practical minimum physical sample size in order for this instrument to function. The rod and excitation can always be scaled to the physical sample size. While the excitation frequencies passing through the rod and the sample are the same, because of the difference in acoustic velocity, the waveforms are almost always longer in the quartz. The sample pathlength is always longer than the reference pathlength. Regardless of the sample size, the waveforms will never be perfectly in phase as they recombine at the detector on the rod. This will result in a characteristic pattern of constructive and deconstructive interferences for every analyte in response to their difference in

thickness. Additionally, when the API and formulations are different chemically between samples, physical manifestations of the chemical differences, such as mass, density, and compressibility are observed in the AR spectra.

Speed of Method. A short experiment was conducted to determine the minimum scan time at which the radio-receiver ARS was still capable of separating tablets. Tablets were scanned again for 15 seconds. Rather than calculating the Fourier transform from the time domain data in its entirety, the FFT was calculated from the first 1 s and from the first 250 ms of data. Principal components and BEST MSDs were calculated from the shorter data sets and plotted to verify that tablet groups were still separable by the same success metrics. For 1 s of time domain data, the average intertablet MSD was 29.30 and the average intratablet MSD was 1.768. For 250 ms of time domain data, the average intertablet MSD was 19.51 and the average intratablet MSD was 1.71. These results suggest that while 15 seconds of data collection offers an advantage in separation, it is not necessary to collect more than 250 ms, an obvious advantage in PAT.

Freedom from Interferences. The main sources of potential interference for the ARS in this configuration are (1) RF cross-talk directly between PZTs, (2) acoustic waves propagating through the support structures of the instrument, (3) spontaneous additional noise bursts in the receiver (this source of interference is unlikely in the FG), and (4) inconsistent sample placement due to the manual loading procedure. Studies were conducted to identify and quantify these effects. (1) To assess the possibility of RF cross-talk, a PZT was suspended (no mechanical contact) 10 cm from the epoxy-fastened receiving PZT. A white noise signal was generated with the suspended PZT and collected with the epoxy-fastened PZT. No signal was recorded with the

sound card, suggesting that the electromagnetic shielding was sufficient on the PZTs, and that when the PZTs are fastened with epoxy to the quartz, they do not move unless driven by the excitation signal through mechanical contact. (2) Where possible, support structures were made of wood to dampen sound. Wood is made up of a cellular network of pores that convert sound energy into heat by frictional and viscoelastic resistance.[68] Because of the high internal friction created by the cellular pore network, wood has more sound dampening capacity than most structural materials (e.g., steel, aluminum, or glass). Because the PZT and the sample support structures were made of wood, there was less sound traveling through the beams than if they had been made of metal. Additional sound dampening was accomplished by coupling the quartz rod to the support structure through rubber grommets. Finally, the support structure was fixed in place and did not change during the course of the experiment, so any sound traveling through it was the same for all sample and reference measurements. (3) Spontaneous noise bursts (e.g., lightning) from the receiver-generated noise might be a problem if they appeared sporadically in some tablet scans and not others. Such noises were eliminated by the use of frequency modulation and VHF frequencies. (4) 280 tablets were scanned without an anomalous spectrum due to manual tablet loading. A contoured sample holder helped to maintain reproducible sample positioning.[36]

Limits of Detection. The multidimensional population translation experiment[30] described in the Materials and Methods section was used to estimate the theoretical limits of detection. Using the populations with the smallest multidimensional separation (ibuprofen and aspirin), translation of one cluster toward the other indicated that tablets with a 0.08mm difference in thickness, a 0.0046g difference in mass, and a 0.01658g/mL difference in density were no longer separable

by ARS. This corresponds to a translation of one cluster 90% of the distance across hyperspace toward the other before the two begin to overlap (<3 SDs, i.e., p=0.0013). Using the populations with the largest multidimensional separation (acetaminophen and vitamin C), translation of one cluster toward the other indicated that tablets with a 0.27mm difference in thickness, 0.0756g difference in mass, and 0.01157 g/mL difference in density are inseparable. This corresponds to a translation of one cluster 93% of the distance across space toward the other before the two begin to overlap. The dynamic range for these properties appears to be about a factor of 10 from these experiments.

ARS may be useful within a PAT paradigm of networked sensors throughout a manufacturing process. The possibility of 'integrating the sensing and processing' (ISP)[7,8] makes ARS a very attractive method for further investigation. An ISP-AR spectrometer now under construction employs a tailored excitation signal so the voltage at the detector becomes directly proportional to the desired analyte concentration or classification. When ARS becomes a standard sensor on a pharmaceutical manufacturing line, speed and accuracy of identification and quantification will be paramount. Collecting and processing full AR spectra for every sample can be too costly in terms of time, hardware, and computing power. ISP can pre-select and weight excitation using only distinguishing frequencies, obviating the need for a full spectrum approach. In this manner, the detector itself will provide the tablet identity signal without the need for post-collection chemometric interpretation. An ISP ARS can be constructed from a CD or MP3 player, a simple, rugged device perfect for PAT. The sensor could be easily reprogrammed to analyze new samples using CDs (or downloaded MP3s) with preprogrammed tracks consisting of calibration signals specifically written for different analytes and different properties.

Conclusion

This research suggests that ARS is a viable PAT method for high-throughput tablet identification and characterization due to its speed, performance, low cost, durability, versatility, and freedom from interferences. The FM receiver used as an excitation source in this research outperformed the function generator, demonstrating that the most effective PAT sensor could be a very inexpensive instrument.

Chapter Two Figures

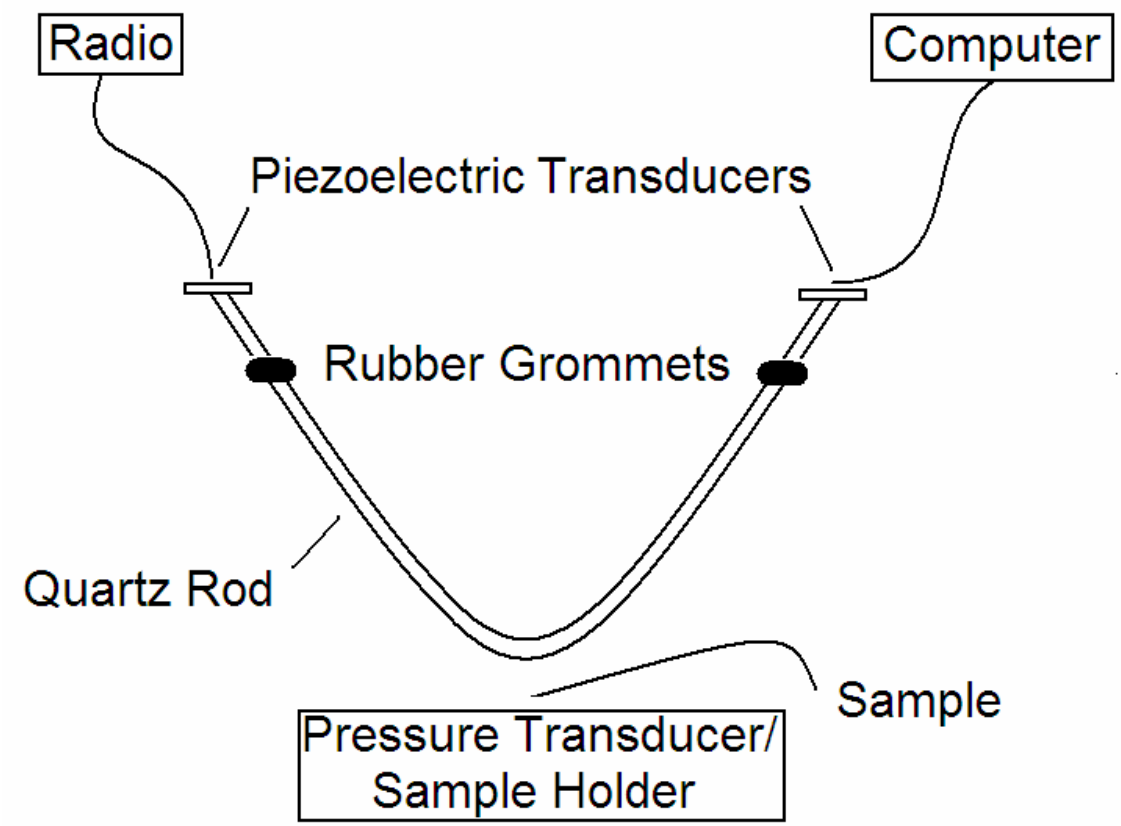

Figure 2.1: ARS schematic illustrating the instrumentation. The PZT on the left receives the excitation signal from the radio, while the PZT on the right receives the transmitted signal through the quartz rod.

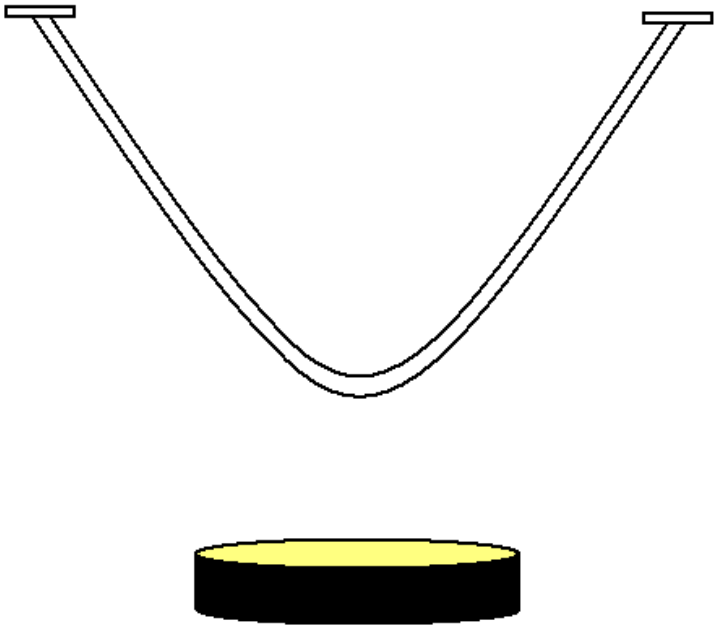

Figure 2.2: Animation demonstrating the principle by which the ARS operates. In the absence of a sample, excitation of the rod results in a standing wave and a consistent frequency spectrum. In the presence of a sample, a second sound path is introduced, passing through the sample and reflecting off the sample holder, and recombining with the original standing wave. This results in an entirely new wave, the interference pattern from the combination of frequencies from the standing wave and the sample wave.

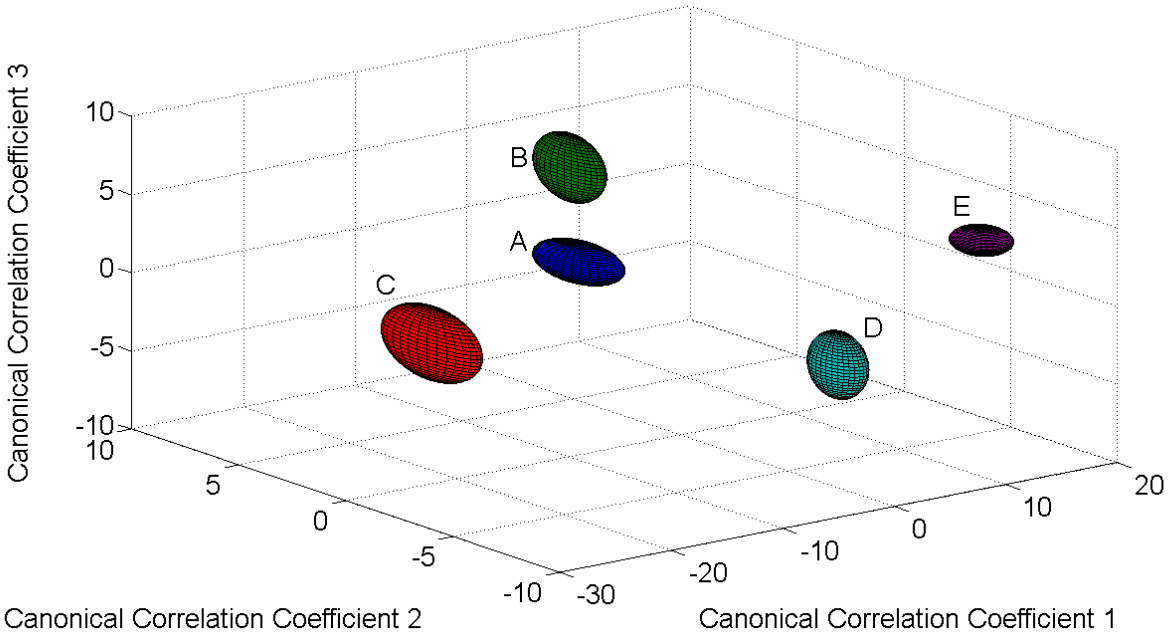

Figure 2.3: Illustrates the first three canonical correlation coefficients calculated from the principal component scores from the receiver noise source. Ellipses are drawn three standard deviations from the center of the clusters. (A) Ibuprofen, (B) aspirin, (C) acetaminophen, (D) vitamin B12, and (E) vitamin C.

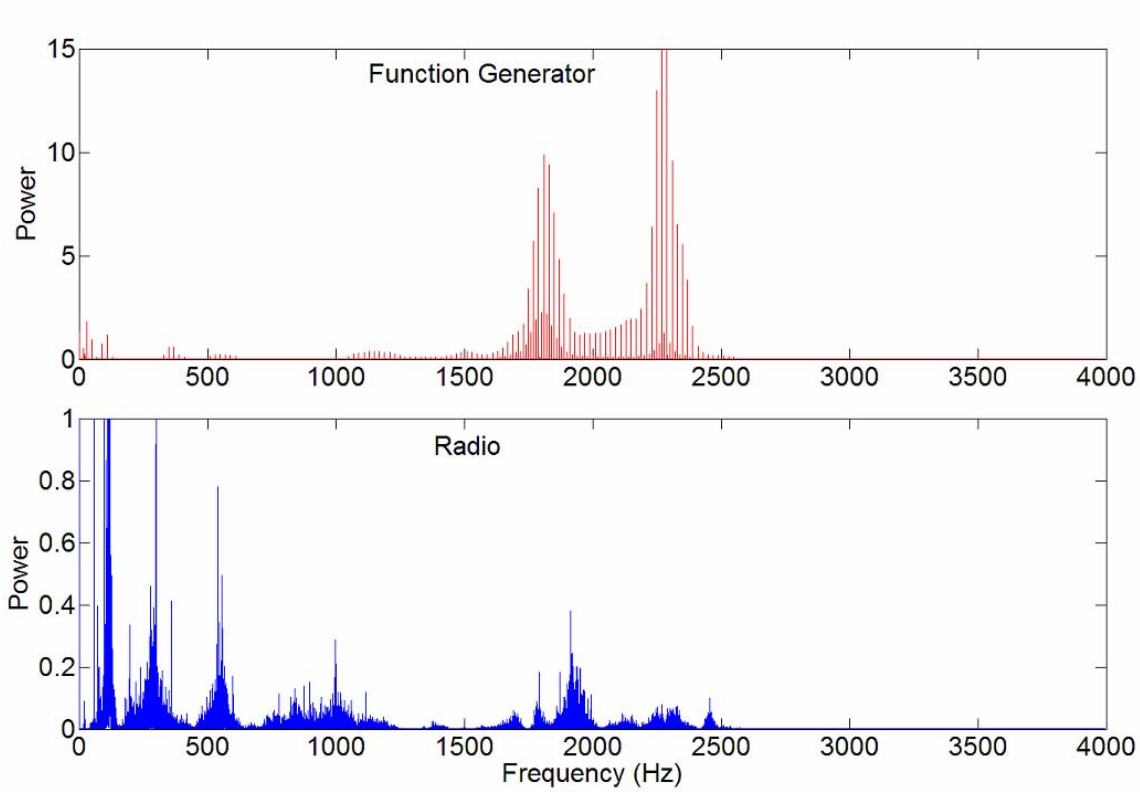

Figure 2.4: Fourier transform spectra comparison of the received signals between the FG and radio.

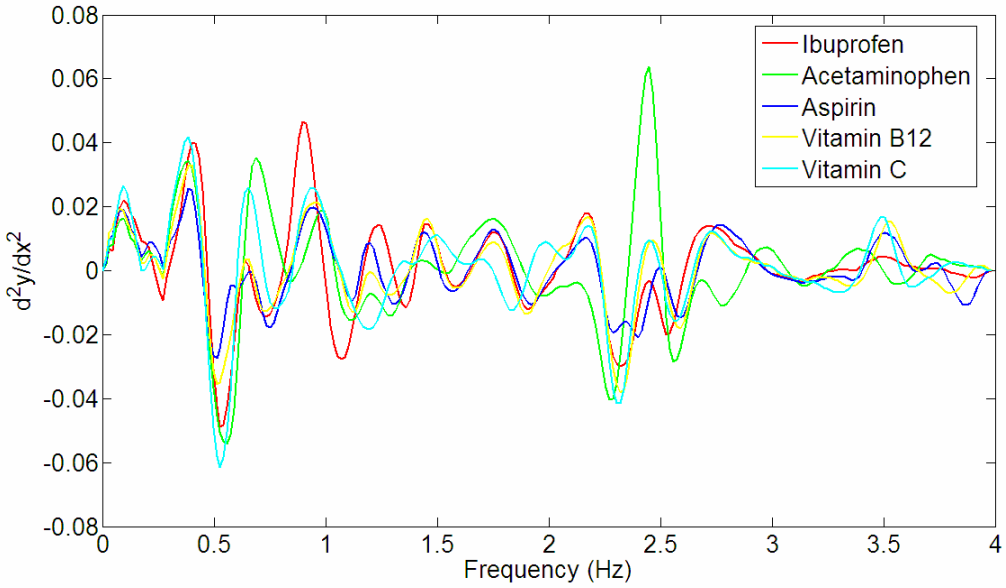

Figure 2.5: The log and second derivatives were calculated from the Fourier transform radio data to more easily visualize the distinct differences between the tablet groups.

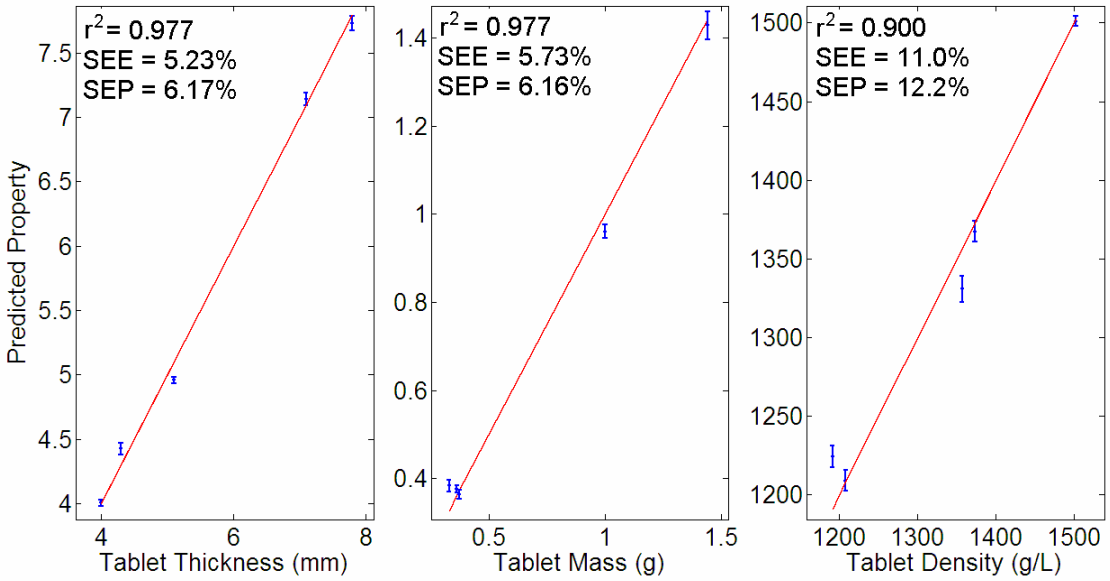

Figure 2.6: Cross validation results demonstrating the ability of ARS to predict tablet thickness, mass, and density from the acoustic-resonance spectra of the tablets.

Chapter Two Tables

Table 2.1: BEST multidimensional standard deviations (MSDs) between the tablet groups for the FM receiver (top) and the FG (bottom). Intragroup tablet MSDs are shown across the diagonal.

Radio MSDs	**Ibu**	**APAP**	**Asp**	**Vit B12**	**Vit C**
Ibu	**2.33**	62.29	13.82	74.82	77.61
APAP		**1.78**	50.08	77.34	105.22
Asp			**1.76**	74.77	83.57
Vit B12				**1.95**	36.94
Vit C					**1.76**

FG MSDs	**Ibu**	**APAP**	**Asp**	**Vit B12**	**Vit C**
Ibu	**2.88**	62.32	24.10	29.24	31.77
APAP		**2.52**	47.49	45.76	75.73
Asp			**1.80**	16.65	39.26
Vit B12				**1.74**	14.21
Vit C					**2.07**

Copyright Statement

Joseph Medendorp, Robert. A. Lodder. *AAPS PharmSciTech*. **2006**; 7(1): Article 25.

Chapter Three - Acoustic-resonance spectrometry as a process analytical technology for the quantification of active pharmaceutical ingredient in semi-solids

Introduction

Colloidal oatmeal (CO) is used as an ingredient in some pharmaceutical lotions, cosmetics, and toiletries. While listed as inert in some lotions but active in others, CO is one of many examples in cosmetics and other household products that fall into a gray area for the US FDA. Between November 2002 and March 2004, there were over 20 FDA recalls on household cosmetic products; shampoos, sprays, hand soaps, toothpastes, and lotions.[69] These recalls, due to mislabeling and active pharmaceutical ingredient (API) concentration errors, caused companies considerable unnecessary expense. A rapid and accurate in-process assay, capable of testing and validating individual samples, would obviate the need for the disposal of entire lots of problem products. The FDA process analytical technology (PAT) initiative calls for the development and implementation of manufacturing processes to guarantee a predefined quality of pharmaceutical materials as warranted by risk analysis.[3] These processes include multivariate data acquisition and analysis tools, and in-process and endpoint monitoring tools. The development of acoustic resonance spectrometry as an in-process and endpoint monitoring device is in harmony with the PAT initiative.

Current assays for analyzing API in various topical lotions are based on HPLC with UV/VIS detection,[70-71] FT-IR spectrometry,[72] and NIR spectrometry.[73] These conventional assays require both more sample preparation and more time for each scan than ARS. Sometimes lotions do not exhibit an optical chromophore[74] so they must be fixed with labeled isotopes.[71] In addition, optical methods provide poor penetration through semi-solids in comparison to acoustic methods.

ARS is typically much more rapid than HPLC and NIR. It is nondestructive and requires no sample preparation as the sampling waveguide can simply be pushed into the lotion. To date, the AR spectrometer has successfully differentiated and quantified sample analytes in various forms; tablets,[36] powders,[38-40] and liquids.[41-42,75] .It has been used to measure and monitor the progression of chemical reactions, such as the setting and hardening of concrete from cement paste to solid.[76] Acoustic spectrometry has also been used to measure the volume fraction of colloids in a dispersion medium, as well as for the investigation of physical properties of colloidal dispersions, such as aggregation and particle size distribution.[77-78] Typically, these experiments are carried out with sinusoidal excitation signals and the experimental observation of signal attenuation. From a comparison of theoretical attenuation to experimental observation, the particle size distribution and aggregation phenomena are inferred. In place of a sinusoidal excitation signal sweeping across the desired frequency range, this research makes use of broadband white noise and standing resonance waves. To the author's knowledge, this research presents the first application of a broadband white noise excitation signal to a resonant system for quantification of colloidal particles in a dispersion medium.

It must be noted that a description of colloid aggregation and particle-particle interaction is beyond the scope of this research. Colloidal oatmeal is a lyophilic colloid, and it is readily hydrated and dispersed evenly through a solution. Therefore, the resonant acoustic signal received at the detector is taken to be approximate representation of the bulk of the sample, regardless of microscopic differences between individual colloids. As with other acoustic studies, the individual colloid particles are considered to be spherical and uniform, thus each particle produces a uniform effect on the bulk physical properties of the surrounding medium.[77]

Figure 3.1 provides a simple schematic of the AR spectrometer for the following discussion. The AR spectrometer used in this research was built in the near-field configuration, where the wavelength of the excitation signal is much larger than the quartz rod or the sample that is applied to the rod. An acoustic signal is applied to one of the piezoelectric transducers (PZTs) and received at the other. The sample, which is in mechanical contact with the vertex of the quartz rod, constitutes a load on the resonant system. With no sample in mechanical contact with the waveguide connecting the PZTs, little excitation signal is lost and the acoustic signal received at the detector is the sum of an ensemble of standing waves. The typical AR spectrum results from the pattern of constructive and destructive interference between the two sound paths; one that travels down the quartz waveguide and through the sample and back on the way to the collecting PZT, and the other that stays in the quartz rod and has no sample interaction. The excitation signal is a broadband white noise source as illustrated in Figure 3.2, where all frequencies between 0-22.05 KHz are excited simultaneously. However, the wave collected at the detector is primarily composed of three large resonance structures as illustrated in Figure 3.3 (2.35-2.7KHz, 9.6-11.2KHz, and 14.15-17.0KHz). When a sample is placed in contact with the vertex of the waveguide, acoustic waves escape and propagate through the lotion/quartz interface and into the sample holder. The added mass effect causes a shift in the resonant frequency of the system, while the frictional or viscous drag force causes a reduction in peak amplitude. This pattern gives rise to the characteristic AR spectrum for any given analyte. In some cases, a third transmitting transducer beneath the sample may increase the analytical signal,[36] although it was not required in this research.

Because the propagation of an acoustic wave is based on longitudinal compressions and rarefactions of the medium through which the sound propagates, each analyte responds differently to the frequency and amplitude of the applied acoustic signal. Acoustic spectra are the direct manifestation of the physical differences between samples such as density, viscosity, acoustic velocity, the degree to which samples and analytes compress and expand in response to an applied acoustic excitation, as well as depth of penetration, and the contact pressure between the resonator and the sample.[79-81] These effects are seen in the acoustic spectra as a shift in the resonant frequency of the system, a reduction or an increase in the peak amplitude, or a mixture of both effects. Chemical changes in samples affect the physical properties measured in acoustic spectra. Figure 3.4 illustrates the shifts in resonance frequency and changes in peak height between acoustic spectra from four different analytes; water, air, metal, and lotion. The forces acting on the standing acoustic wave come from the inertial effect as the sample is moved by the vibrating quartz rod, and from the dissipative effect due to the viscous drag force. The quartz rod was chosen as the waveguide due to its high characteristic sound velocity and because quartz is an electrical insulator, which prevents electromagnetic standing waves in the waveguide. The velocity keeps acoustic impedance to a minimum according to $Z_{ac} = P/vA$, where Z is acoustic impedance, P is sound pressure, v is sound velocity, and A is cross sectional area.[57,66]

Chemometrics . This research compared two different calibration methods, net analyte signal (NAS) and principal component regression (PCR). PCR has been described previously.[82] For NAS calibration, a vector of instrumental responses $\mathbf{r}_k$ is the sum of two independent signals, the signal from all interferences $\mathbf{r}_k^{=}$, and the signal from the analyte of interest $\mathbf{r}_k^{\perp}$, which is orthogonal to the contribution from the interferences.[22] This orthogonal portion is termed the net

analyte signal (NAS), and is the portion of the signal used for multivariate calibration. A matrix **R** *(J x I)* without the analyte of interest must be available, where *J* is the number of wavelengths and *I* is the number of samples. A projection matrix $\mathbf{P}_k^{\perp}$ can be calculated according to Equation 3.1:

$$\mathbf{P}_k^{\perp} = (\mathbf{I} - \mathbf{R}_{-k}\mathbf{R}_{-k}^{+}) \qquad \textbf{3.1}$$

where **I** is the identity matrix, and the '+' superscript indicates the Moore-Penrose pseudoinverse. Using one-point calibration with spectrum $\mathbf{r}_{cal}$, the NAS vector $\mathbf{r}_{cal}^{\perp}$ can be calculated with Equation 3.2.

$$\mathbf{r}_{cal}^{\perp} = \mathbf{P}_k^{\perp}\mathbf{r}_{cal} \qquad \textbf{3.2}$$

This vector is then normalized to length one with Equation 3.3:

$$\mathbf{r}_k^{NAS} = \frac{\mathbf{r}_{cal}^{\perp}}{\left\|\mathbf{r}_{cal}^{\perp}\right\|} \qquad \textbf{3.3}$$

The slope of the calibration line is calculated from Equation 3.4:

$$s = \frac{\left\|\mathbf{r}_{cal}^{\perp}\right\|}{c_{cal}} \qquad \textbf{3.4}$$

where s is the slope and c_{cal} is the analyte concentration of the calibration spectrum. The first step of NAS calibration includes using Equations 3.2-3.4 to find the NAS direction and determine the length of the NAS vector. When the NAS direction and magnitude are known, the unknown spectrum $\mathbf{r}_{un}$ can be projected in the NAS direction with Equation 3.5 and its magnitude compared to the calibration magnitude.

$$y_{un}^{NAS} = \mathbf{r}_{un}^{T}\mathbf{r}_k^{NAS} \qquad \textbf{3.5}$$

Equation 3.6 can now be used to calculate c_{un}, the unknown analyte concentration.

$$c_{un} = \frac{1}{s} y_{un}^{NAS} \qquad \textbf{3.6}$$

The NAS approach allows for the calculation of figures of merit from multivariate data sets. In severely overlapping spectra, it has historically been difficult to quantify selectivity, sensitivity, and signal-to-noise (S/N) because of the inability to distinguish between interferences and the analyte of interest.[19-20] With the NAS, these quantities can be measured directly. Selectivity is defined as the scalar degree of overlap, α, between the NAS vector and the calibration spectrum according to Equation 3.7:

$$\alpha = \frac{\left\| \mathbf{r}_{cal}^{\perp} \right\|}{\left\| \mathbf{r}_{cal} \right\|} \qquad \textbf{3.7}$$

The selectivity is a measure from 0 to 1 indicating how unique the analyte of interest is compared to the interferences. The sensitivity is a measure of the analyte variation in response to a change in concentration. This quantity can be expressed as Equation 3.8:

$$\mathbf{s}_k = \mathbf{r}_k^{NAS} / c_k \qquad \textbf{3.8}$$

where c_k is the concentration of the k-th analyte. Sensitivity should be the same for each concentration and each NAS vector.[21] The S/N ratio can be expressed as Equation 3.9:

$$S/N = \frac{c_k \left\| \mathbf{r}_k^{NAS} \right\|}{\left\| \boldsymbol{\varepsilon} \right\|} \qquad \textbf{3.9}$$

where $\boldsymbol{\varepsilon}$ is the random instrumental error.

ISP-ARS for PAT. Integrated sensing and processing (ISP) is a paradigm for instruments in which large physical fields of data are reduced to high-level information as the data are sensed, before the data are passed to a computer.[83] In many cases, the signal transmitted from the detector to the computer is directly proportional to the analyte concentration, or is a classification

that represents the sample identity. Using ISP, little or no post-collection data processing is required, making analyses significantly faster and perhaps even more accurate. An imminent AR spectrometer, already under construction and validation, employs ISP by encoding a special excitation signal to resemble the spectral features required for the quantification of colloidal oatmeal in lotion, so the signal received at the detector is directly proportional to the analyte concentration without further processing. ISP-ARS has great potential as a process analytical technology (PAT) for sensing in the pharmaceutical industry. The excitation can be specifically tailored to meet the needs of demanding analytical challenges. For example, if the need is to differentiate between 1% and 5% colloidal oatmeal creams, an ISP waveform created for this purpose could be easily downloaded from an online database in the form of an .MP3 file. The hardware of the ISP-ARS comprises inexpensive COTS (commercial off-the-shelf) components, simplifying its manufacture and deployment.

Materials and Methods

Sample Preparation. Colloidal oatmeal samples were prepared using Gold Bond Sensitive Lotion (Lot #03518). Lotion was weighed into an 800 mL beaker. Five concentrations (2.0 %, 2.5 %, 3.0 %, 4.0 %, 5.0 %) of CO were prepared by gravimetric addition of CO to the lotion samples for a total sample mass of 350g. Samples were heated to 50^{o}C on a hot plate while mixing with a paddle blade. When lotions reached 50^{o}C, the CO (Vendor Lot #22915) was added and the sample was mixed (Heidolph RZR 50, Frankfort, Germany) at a sufficient RPM to create a gentle vortex. The heat was then turned off and samples were mixed until they cooled to 35^{o}C.

ARS Data Collection. Rubber grommets were positioned on the rod so that they firmly held it in place, but dampened the signal minimally. Previous investigations with a similar ARS model made use of a third transducer beneath the vertex of the rod, which acted as an interferometer.[54] Our investigation indicated that the interference pattern inherent to the two sound paths and its effects on the resonance frequency and peak amplitudes was sufficient to negate the need for an additional transducer. A white noise signal played through an ordinary MP3 player was used as the source of random noise across the range from 0-22.05 KHz. The ARS instrument was designed to output a voltage signal to a computer sound card (Realtek AC97, Realtek Semiconductor Japan Corp., Yokohama, Kanagawa, Japan), operated with a graphic user interface in Matlab 7.0.1 (The Mathworks Company, Natick, MA, USA). Lotion samples were placed in a small plastic cup and the vertex of the quartz rod was plunged into the cup so the two were in mechanical contact for the duration of the scan. The penetration depth of the rod into the lotion was kept constant at 5mm. All data were collected for three seconds at a sample rate of 44.1 KHz. Data were transformed from the time domain into the frequency domain by an FFT, resulting in a signal-to-noise ratio of 110/1. Because the data set dimensions are a product of the sample rate and duration of the collection, each spectrum collected with these instrument parameters was 132,300 data points long. A twenty-point moving boxcar average was performed on the frequency domain data as a smoothing function. In its current configuration, there are significant resonance peaks at 2.35-2.70 KHz, 9.6-11.2 KHz, and from 14.5-17.0 KHz. All data analysis was performed on these three frequency regions for the duration of the experiment.

NIR Data Collection. A layer of lotion approximately 1 mm thick was spread on a single-well depression microscope slide (Gold Seal Products, Portsmouth, NH, USA). NIR scans were collected with a scanning monochromator NIR instrument described previously.[84] Scans were collected from samples inside light-proof chamber in a darkened room to eliminate stray light interference. Samples were scanned in random order to eliminate the effects of drift over time on the results. All data were exported to Matlab 7.0.1 for all processing.

Data Analysis. The objective of this experiment was to test ARS in lotion analysis, and compare the results with NIR spectrometry for the measurement of concentration change of API in lotion. NIR data were multiplicative scatter-corrected to eliminate baseline variations due to pathlength differences.[85] One-point NAS calibration was compared with principal component regression (PCR)[82] for the measurement of CO from ARS and NIR spectra. Detection limits were estimated by a BEST multidimensional population translation of one cluster toward another.[32] The clusters used in the translation were unprocessed AR FFT spectra and PC scores. Other figures of merit such as selectivity, sensitivity, and signal to noise ratio were calculated using the net analyte signal (NAS). All algorithms were written by the authors.

Results and Discussion

ARS vs. NIR. The results from the calibration models are summarized in Table 3.1. NAS calibration outperformed PCR for both NIR and ARS. NIR for the measurement of CO gave an r^2 prediction accuracy of 0.971, and a standard error of performance (SEP) of 0.517 %CO. ARS for the API resulted in an average $r^2 = 0.983$, calculated by averaging the r^2 from the three resonance peaks, and an SEP = 0.317 %CO. Each NIR scan took almost 90 seconds to complete

(the long integration time was selected to get a S/N competitive with ARS), while each ARS scan required only a maximum of three seconds to achieve similar sample predictive ability.

ARS Detection Limits. ARS spectra collected at *n* frequencies can be reduced to a single point in an *n*-dimensional hyperspace using the bootstrap error-adjusted single sample technique (BEST). The resulting points usually form complex Lissajous patterns (or curved portions thereof) in hyperspace as analyte concentrations change due to the nonlinear nature of acoustic interactions.[62] BEST is a nonparametric cluster analysis algorithm based on the premise that spectra from similar samples tend to cluster in the same region of multidimensional hyperspace.[30] In order to demonstrate clustering quantitatively, multidimensional standard deviations (MSDs) can be used to measure the separation between clusters of different samples.[30,32,54] Using the BEST, intercluster MSDs greater than three indicate distinct cluster populations, while clusters less than three MSDs apart are inseparable with statistical significance. A multidimensional translation operation can be performed with this algorithm to estimate the theoretical detection limits of the instrument.[32] The translation operation is based on the principle that clusters from pure component spectra form distinct and separate populations, and a translation of one toward the other corresponds to a mixture of the two components when Beer's Law holds.

To estimate the theoretical detection limits for dynamic range calculations, intragroup spectra from the two lotion populations ($\mathbf{P}_1$ and $\mathbf{P}_2$) were used as *m x n* matrices.[32] The columns of the matrices were averaged by Equations 3.10 and 3.11, giving two *1 x n* vectors.

$$\overline{\mathbf{P}}_1 = \frac{1}{m}\sum_{i=1}^{i=m}\mathbf{P}_{1i} \qquad \textbf{3.10}$$

$$\overline{\mathbf{P}}_2 = \frac{1}{m}\sum_{i=1}^{i=m}\mathbf{P}_{2i} \qquad \textbf{3.11}$$

A difference spectrum $\mathbf{X}$ was calculated from $\overline{\mathbf{P}}_2 - \overline{\mathbf{P}}_1$. One population was spatially translated toward the other, $\mathbf{P}_A = \mathbf{yX} + \mathbf{P}_2$, where $\mathbf{y}$ (defined on the interval $\{0<y\leq 1\}$) started at zero, and increased in increments of 0.01 until $\mathbf{P}_1$ and $\mathbf{P}_A$ were inseparable by the BEST metric. The detection limits were estimated by the BEST multidimensional translation experiment using raw AR spectra in the selected frequency regions and PC scores calculated from the selected frequency regions. The limits were calculated using translation of two clusters, the translation from 2%CO to 2.5%CO, and the translation from 2%CO to 5%CO, all results are summarized in Table 3.2. The best detection limits were seen using PC scores rather than raw AR spectra. This is an expected result as PC scores are inherently the representation of the largest difference between spectral groups. Therefore, separation has been optimized prior to estimation of the detection limits by PC scores.

Speed of ARS Method. To assess the maximum speed at which the ARS is still capable of quantifying colloidal oatmeal, the FFT was calculated for progressively shorter blocks of time rather than calculating the FFT from the time domain data in its entirety. For example, the FT was calculated for three seconds, two seconds, and one second. NAS calibration and a PC regression were performed on the FFT. The speed of method results are presented in Table 3.1. These results suggest that longer periods of data collection offer an advantage in quantification, though it is not necessary to collect more than 1s for a high predictive ability. There is a trade-

off between scan time and performance. If the experimenter can spend the time to collect only 1s of data, the process is much faster but the predictive ability of the instrument suffers slightly.

Selectivity, S/N Ratio, and Sensitivity of ARS. Estimation of figures of merit is a straightforward task when the measured response is a scalar, such as in zeroth-order calibration methods. The measured response in zeroth-order methods is simply the contribution of the analyte plus a constant instrumental background or constant interference. However, in first-order data where the measured response is a vector of scalars, the total instrument response is not sufficient to determine the figures of merit. The presence of multiple chemical and instrumental interferences varying simultaneously requires an extra step for multivariate calibration. Therefore, the portion of the signal orthogonal to the interfering chemical species and background, termed the net analyte signal (NAS), must be determined prior to calculation of figures of merit.[19-22] Because the ARS functions by the interference between the analyte acoustic signal and a standing wave, there is usually a substantial instrument response even in the absence of a sample. For this reason, calculation of the NAS with acoustic data tends to behave much like an additional smoothing function. The NAS and figures of merit were calculated according to Equations 3.7-3.9. Selectivity is a unitless scalar ratio, indicative of the degree of overlap between the two vectors, and was calculated to be 0.5963. The signal-to-noise ratio (S/N) was 110/1. The sensitivity of the ARS for colloidal oatmeal in lotion is 0.230, given in instrumental response per unit concentration change.

Versatility and Flexibility. In addition to being rapid, effective, and inexpensive, the ARS has also proven to be an extremely versatile instrument. The same instrument and chemometrics can

be applied to multiple analytes including tablets,[36] powders,[38-40] and liquids.[41-42,75] Because the system is set up such that a standing wave forms in the quartz rod in the absence of a sample, the addition of an analyte to the vertex of the rod acts as a load on the resonance of the system and necessarily disrupts the pathlength of sound in the waveguide. While the excitation frequencies passing through the rod and the analyte are the same, because of the large differences in their velocities, the wavelengths are usually longer in the quartz than in the sample. As a result, regardless of the colloid size, shape and concentration, the waveforms from the waveguide and the sample will never be perfectly in-phase as they recombine at the vertex of the rod. This phase change will result in a characteristic pattern of constructive and deconstructive interferences for every analyte because of differences in physical properties.

Freedom from Interferences. The main sources of interferences for the ARS in this configuration are (1) acoustic waves propagating through the support structure of the instrument, (2) RF cross-talk directly between PZTs, and (3) inconsistent analyte and quartz rod alignment due to a manual loading procedure. Each of these points is addressed in the following discussion. (1) The quartz rod was mounted on a frame of wood, which is made up of a cellular network of pores that convert sound energy into heat by frictional and viscoelastic resistance.[68] The cellular pore network creates high internal friction, thus wood has more sound dampening capacity than most structural materials. Because the mounting framework is constructed from wood, there is less sound traveling through the beams, and any sound that may get through is dampened due to the placement of the rubber grommets holding the quartz rod in place. It should be noted that the wood construction for the purposes of this experiment was the low cost alternative to other potential materials. For example, if this instrument were needed in a cGMP environment for

PAT, it could just as easily be constructed from other acoustic dampening materials such as concrete and rubber. (2) To assess direct RF cross-talk, a PZT was suspended 10 cm from the epoxy-fastened receiving PZT. A white noise signal was generated with the suspended PZT and collected with the epoxy-fastened PZT. No signal was recorded with the sound card, suggesting that when the PZTs are fastened with epoxy to the quartz, they do not move unless driven by the excitation signal through mechanical contact. (3) More than 90 lotion samples have been scanned without an anomalous spectrum due to manual loading. Because the position of the quartz rod is regulated by a vertical translation stage, the depth of penetration into the lotion is consistent; therefore, spectra are reproducible, demonstrating the durability of the ARS.

Conclusion

This research demonstrates that acoustic resonance spectrometry is faster, less expensive, and outperforms near infrared spectrometry for the quantification colloidal oatmeal in lotion. Therefore, it has demonstrated its potential as a process analytical technology for the quantification of active pharmaceutical ingredient in semi-solids.

Chapter Three Tables

Table 3.1: Three seconds of data collection gave the highest r2 and the lowest RMSEP for each frequency range and both chemometric models.

Method	Frequency (KHz)	Analysis Time (s)	NAS r^2	NAS RMSEP (% CO)	PCR r^2	PCR RMSEP (% CO)
	2.35-2.70	1	0.892	1.532	0.926	0.306
		2	0.985	0.787	0.943	0.260
		3	0.991	0.504	0.946	0.265
ARS	9.60-11.20	1	0.962	0.491	0.855	0.427
		2	0.973	0.299	0.879	0.391
		3	0.983	0.251	0.948	0.255
	14.5-17.0	1	0.968	0.398	0.813	0.511
		2	0.979	0.262	0.816	0.495
		3	0.991	0.180	0.855	0.488
NIR	1100-2500nm	120	0.971	0.517	0.956	0.314

Table 3.2: Theoretical detection limits as estimated by the BEST population translation. Translations were performed with the two closest concentrations and the two most distant concentrations.

	Raw AR Spectra (Detection Limit %CO)		PC Scores (4 PCs) (Detection Limit %CO)	
Frequency (KHz)	2% → 2.5%	2% → 5.0%	2% → 2.5%	2% → 5.0%
2.35-2.70	0.210	1.26	0.100	0.150
9.0-11.20	0.175	1.23	0.050	0.210
14.5-17.0	0.235	1.41	0.160	0.090

Chapter Three Figures

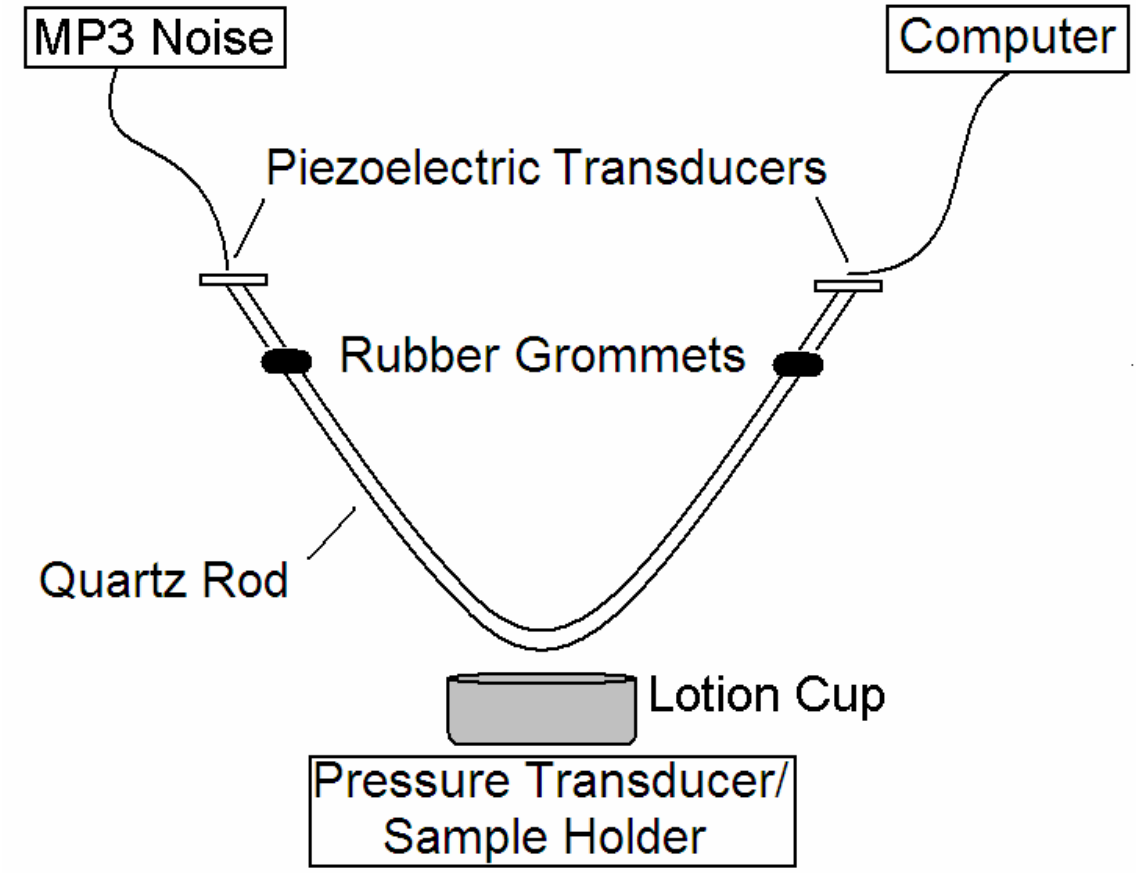

Figure 3.1: A schematic diagram of the ARS instrument.

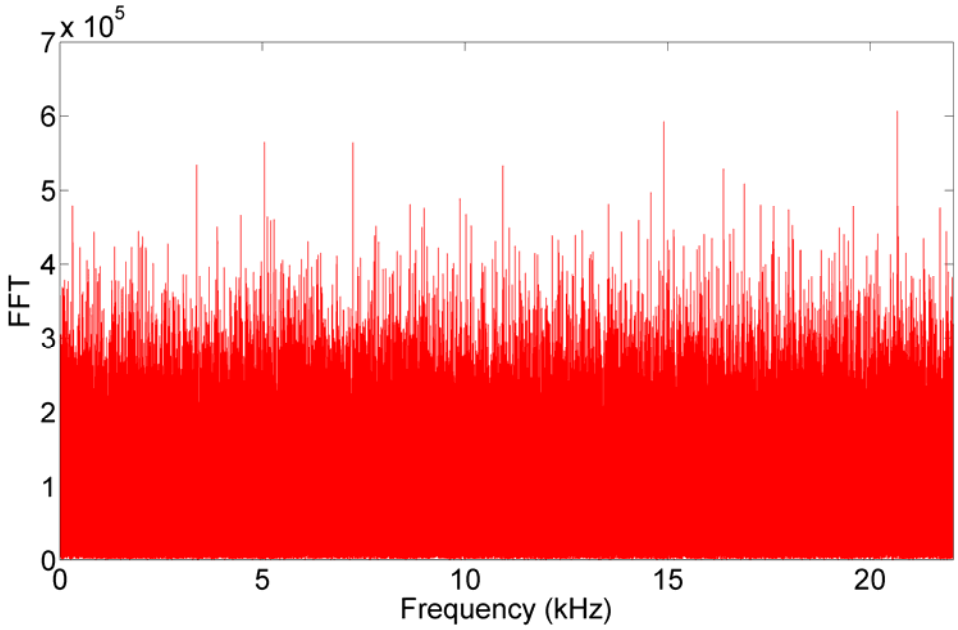

Figure 3.2: Frequency domain representation of the white noise excitation signal source delivered from an MP3 player to the PZT. All frequencies between 0-22.05KHz are excited simultaneously.

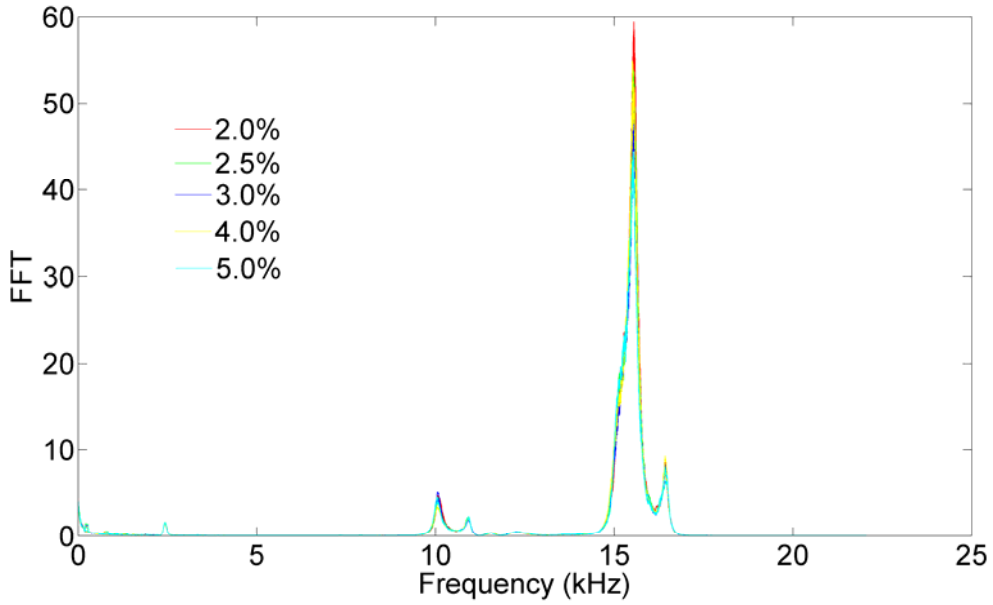

Figure 3.3: Three large resonance structures are prevalent for this design of the AR spectrometer despite excitation with a broadband white noise source. This figure illustrates AR scans of the different colloidal oatmeal concentrations in lotions.

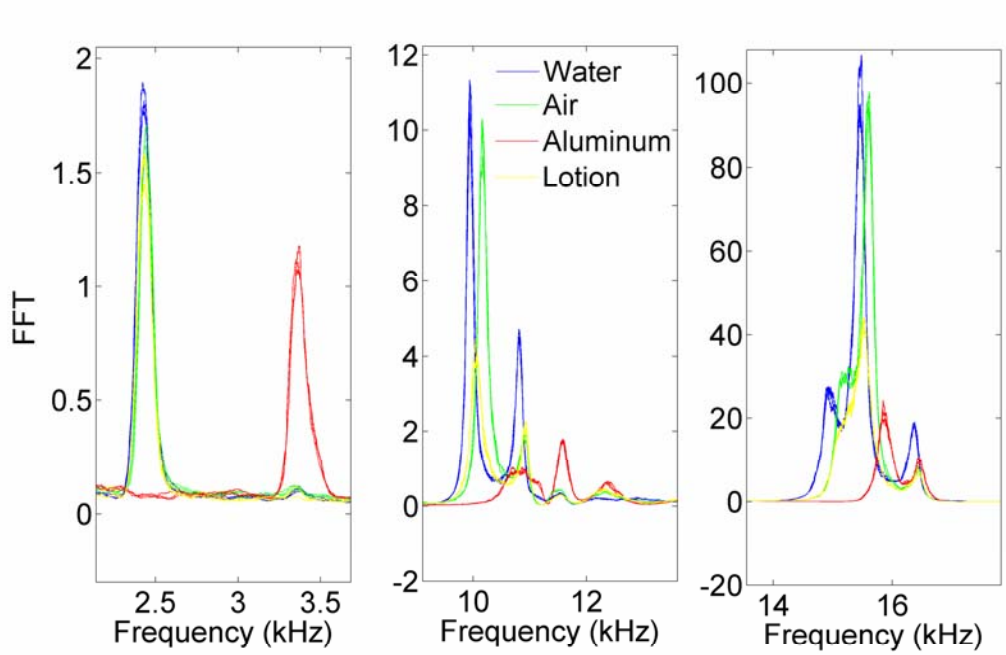

Figure 3.4: A zoomed view of the three large resonance peak structures considered in this study.

Copyright Statement

Joseph Medendorp, Robert G. Buice, Jr., Robert. A. Lodder. *AAPS PharmSciTech*. **2006;** 7(3): Article 59.

Section III: Near Infrared Spectrometry

Chapter Four - Near-infrared spectrometry for the quantification of dermal absorption of econazole nitrate and 4-cyanophenol

Introduction

Application. Methods to determine the bioequivalence of drugs from topical administration have been discussed formally by the FDA for many years.[86] A vasoconstrictor assay measuring skin blanching has been used to determine the bioequivalence of topical steroids.[87] This steroid blanching is the only topical bioequivalence test that has been agreed upon among the majority of the topical drug delivery scientists. Punch biopsies of the dermis and epidermis have been done to determine skin drug concentrations in humans in vivo, but this is very invasive and leaves permanent scarring. Sampling of just the stratum corneum using tape stripping has generated the most interest of all the current methods under investigation, however, this method has many variables and intricacies that result in inter- and intra- lab variation in the results.[88] One problem noted with tape stripping is that chemical continues to diffuse through the skin sample in the time it takes to collect the tape strips. In addition, tape stripping is not ideal for volatile chemicals as they tend to evaporate faster than the required analysis time. As an alternative to processing the tape strips with solvent extraction and HPLC analysis, attenuated total reflectance-Fourier transform infrared spectrometry (ATR-FTIR) has been used to quantify chemicals on the tape with some degree of success.[89-91] The tape stripping technique is not exactly non-invasive, as substantial skin irritation is also generated by complete removal of the stratum corneum. Microdialysis has also been explored for topical bioequivalence measurement, but this is also a relatively invasive and inflammation-inducing procedure.[92] Non-invasive imaging techniques of the skin have become more popular over the last few years.[93]

Currently in vitro human skin drug-diffusion studies are done to study topical drug products. In these experiments, tissue samples are often extracted whole or sectioned and extracted, and then

the solvent extract is analyzed by high pressure liquid chromatography (HPLC) or some other analytical method for drug content. The skin-sample analysis time doubles the length of the experimental time, and doubles the cost of the experiment as well.

There is a significant need to find a non-invasive and simple method to determine topical drug product bioequivalence. There is also a substantial need to be able to do topical drug-product development research with an instrument that would cut development time and costs in half, if there was an easier way to analyze and image skin drug concentrations.

The best solution, however, in terms of time and ease of application is an all-optical approach. A rapid all-optical approach effectively eliminates the possibility of continued diffusion while scanning. In addition, different volatilities of chemicals are less relevant because scans are collected instantaneously from all depths in the skin sample. For these reasons, NIR spectrometry was chosen as the method of analysis.

Photonic techniques can detect and measure chemically and medically significant optical parameters not accessible to other clinical imaging modalities.[94-95] However, in contrast to X-rays, for example, the use of optical methods to probe human tissue properties presents major problems because human tissue is decidedly scattering, which sets up uncertainty in the interpretation of transmission and scattering data.

A number of types of experiments have been utilized in medical applications. One is the continuous-wave (CW) experiment employed in this research. In a CW experiment, a continuous

modulated beam (often a laser) injects photons into a tissue like skin and the intensity of reradiated light on the same surface is measured as a function of distance from the injection point. A second kind of experiment entails time gating, in which, after a pulse of photons is injected in the tissue, the reradiated photons are measured at a fixed distance from the injection point as a function of time. Models for these experiments normally presume that the tissue is semi-infinite and bounded by a planar surface. This is typically a fine assumption because photons at the energies used are able to penetrate only a few millimeters of the tissue. A third kind of experiment, not particularly useful in skin studies in vivo, is termed the transillumination experiment. In transillumination, photons traverse a thin slab of tissue. Measurements of the transmitted light are made either with a steady input beam or as a function of time.

Because tissue is greatly scattering, theory must be applied to decipher the measurements. The most rigorous formulation of the theory entails solving a transport equation (TE). Solving the transport equation involves intense computation and because scattering cross-sections can only be approximated, these calculations are not used as frequently as calculations based on photon diffusion theory or random walk theory.[96-97] Unluckily, diffusion theory is intrinsically inaccurate in close proximity to interfaces, which exist in skin. A further impediment to use of optical diffusion theory is that it cannot imitate very short-time ballistic behavior, which can include important information in time-gated experiments, and which would also be required to achieve depth resolution in skin.

Instrumentation. Econazole nitrate (EN) and 4-cyanophenol (4-CP) were chosen as the analytes for the following research. 4-CP is a model chemical with an NIR chromophore unique from skin.[89] Both compounds exhibit strong NIR chromophores, thus it is a simple matter to analyze

their spectra. Because NIR can have a tissue penetration depth of many millimeters, it has previously been used in studies for the noninvasive analysis of drugs in tissue.[98] It has also been used to quantify the depth of photon penetration in tissues.[99-100]

A future instrument already under construction will replace the bulky full spectrum NIR with a small and portable hand-held device termed a solid-state spectral imager (SSSI). It is composed of an aluminum block with three concentric rings of light-emitting diodes (LED). Each ring has a different diameter and is centered on a single lead sulfide photodetector. With different angles of incidence, each ring has a different angle of skin penetration. To be collected at the photodetector, light from the ring farthest from the center must penetrate the skin more deeply than from the ring closest to the center. Though only one signal is collected, the LEDs are frequency modulated such that the contribution of each individual LED can be deconvolved from the total signal. In this manner, a spectral profile can be constructed at different depths in the tissue, allowing for the rapid determination of drug penetration.

Theory. A three-dimensional version of the TE for the electric field E can be calculated using Maxwell's equations. Assuming a constant charge density, Maxwell's equations are:

$$\mu \frac{\partial H}{\partial t} = -\nabla \times E, \quad \sigma E + \varepsilon \frac{\partial E}{\partial t} = \nabla \times H \qquad \textbf{4.1}$$

where σ is the conductivity, ε is the dielectric permittivity, and μ is a parameter expressed in terms of impedance.[97] If **H** is removed from these equations then **E** fulfills the three-dimensional TE

$$\frac{\partial^2 E}{\partial t^2} + \frac{\sigma}{\varepsilon}\frac{\partial E}{\partial t} = \frac{1}{\mu\varepsilon}\nabla^2 E \qquad \textbf{4.2}$$

NIR models for chemical composition based on scattering data are generally derived statistically based upon calibration data, and as such are incomplete without a description of the multivariate statistics needed to analyze the results. In search for the most descriptive model, the present research explored the use of a number of different chemometric approaches including principal component regression (PCR),[69] interval PCR (iPCR), and PCR-uninformative variable elimination (PCR-UVE).[101-102] For each of these models, principal components were calculated by a singular value decomposition of matrix **A** according to Equation 4.3:

$$\mathbf{A} = \mathbf{USV} \qquad \textbf{4.3}$$

where **A** is the matrix of original spectra, **U** is the matrix of eigenvalues (scores), **S** is a diagonal matrix of singular values, and **V** is the matrix of eigenvectors (loadings).

Shown in Equation 4.4, a regression of **U** indicates which of the components have the strongest correlation to a change in drug concentration **y**, where **a** is the y-intercept, **b** is a vector of regression coefficients, and **c** is the residual.

$$\mathbf{y} = \mathbf{a} + \mathbf{bU} + \mathbf{c} \qquad \textbf{4.4}$$

Equation 4.5 demonstrates how a leave-one-out cross validation can be used to predict the concentration of topical drug present from the NIR spectra, where σ^2 is the variance, $f_i(U_i)$ is the prediction of the model for the *i-th* pattern *m* in the training set, after it has been trained on the *m-1* other patterns.

$$\sigma^2_{LOO} = \frac{1}{m}\sum_{i=1}^{i=m}(\mathbf{y}_i - f_i(\mathbf{U}_i))^2 \qquad \textbf{4.5}$$

In the case of a simple two-component system it is a simple matter to observe a linear change in the analyte concentration. When only two system components are present, only one principal component is needed to describe the concentration change accurately. The loading

corresponding to this principal component accurately reflects the contribution of each wavelength to the overall classification. However, in complex systems, such as tissues, where many system components change simultaneously, multiple principal components are needed for the prediction model.

Interval PCR performs the same analysis as above except on smaller subsets of the data set rather than the data set in its entirety.[102] For example, when the experimenter specifies an interval (I) of 200 wavelengths, the algorithm performs PCR followed by principal component selection and cross-validation on intervals of 200. By using a moving boxcar, all wavelengths are paired with all other wavelengths inside of ±I. For example, after the algorithm analyzes 1501-1700nm, the next iteration analyzes 1502-1701nm, and so on. At the final wavelength, the first I wavelengths are added to the end for the final iterations. In this manner, each wavelength is included in a new model I*2 times. The goal of interval selection is the minimization of the standard error of performance in Equation 4.6, which indicates the interval with the highest correlation to the change in drug concentration:

$$SEP = \frac{\left(\frac{c^2}{n-1}\right)^{1/2}}{\max(\mathbf{R}) - \min(\mathbf{R})} \qquad \textbf{4.6}$$

where c is the residual, n is the number of spectra, and R is a concentration vector.

From the description of iPCR, it is apparent that while one interval may give an acceptable predictive ability, it does not take into account the interaction among variables and peaks outside of the specified interval. Uninformative variable elimination (PCR-UVE) compensates for this by randomly selecting combinations of I variables.[102] Loading vectors are retained and summed

for each combination of variables so that every variable contains a weight from a number of different local models. The final loading vector indicates the individual wavelengths with the highest correlation to the desired change in drug concentration. The variables with the largest loading are used in the final calibration with the ultimate goal of minimizing the SEP.

Materials and Methods

Materials. 4-Cyanophenol (4-CP), Hanks' balanced salts modified powder, sodium bicarbonate and propylene glycol were purchased from Sigma Chemical (St. Louis, MO). Econazole nitrate (EN), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), gentamicin sulfate, trifluroacetic acid (TFA), triethylamine (TEA), potassium phosphate (monobasic, anhydrous), hydrochloric acid, polyethylene glycol 400, methanol and acetonitrile (ACN) were obtained from Fisher Scientific (Fair Lawn, NJ). 1-octane sulfonic acid sodium salt was purchased from Chrom Tech® (Apple Valley, MN). Chemical structures of 4-CP and EN can be found in Appendix B.

Donor solutions and creams. Saturated donor solutions of EN were prepared in propylene glycol (20 mg/mL). Saturated donor solutions of 4-CP were prepared in nanopure water (35 mg/mL). One percent econazole nitrate creams were used from four different manufacturers, Clay Park, Johnson and Johnson (Spectazole®), Fougera, and Taro. Propylene glycol, nanopure water, and a placebo cream were also investigated in order to collect skin spectra from vehicle interference.

In vitro diffusion studies. Teflon MatTek Permeation Devices (MPD, MatTek Corporation, Ashland, MA, USA) were used for the in vitro skin diffusion studies. The MPD is essentially a modified form of a Franz diffusion cell that is meant to be used for in vitro tissue culture

permeation studies. It is a practical diffusion cell choice because it requires a smaller amount of valuable skin and it is a portable and unbreakable Teflon design, as opposed to the larger glass Franz cells, which must be used on a large water-circulating/stirring bench. Spacers are used in the wells of the diffusion cells to hold the skin in place. Three MPDs were used for each treatment (donor solution/cream) and exposure time (three per formulation treatment), for a total of nine MPDs per formulation treatment. The purpose of the different exposure times was to create different skin concentrations for measurement and correlation. The receiver solution (simulated "blood flow" compartment under the skin in the MPD) was composed of 60% Hanks' pH 7.4 buffer and 40% polyethylene glycol 400 (PEG). The PEG is added to the buffer in order to help solubilize hydrophobic drugs without damaging the skin.

The diffusion studies were conducted with dermatomed (split-thickness, ~250μm) hairless guinea pig skin. Skin samples were stored at -20^{o}C until use. Animal use was approved by the University of Kentucky Institutional Animal Care and Use Committee. Hairless guinea pig skin was used for these experiments instead of human skin, in order to develop and validate the methodology in a tissue that is more readily available and economical than human tissue. Hairless guinea pigs are the best small animal model skin match to human skin for diffusion studies.[103] Skin samples were secured into the MPD and placed in a tightly sealed glass chamber with 10 mL of the receiver solution so that the level of receiver solution remained constant and in contact with the dermis side of the skin. Two-hundred μl of donor solution or 400 μl of cream were added directly onto the stratum corneum from the upper opening of the cell. To prevent the evaporation of the donor solution, cells were capped with vial caps (Waters, Milford, MA, USA). Micro-stirring bars were centered below each diffusion cell and set to stir at a constant rate

throughout the experiment. At the end of the diffusion experiment, skin samples were removed from the diffusion cells and rinsed with nanopure water 3 times for 10 seconds each. In the case of the creams, an alcohol wipe was used to remove excess surface formulation gently. Samples were placed on a paper towel and blotted, then two tape strips (Scotch Book Tape 845) were done to remove surface drug. The skin was rinsed once more with nanopure water, blotted dry with paper towel, and the treated skin area was excised from the center of the skin sample. NIR analysis followed immediately.

Skin extraction and HPLC analysis. Immediately following NIR analysis, the tissue was weighed. The sample was then minced with a scalpel and placed in a vial with 5 mL of acetonitrile. This vial was then sonicated for 10 minutes and shaken for 15 hours at 32°C to extract the drug from the tissue into the acetonitrile. The tissue extract was then analyzed for drug concentration by HPLC analysis, and expressed as μg of drug per wet g of tissue weight.

The high-pressure liquid chromatography (HPLC) assay was done with a Perkin-Elmer Series 200 Autosampler, Pump, Column Oven, and a 785A UV/VIS Detector with Turbochrom Professional Version 4.1 Software. A Brownlee® C18 RP Spheri-5 μm column (220 x 4.6 mm) with a C18 RP 7 μm guard column (15 x 3.2 mm) was used with the UV/VIS Detector set at a wavelength of 215 nm for econazole nitrate and 246 nm for 4-cyanophenol. The mobile phase used for econazole nitrate was 70:30 ACN:25mM potassium phosphate buffer with 0.65 g/L 1-octane sulfonic acid sodium salt. The mobile phase used for 4-cyanophenol was 50:50 ACN:0.1% TFA adjusted with TEA to a pH of 3.0. The flow rate of the mobile phase was at 1.5 mL/min for econazole nitrate and 1.0 mL/min for 4-cyanophenol with 100 μL and 30 μL

injections of the sample, respectively. Standards were analyzed with each set of diffusion samples and exhibited excellent linearity over the concentration range employed. The retention times for econazole nitrate and 4-cyanophenol were 5.94 ± 0.05 min and 3.33 ± 0.04 min, respectively. The sensitivity of each assay was 25 ng/mL.

NIR Analysis

Drug powders. Pure powders were scanned first to ensure that econazole nitrate (EN) and 4-cyanophenol (4-CP) had distinct near infrared (NIR) chromophores. A uniform layer of pure drug approximately 1mm thick was loaded on a one-well depression microscope slide (Gold Seal Products, Portsmouth, NH). NIR spectra from 1100-2500nm were collected in steps of two nm (1100, 1102, 1104…) with a NIR spectrometer (Technicon InfraAlyzer 500, Tarrytown, NY) interfaced to a computer (OptiPlex GXM 5166, Dell, Round Rock, TX, USA) running SESAME 3.1 (Bran+Luebbe, Norderstedt, Germany). Scans were collected inside the instrument drawer. To maximize light scatter, microscope slides and samples were placed on top of a conical reflecting cup, designed such that when a sample is placed along the axis of radial symmetry of the cone, specular reflection at the detector is minimized while diffuse reflectance is maximized.[104] All data were exported to Matlab 7.0.1 (The Mathworks Company, Natick, MA) for processing and analysis. Samples were each scanned three times, rotating them 120 degrees between each scan.

Drugs in Solution. To demonstrate how effectively NIR could predict drug concentrations from NIR spectra of drugs in solution, spectra were collected from EN and 4-CP in 75% acetonitrile (ACN) and 25% buffer:PEG receiver solution. The drug that diffuses through the skin can be

analyzed in the receiver solution diluted with 75% ACN for HPLC analysis, so the same solution composition was compared by NIR. Samples were placed inside the instrument drawer and scanned on top of a 135-degree conical reflecting cup to maximize light collection by the integrating sphere and detector. Each sample was scanned three times, and rotated 120 degrees between each subsequent scan. For data analysis, data were divided into their respective groups.

Saturated Solutions and Creams on Skin. Exposure times for EN solutions were 2 min, 2 hrs, and 15 hrs, EN creams were 2 hrs, 8.5 hrs, and 15 hrs, and exposure times for 4-CP were 2 min, 10 min, and 1 hr. Following exposure to drug treatment, skin samples were removed from the MPD, the drug-exposed skin section (approximately 1.0 cm in diameter) was excised and the surface formulation was carefully removed, loaded on a one-well depression slide, and covered with a cover slip (Gold Seal Products, Portsmouth, NH). NIR spectra were collected from covered skin samples. Samples were placed inside of the instrument drawer and on top of the conical reflecting cup and scanned three times each, rotating them 120 degrees between each scan. For the data analysis, data were divided into their four respective groups; EN epidermis (outside/donor surface facing NIR probe), EN dermis (inside/receiver surface facing NIR probe), CP epidermis, and CP dermis.

Data Analysis Method. To eliminate baseline variation due to different tissue thicknesses and sample preparation techniques, all data were multiplicative scatter-corrected.[85] Data were presmoothed with a cubic spline operation. The following chemometric processing was applied to each of the data sets individually. From the scatter-corrected NIR spectra, principal components were calculated.[9] A multiple least-squares regression (MLR) identified the PCs that

correlated most highly with the reported drug concentrations from the HPLC analysis. These components were used in a leave-one-out cross-validation to determine how effectively the NIR method predicted drug concentrations from the spectra. To find the most effective data analysis method, PCR was compared with iPCR and PCR-UVE.

Testing for interferences. To prove the strong correlations between HPLC and NIR were not artifacts, and were in fact coming from 4-CP and EN, the contributions of solvents, drug vehicles, and placebo cream were evaluated and appropriate corrections were made. For example, normalized ACN and 25% buffer:PEG control spectra were subtracted from the 4-CP and EN solution spectra and statistical analysis was repeated. The solvent corrected spectra still demonstrated a high correlation to known concentrations, suggesting the effects were entirely due to the analyte of interest. In addition, the absence of the solvent peaks allowed for the identification of the most significant regions in the NIR spectrum for the quantification of each drug. Rather than using a full-spectrum approach where the likelihood of chance correlations is higher, these specific wavelength regions were used for the quantification of each drug for saturated solution and cream experiments. In its identified wavelength region, the EN cream demonstrated a high correlation to HPLC concentrations, while at the same region, there was no predictive ability for the placebo cream, proving this was the correct region.

Results

Interval PCR with 200 wavelengths resulted in the best standard errors of performance. Leardi et. al. suggest the inclusion of more variables increases the likelihood of overfitting and chance correlations.[101] For 4-CP, the region spanning 1470-1870nm demonstrated the highest

correlation to HPLC concentrations, and for EN the best region spanned 1936-2336nm. Therefore, these regions were selected for all quantification experiments. Figures 4.1-4.4 illustrate calibration curves for 4-CP and EN in solution and in tissue. Near-infrared spectrometry successfully predicted the concentration of both types of analytes in solution, as shown in Figure 4.1. The left side of Figure 4.1 illustrates that NIR analysis of 4-CP in solution gives a relative standard error of estimation (SEE) over the concentration range of 4.36%, relative standard error of performance (SEP) over the concentration range of 5.74%, and r^2=0.978. The right side of Figure 4.1 illustrates that NIR analysis of EN in ACN/PEG gives a SEE=4.73%, SEP=6.38%, and r^2=0.974.

Figure 4.2 (left) illustrates that the epidermal NIR calibration line from skin samples treated with an applied dose of saturated EN in propylene glycol gave a SEE=4.95%, SEP=6.28%, and r^2=0.965. Figure 4.2 (right) illustrates that the dermal NIR calibration line gave a SEE=4.77%, SEP=6.73%, and r^2=0.959. These calibration lines were based on the leave-one-out cross-validation using the HPLC concentrations for the regression.

Skin samples treated with an applied dose of saturated 4-CP in water and scanned from the epidermal side gave a SEE=7.31%, SEP=9.17%, and r^2=.0952, shown in Figure 4.3 (left). The same samples scanned from the dermal side gave an SEE=5.51%, SEP=6.84%, and r^2=0.973, shown in Figure 4.3 (right).

The most important results obtained are in Figure 4.1 and Tables 4-1 and 4-2, showing the excellent correlation of the 1% econazole nitrate cream-treated skin data (analyzed by NIR) with

the reference HPLC assay. Skin samples treated with an applied dose of 0.4 mL Clay Park (1% EN) scanned from the epidermis gave a SEE=4.16%, SEP=5.39%, r^2=0.975, shown in Figure 4.4 (left). The same samples scanned from the dermis gave a SEE=4.17%, SEP=5.05%, r^2=0.978, shown in Figure 4.4 (right). An expanded list of Clay Park results are given in Table 4.1 and concise results for the other three lotions are given in Table 4.2. It is this proof of concept data that leads us to believe that it will be possible to analyze many other topical drug products using an NIR SSSI. Additionally, we are not limited by NIR, because UV, visible, and NIR LEDs plug into the SSSI and are easily changed.

Discussion

Great Flexibility and Durability. This experiment applied NIR spectroscopic analysis in four distinct applications; drug powders, drug in solution, and tissue concentrations of drug after exposure to solutions and creams. The relative standard errors of performance and estimation, and the strong correlation of NIR concentration prediction to the HPLC results suggest that NIR spectrometry is a flexible technique for analysis of tissue concentrations of drugs and other chemicals after topical exposure. In total, 270 NIR spectra were collected from skin samples. Only two spectra were discarded because of anomalous scans (short noise spikes from instrument preamplifier), suggesting that NIR is a very durable instrument.

NIR is also a nondestructive and rapid method of analysis, taking less than two minutes to complete a scan. No special sample preparation is required. In this research, samples were simply placed on a microscope slide and scanned. NIR also has the capability of being a noninvasive method of analysis. Using a fiber-optic probe it is possible to scan skin tissue *in*

vivo in whole animal or clinical studies. In this case, skin samples were scanned inside the instrument drawer, thus external light and room noise were not a factor. In the drawer configuration, NIR spectra are essentially free from external noise and interference. Freedom from interference can be measured. An NIR "contaminant" spectrum (e.g. EN in the presence of trace amounts of 4-CP) can be mathematically translated in hyperspace toward one of the drugs until the two are chemically indistinguishable by the BEST metric.[32] Spectra from similar samples tend to cluster in the same region of hyperspace, thus pure component spectra cluster in distinct regions. The space between the two cluster centers is composed of theoretical mixtures of the two components. The distance that one cluster is translated toward the other before they are spatially indistinguishable corresponds to the theoretical NIR detection limit. When applied to EN and 4-CP, this operation indicated that both clusters were mathematically distinguishable until the mixture consisted of 98.0% contaminant.

The success of this phase of the experiments suggests the utility of a new portable NIR sensor for the rapid and accurate in vivo measurement of dermal absorption and topical bioequivalence.[106] The concentric circle design of a new NIR SSSI allows the instrument to focus on all tissue depths simultaneously. When a scan is collected from the epidermal side, the signal detected from the innermost circle reflects tissue scattering events only from the surface. However, the signal detected from the middle circle reflects tissue scattering events from both the surface and a moderate number of skin layers. Likewise, the signal detected from the outermost circle reflects tissue scattering events from all depths simultaneously. With this information in hand, the inner and middle circles can be used as reference channels, and their drug spectra can be subtracted to yield pure interference-free spectra at lower depths. Removal of formulation prior

to scanning tends to be highly variable; however, the SSSI will make it possible to measure drug penetration at lower tissue depths without removing formulation from the surface.

Conclusion

As the use of topical dosage forms becomes more common, the need has arisen for an analytical method capable of noninvasively assessing dermal absorption and topical bioequivalence. This research demonstrates that for the rapid and accurate measurement of topical bioequivalence of econazole nitrate formulations, NIR spectrometry is an analytical method worth further exploration. A portable NIR instrument is currently under construction that provides for depth profiling at a number of discrete NIR wavelengths based on a CW experiment format using spatial feature encoding and integrated computational imaging.[107] With an array of light-emitting diodes, and the use of frequency modulation, all wavelengths are collected simultaneously. The instrument is designed so a NIR spectrum is collected at various depths in the skin sample, so that the experimenter can locate a spectral signature at different depths in the sample. With this design, the entire analysis takes merely 15 seconds, eliminating the problem of thermal drift and chemical evaporation.

Chapter Four Tables

Table 4.1: This table presents the skin concentrations from a 0.4 mL applied dose of 1% Clay Park cream at three different exposure times as measured by the HPLC and NIR, and the strength of correlation between the two analytical instruments. Each tissue was scanned three times in the NIR prior to HPLC analysis.

Exposure Time	Tissue Number	HPLC (μg/g)	NIR (μg/g)	NIR Std Error	HPLC NIR Correlation
2 hr	Tissue #1	216	249		0.867
		216	191	17.3	0.886
		216	207		0.956
2 hr	Tissue #2	196	194		0.992
		196	192	1.64	0.979
		196	189		0.963
2 hr	Tissue #3	394	413		0.955
		394	416	7.18	0.946
		394	383		0.972
8.5 hr	Tissue #4	116	149		0.780
		116	136	10.5	0.854
		116	113		0.973
8.5 hr	Tissue #5	235	223		0.948
		235	228	10.5	0.970
		235	202		0.862
8.5 hr	Tissue #6	489	494		0.991
		489	501	18.1	0.975
		489	444		0.907
15 hr	Tissue #7	408	394		0.966
		408	414	5.81	0.985
		408	404		0.989
15 hr	Tissue #8	212	227		0.936
		212	207	14.2	0.978
		212	178		0.838
15 hr	Tissue #9	108	143		0.755
		108	101	15.6	0.931
		108	93		0.863

Table 4.2: This table illustrates the strength of correlation between the HPLC and NIR for the four different brands of econazole nitrate lotions tested. The statistics (r^2, standard error of estimation (SEE), and standard error of performance (SEP)) come from a leave-one-out cross validation.

Lotion Brand	**Spectazole®**		**Clay Park**	
	Epiderm.	**Dermis**	**Epiderm.**	**Dermis**
r^2	0.957	0.931	0.975	0.978
SEE	5.46	6.53	4.16	4.17
SEP	6.43	8.17	5.39	5.05

Lotion Brand	**Taro**		**Fougera**	
	Epiderm.	**Dermis**	**Epiderm.**	**Dermis**
r^2	0.936	0.961	0.934	0.960
SEE	7.42	5.38	7.36	5.78
SEP	8.94	6.95	8.97	6.96

Chapter Four Figures

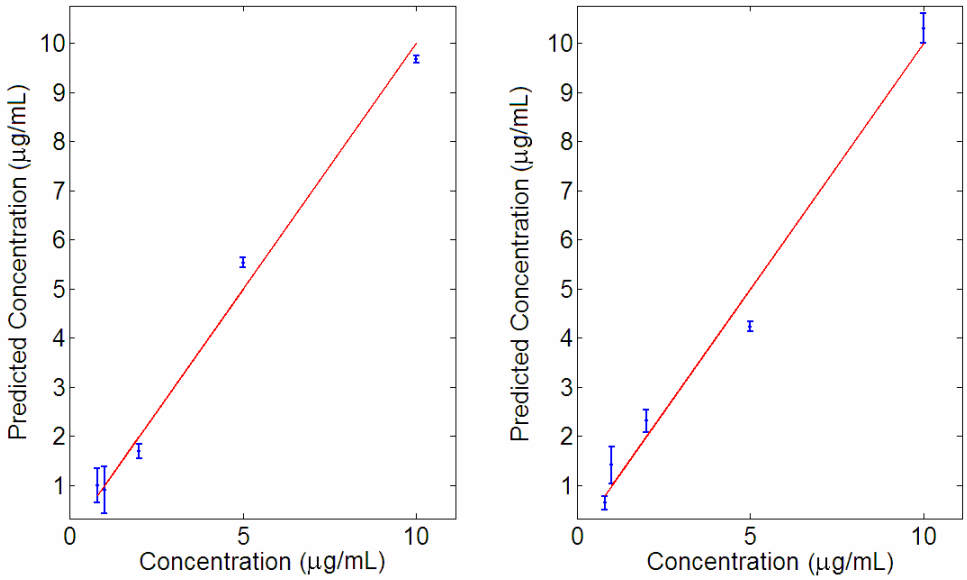

Figure 4.1: (Left) Calibration line for 4-cyanophenol (4-CP) in solution, relative standard error of estimation (SEE) over the concentration range is 4.36%, relative standard error of performance (SEP) over the concentration range is 5.74%, r^2=0.978. (Right) Calibration line for EN in solution, SEE=4.73%, SEP=6.38%, r^2=0.974.

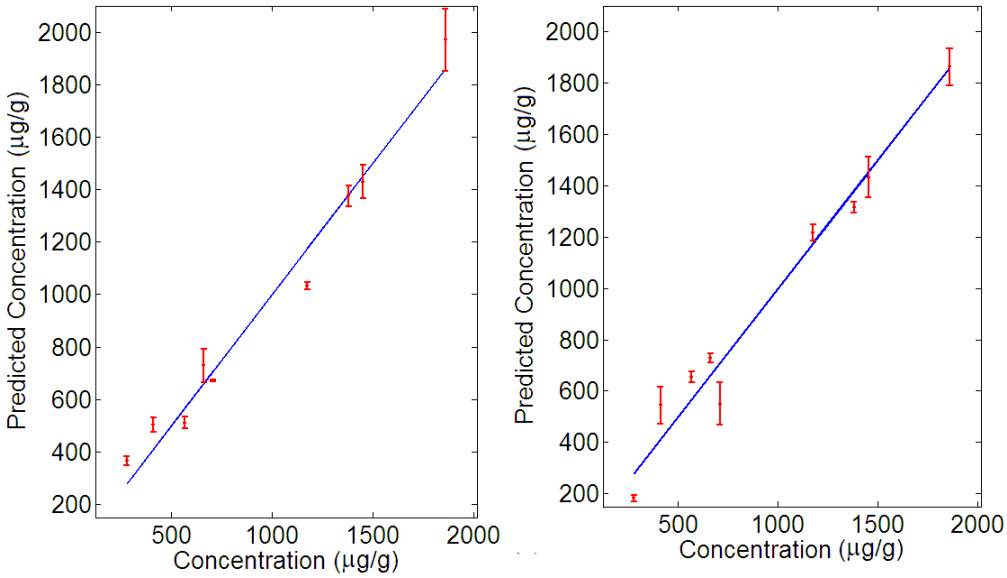

Figure 4.2: NIR calibration line from skin samples treated with an applied dose of saturated EN in propylene glycol, (left) epidermis calibration, SEE=4.95%, SEP=6.28%, r^2=0.965, (right) dermis calibration, SEE=4.77%, SEP=6.73%, r^2=0.959. The calibration line is based on the leave-one-out cross-validation using the HPLC concentrations for the regression. The diagonal line is the mean of the HPLC results, and the error bars are the means and the standard errors of the NIR predictions.

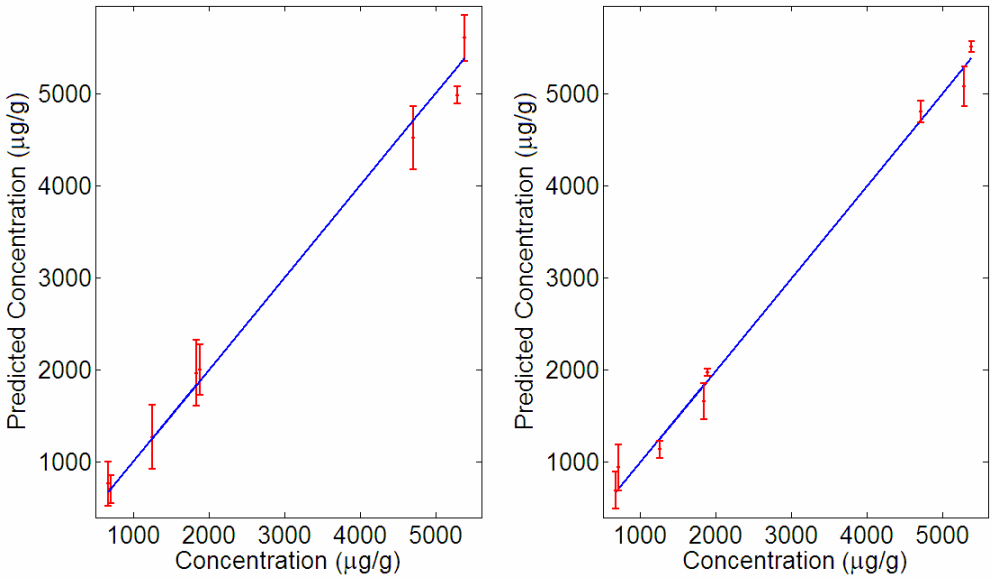

Figure 4.3: NIR calibration line from skin samples treated with an applied dose of saturated 4-CP in water, (left) epidermis calibration, SEE=7.31%, SEP=9.17%, r^2=0.917, (right) dermis calibration, SEE=5.51%, SEP=6.84%, r^2=0.973. The diagonal line is the mean of the HPLC results, and the error bars are the means and the standard errors of the NIR predictions.

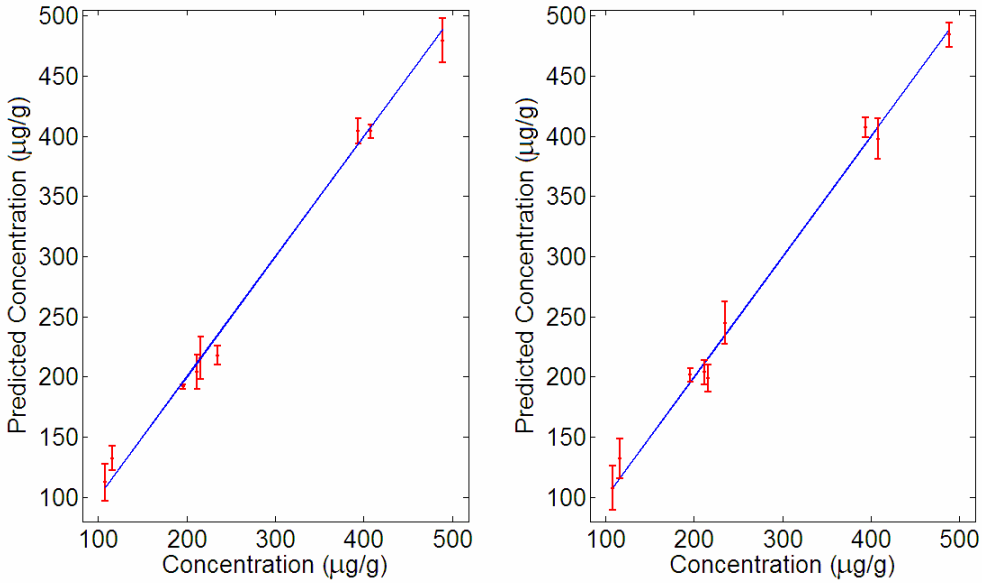

Figure 4.4: NIR calibration line from skin samples treated with an applied dose of 0.4 mL Clay Park (1% EN), (left) epidermis calibration, SEE=4.16%, SEP=5.39%, r^2=0.975, (right) dermis calibration, SEE=4.17%, SEP=5.05%, r^2=0.978. The diagonal line is the mean of the HPLC results, and the error bars are the means and the standard errors of the NIR predictions.

Copyright Statement

Joseph Medendorp, Jhansi Yedluri, Dana C. Hammell, Tao Ji, Robert A. Lodder, Audra L. Stinchcomb. *Pharmaceutical Research*. **2006**; 23(4): 835-853.

Chapter Five - Near infrared spectrometry for the quantification of human dermal absorption of econazole nitrate and estradiol

Introduction

The drugs used in this study cover two different broad classes of biopharmaceutical goals; a topical drug product category requiring development of a rapid and noninvasive bioavailability/bioequivalence test, and a popular transdermal drug product category requiring development of a rapid and noninvasive test for examining skin drug concentrations after changes in formulation strategy. Whereas it is obvious that topical drug products for local skin treatment need to have a method for biosampling the tissue after product application, it is not as obvious that measuring skin drug concentrations after transdermal treatments are relevant for optimization of drug partitioning and investigation of skin depot effects. Transdermal studies also require standard systemic plasma sampling for complete bioavailability/bioequivalence determination, but if drug concentrations do not exist in the skin of human subjects after application, then there is no need to continue with the time and expense of systemic sampling. The two drugs investigated in this study were econazole nitrate (EN) and estradiol (EST). EN is an antifungal agent that is indicated for the treatment of infections caused by susceptible dermatophyte and candida species, including tinea pedis, tinea cruris, tinea corporis, tinea versicolor and cutaneous candidiasis. A 1% EN topical cream prescription drug product is commercially available from multiple generic manufacturers and Johnson and Johnson (Spectazole® brand name). EST is typically prescribed for hormone replacement therapy in post-menopausal women. In addition to the multiple oral, vaginal, and injectable EST formulations on the market, many brands of transdermal delivery dosage forms are also available. Gels, emulsions, and patches comprise the popular options for non-oral hormone replacement, including Estrogel®, Estrasorb®, Climara-Pro®, Combipatch®, Esclim®, Estraderm®, Menostar™, Vivelle®, Vivelle-Dot® and others

Current methods for measuring dermal absorption of drugs such as EN and EST include tape stripping,[88] punch biopsies, and microdialysis.[92] Each of these methods is invasive, causing skin irritation or permanent scarring for the patient or healthy human volunteer. Additionally, quantification of the drugs sampled during these procedures requires a very time-consuming tissue/fluid extraction step and chromatography-based assay. To illustrate why this is so problematic, consider briefly the phases of the drug development cycle of a topical pharmaceutical product. In preclinical testing, the initial research is conducted and a formulation is developed, followed by phase I, II, and III clinical trials. At each step of the development process, topical bioavailability/bioequivalence must be measured and guaranteed to fall within required therapeutic limits. Preclinical testing alone includes both *in vivo* animal studies and *in vitro* tissue extraction studies, requiring perhaps thousands of tissue samples. Phase I, II, and III clinical trials require 10-80, 100-300, and 1000-3000 volunteer test subjects, respectively.[108] Before one even takes into account potential studies needed for manufacturing process changes, formulation changes, or generic product development, this process requires a very high number of tissue samples for the development of a single drug product. With inefficient and invasive methods of analysis as the industry standard, an all-optical alternative for measuring dermal absorption is comparatively attractive. An all-optical bioanalytical method is ideally suited as a high-throughput method of analysis for such a demanding drug development process.

As an alternative to processing the tape strips from dermatopharmacokinetic studies with solvent extraction and HPLC analysis, attenuated total reflectance-Fourier transform infrared spectrometry (ATR-FTIR) has been used to quantify chemicals on the tape strips.[89-91] While ATR-FTIR is an improvement over tissue extraction and HPLC, it still requires tape stripping.

Because tape stripping is an inefficient method of biosample collection, drug has been demonstrated to continue diffusing faster than the tape strips can be collected.[91] This inaccuracy and inconvenience of sample collection demonstrates that there is clearly a place for an all-optical approach for measuring dermal absorption. In our previous study,[82] we have demonstrated the utility of NIR spectrometry for the *in vitro* quantification of EN and 4-cyanophenol in hairless guinea pig skin. The purpose of the current study was to investigate the use of NIR spectrometry for the *in vitro* quantification of EN and EST in human skin.

Theory and Instrumentation. NIR spectrometry can have a tissue penetration depth of many micrometers, millimeters or centimeters depending upon wavelength; making it an analytical method of choice for this research.[98,109] Beer's Law states that NIR spectra are the result of a linear combination of the pure component spectra that comprise the sample; therefore, linear multivariate statistics can be used to quantify the analyte of interest even as all other system components change.[110] A vector of instrumental responses $\mathbf{r}_k$ can be represented as the sum of two independent signals, the signal from all interferences $\mathbf{r}_k^{=}$, and the signal from the analyte of interest $\mathbf{r}_k^{\perp}$, which is orthogonal to the contribution from the interferences.[22] This orthogonal portion is termed the net analyte signal (NAS), and is the portion of the signal used for multivariate calibration. Figure 5.1 illustrates the concept of the NAS. For NAS calibration, only the presence of interferences and their specific spectra must be known *a priori*, not their specific concentrations. A matrix **R** *(J x I)* without the analyte of interest must be available, where *J* is the number of wavelengths and *I* is the number of samples. For the solution exposure experiments, the matrix **R** is composed of drug vehicle and a blank tissue spectrum, and for the

cream experiments the matrix **R** is composed of placebo cream and a blank tissue spectrum. A projection matrix $\mathbf{P}_k^{\perp}$ can be calculated according to Equation 5.1:

$$\mathbf{P}_k^{\perp} = (\mathbf{I} - \mathbf{R}_{-k}\mathbf{R}_{-k}^{+}) \qquad \textbf{5.1}$$

where **I** is the identity matrix, and the '+' superscript indicates the Moore-Penrose pseudoinverse. Using one-point calibration with spectrum $\mathbf{r}_{cal}$, the NAS vector $\mathbf{r}_{cal}^{\perp}$ can be calculated with Equation 5.2.

$$\mathbf{r}_{cal}^{\perp} = \mathbf{P}_k^{\perp}\mathbf{r}_{cal} \qquad \textbf{5.2}$$

This vector is then normalized to length one with Equation 5.3:

$$\mathbf{r}_k^{NAS} = \frac{\mathbf{r}_{cal}^{\perp}}{\left\|\mathbf{r}_{cal}^{\perp}\right\|} \qquad \textbf{5.3}$$

The slope of the calibration line can be calculated from Equation 5.4:

$$s = \frac{\left\|\mathbf{r}_{cal}^{\perp}\right\|}{c_{cal}} \qquad \textbf{5.4}$$

where s is the slope and c_{cal} is the analyte concentration of the calibration spectrum. Illustrated in Figure 5.2a, the first step of NAS calibration includes using Equations 5.2-5.4 to find the NAS direction and determine the length of the NAS vector. As illustrated in Figure 5.2b, when the NAS direction and magnitude are known, the unknown spectrum $\mathbf{r}_{un}$ can be projected in the NAS direction with Equation 5.5 and its magnitude compared to the calibration magnitude.

$$y_{un}^{NAS} = \mathbf{r}_{un}^{T}\mathbf{r}_k^{NAS} \qquad \textbf{5.5}$$

The ultimate goal of the multivariate calibration is the calculation of the unknown analyte concentration c_{un} which can now be derived with Equation 5.6.

$$c_{un} = \frac{1}{s} y_{un}^{NAS} \qquad \textbf{5.6}$$

The NAS approach allows for the calculation of figures of merit from multivariate data sets. In severely overlapping spectra, it has historically been difficult to quantify selectivity, sensitivity, and signal-to-noise (S/N) because of the inability to distinguish between interferences and the analyte of interest.[19-20] With the NAS, these quantities can be measured directly. Selectivity is defined as the scalar degree of overlap, α, between the NAS vector and the calibration spectrum according to Equation 5.7:

$$\alpha = \frac{\left\| \mathbf{r}_{cal}^{\perp} \right\|}{\left\| \mathbf{r}_{cal} \right\|} \qquad \textbf{5.7}$$

The selectivity is a measure from 0 to 1 indicating how unique the analyte of interest is compared to the interferences. The sensitivity is a measure of the analyte variation in response to a change in concentration. This quantity can be expressed as Equation 5.8:

$$\mathbf{s}_k = \mathbf{r}_k^{NAS} / c_k \qquad \textbf{5.8}$$

where c_k is the concentration of the k-th analyte. Sensitivity should be the same for each concentration and each NAS vector.[21] The S/N ratio can be expressed as Equation 5.9:

$$S/N = \frac{c_k \left\| \mathbf{r}_k^{NAS} \right\|}{\left\| \boldsymbol{\varepsilon} \right\|} \qquad \textbf{5.10}$$

where $\boldsymbol{\varepsilon}$ is the random instrumental error.

Materials and Methods

Materials. Estradiol (EST), Hanks' balanced salt powder, sodium bicarbonate, ethanol and propylene glycol were purchased from Sigma Chemical (St. Louis, MO). Econazole nitrate (EN), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), gentamicin sulfate,

trifluroacetic acid (TFA), triethylamine (TEA), potassium phosphate (monobasic and dibasic), sodium hydroxide, polyethylene glycol 400, methanol and acetonitrile (ACN) were obtained from Fisher Scientific (Fair Lawn, NJ). The chemicals were used to make the buffers, donor and receiver solutions for the skin diffusion studies, and for the HPLC mobile phases as described in the respective section below. Chemical structures for estradiol and econazole nitrate can be found in Appendix B.

Human skin. The diffusion studies were conducted with human skin. Human skin from abdominoplasty surgery was obtained from the National Cancer Institute's Cooperative Human Tissue Network (CHTN). The samples were dermatomed immediately upon arrival to a thickness of approximately 200 μm and frozen at – 20^{o}C. On the day of the experiment, the skin was thawed and used for the studies. Skins from three different individuals were used for three different treatments (EST solution, EN solution and EN cream studies).

Donor solutions and creams. Saturated donor solutions of EN were prepared in propylene glycol (20 mg/mL). Donor solutions of EST (0.25%) were prepared in ethanol, a concentration relevant to current topical formulations for transdermal delivery. One percent EN cream and a corresponding placebo cream were used. Propylene glycol and ethanol were also investigated in order to collect skin spectra from drug vehicle interferences.

In vitro diffusion studies. Teflon MatTek Permeation Devices (MPD, MatTek Corporation, Ashland, MA, USA) were used for the *in vitro* skin diffusion studies with EN and EST solutions. The MPD is essentially a modified form of the Franz diffusion cell that is designed for *in vitro*

tissue culture permeation studies. The MPD was the diffusion cell of choice because it requires a smaller amount of valuable skin, and is a portable and unbreakable Teflon design in contrast to the larger glass Franz cells, which must be used on a large water-circulating/stirring bench. Spacers were used in the wells of the diffusion cells to hold the skin in place. Three MPDs were used for each treatment (donor solution) and exposure time (three per donor solution treatment). The receiver solution (simulated "blood flow" compartment under the skin in the MPD) was composed of 60% Hanks' pH 7.4 buffer and 40% polyethylene glycol 400 (PEG). The PEG was added to the buffer in order to help solubilize hydrophobic drugs without damaging the skin.

Skin samples were secured into the MPD and placed in a tightly sealed glass chamber with 10 mL of the receiver solution so that the level of receiver solution remained constant and in contact with the dermis side of the skin. Two-hundred μL of donor solution were added directly onto the stratum corneum from the upper opening of the cell. To prevent the evaporation of the donor solution, cells were capped with vial caps (Waters, Milford, MA, USA). Micro-stirring bars were centered below each diffusion cell and set to stir at a constant rate throughout the experiment.

EN cream studies were conducted in PermeGear modified-Franz flow-through diffusion cells with an area of 1.77 cm^2 and heating blocks maintained at 32°C (PermeGear, Riegelsville, PA), a retriever IV fraction collector (ISCO, Inc., Lincoln, NE), and a Pumppro MPL static pump (Watson Marlow, Wilmington, MA). The FDA has previously recommended 1.77 cm^2 Franz diffusion cells for the *in vitro* testing of topical products. Thus, the present study encompasses both the MPDs, a more practical and portable Teflon-design cell, as well as the standard 1.77 cm^2 diffusion cells**.** The diffusion experiment was initiated by charging the donor compartment

with 50 mg of the 1% EN or placebo cream. The cream was applied evenly on the skin with a Teflon rod taking care not to damage the skin. The receiver solution was pumped through the diffusion cells at a flow rate of 1.0 mL/h for 2h.

At the end of the diffusion experiment, skin samples were removed from the diffusion cells and rinsed with nanopure water three times for ten seconds each. In the case of the creams, an alcohol wipe was used to remove excess surface formulation gently. Samples were placed on a paper towel and blotted, and two tape strips (Scotch Book Tape 845) were applied to remove any surface drug. The skin was rinsed one more time with nanopure water, blotted dry with paper towel, and the treated skin area was excised from the center of the skin sample. NIR analysis followed immediately.

Near infrared analysis. Drug powders were scanned first to verify the presence of distinct NIR chromophores for EN and EST. A uniform layer of pure drug approximately 1mm thick was loaded on a one-well depression microscope slide (Gold Seal Products, Portsmouth, NH). NIR spectra from 1100-2500nm were collected in steps of two nm with a scanning monochromator system described earlier[84] interfaced to a computer (OptiPlex GXM 5166, Dell, Round Rock, TX, USA) running SESAME 3.1 (Bran+Luebbe, Norderstedt, Germany). To maximize light scatter, microscope slides and samples were placed on top of a 135-degree conical reflecting cup, designed such that when a sample is placed along the axis of radial symmetry of the cone, specular reflection at the detector is minimized while diffuse reflectance is maximized.[104] All data were exported to Matlab 7.0.1 (The Mathworks Company, Natick, MA) for processing and analysis. Powder samples and solutions were scanned six times each, rotating them 120 degrees

between each scan. Tissue samples were scanned three times from the epidermal side, and three times from the dermal side.

Chemometric analysis. All chemometric algorithms were written by the authors. NAS calibration requires a matrix of spectra without the analyte such that the region of space spanned by the interferences is included in the calibration model. For quantification of drug absorbed from solutions of EST and EN applied to the tissue, the matrix consisted of a pure component spectrum from drug vehicle (ethanol or propylene glycol) and a blank tissue spectrum exposed to drug vehicle alone. For quantification of EN cream, the matrix consisted of a placebo cream spectrum and a blank tissue spectrum exposed to placebo. With this approach, it was not necessary to identify the spectrum regions of the spectrum where the target drug could be quantified. NAS calibration automatically finds the regions of the spectrum orthogonal to the space spanned by the interferences, thus canceling out their effects on the net analyte signal. Figures of merit, such as sensitivity, selectivity, S/N ratio, limit of detection, bias, and precision were also calculated directly from the NAS.

Skin extraction and HPLC analysis. Immediately following NIR analysis, the tissue was weighed. The sample was then minced with a scalpel and placed in a vial with 1 or 2 mL of ACN for solution and cream studies, respectively. This vial was sonicated for 10 minutes and shaken for 15 hours at room temperature to extract the drug from the tissue into the ACN. The tissue extract was then analyzed for drug concentration by HPLC analysis, and expressed as μg of drug per wet g of tissue weight.

The high-pressure liquid chromatography (HPLC) assay was done with a Perkin-Elmer Series 200 Autosampler, Pump, Column Oven, and a 785A UV/VIS Detector with Turbochrom Professional Version 4.1 Software. A Brownlee® C18 RP Spheri-5 μm column (220 x 4.6 mm) with a C18 RP 7 μm guard column (15 x 3.2 mm) was used with the UV/VIS Detector set at a wavelength of 215 nm for EN and 205 nm for EST. The mobile phase used for EN was 70:30 ACN:0.1% TFA adjusted with TEA to a pH of 3.0. The mobile phase used for EST was 45:55 ACN:methanol and distilled water (5% and 50%, respectively). The flow rate of the mobile phase was at 1.0 mL/min for EN and EST.

Results

For HPLC, standards were analyzed with each set of diffusion samples and exhibited excellent linearity over the concentration range employed. The retention times for EN and EST were 10.02 ± 0.05 min and 10.37 ± 0.03 min, respectively. The sensitivities of the assays for EN and EST were 25 ng/mL and 50 ng/mL, respectively.

To test whether or not there were appreciable differences between the NIR ability to measure drug absorption from the dermal side and the epidermal side, scans were collected from each side and analyzed separately. Using NAS calibration, NIR spectrometry measured an applied dose of 200 μL of saturated EN in propylene glycol on human skin with an r^2=0.988, a standard error of estimate over the concentration range (SEE) of 2.46%, a standard error of performance over the concentration range (SEP) of 2.86% from the epidermal side, and an r^2=0.996, SEE=1.98%, SEP=2.12% from the dermal side. (See Figure 5.3 for the calibration results.) Figure 5.4 illustrates that NIR spectrometry measured different concentrations of EST in ethanol using NAS

calibration with an r^2=0.988, SEE=3.46%, SEP=4.01%. Figure 5.5 illustrates that NIR measured an applied dose of EST in ethanol on human tissue with an r^2=0.987, SEE=3.30%, SEP=5.66% from the epidermal side, and an r^2=0.967, SEE=5.53%, SEP=6.83% from the dermal side. NIR measured an applied dose of 50 mg of either placebo cream or 1% EN cream with an r^2=0.987, SEE=2.30%, SEP=2.66% from the epidermal side, and an r^2=0.987, SEE=2.53%, SEP=2.83% from the dermal side (data not shown). The drug absorption was extremely high for one tissue sample with these data, however, and these statistics were calculated with the inclusion of a high-end outlier point. This outlier was most likely due to an insufficient surface formulation wash-step in this particular sample; however, it was encouraging to note that both assay methods consistently detected this procedural artifact. Figure 5.6 illustrates the same scenario with the high-end outlier excluded from the model. In this case, NIR resulted in a calibration r^2=0.981, SEE=5.67%, SEP=6.07% from the dermal side and an r^2=0.980, SEE=5.93%, SEP, 6.71% from the epidermal side.

The sensitivity (s_k), selectivity (α), and S/N ratios were calculated according to Equations 5.7-5.9, respectively. For solutions of EST on human skin, s_k = 0.0455, α = 0.143, and the S/N = 29.54. For saturated solutions of EN on human skin, s_k = 0.0104, α = 0.0205, and the S/N = 10.04, and for EN cream, s_k = 0.052 , α = 0.0778, and the S/N = 7.763. Selectivity is a unitless quantity, and sensitivity is given in units of signal/concentration. Bias was calculated as the mean of the difference between replicate measurements of each concentration and the reported HPLC true concentration, and precision was calculated by the standard deviation of the difference between replicate measurements at each concentration and the reported HPLC true concentration. Because bias and precision values were calculated for each concentration, the

final reported values are the means calculated from the data set in its entirety. For EST on human skin, bias = -16.37 and precision = 1.149. For saturated solutions of EN on human skin, bias = 0.4706, and precision = 3.171, and for EN creams on human skin, bias = 16.03, and precision = 1.357. Both bias and precision are given in units of μg/g.

Discussion

Some multivariate calibration models offer the capability of wavelength and feature selection, such as principal component regression,[101] interval partial least squares,[102] and uninformative variable elimination.[111] These chemometric models require the construction of a large database of calibration samples, to ensure that all sample spectral variations are included in the calibration model. Then, the calibration model is validated by using it to predict the analytical results of a new set of samples through their spectra. This prediction is compared to the results from a reference analytical method (usually the same method employed to provide reference values for the calibration samples).

NAS calibration offers a distinct advantage over these methods, namely the potential for one-point calibration. When the experimenter has a set of pure component spectra for each of the system components, and a matrix of spectra without the analyte, it is a simple matter to identify the portion of the signal orthogonal to the interferences. If the NAS is the portion of the signal the experimenter is looking for, it will correlate highly to a change in concentration of the analyte of interest. If it is not the right signal, it will not correlate to the desired change in concentration. Therefore, NAS accomplishes with one-point calibration what other chemometric

models require an entire calibration set to accomplish. This offers a distinct advantage in time and in the computational burden for analysis of NIR spectra.

In our previous *in vitro* study with guinea pig skin, the skin contents of EN were found to be generally higher than with the human skin studied here.[82] For example, a statistically significant ($p<0.05$) 14-fold difference between the EN content in the human skin (103±21 μg/g) and the guinea pig skin (1497±345 μg/g, (7)) after a 15 h exposure to the saturated EN solution was observed (Student's t-test, SigmaStat, SPSS, Inc., Chicago, IL). For topical and transdermal studies, hairless guinea pig skin is a close approximation of human skin. Skin permeability studies with many compounds indicate that hairless guinea pig skin is more similar to human skin than either rat skin or hairless mice skin.[111-112] However, some morphological and biochemical differences do exist between the two, as the stratum corneum in the hairless guinea pig is composed of only a few layers.[103] The increased permeability of drugs through guinea pig skin, as compared to human skin, is most likely due to the lack of the multilayered stratified and resistant stratum corneum structure. Other studies have shown that there are also some immune and biochemical differences between guinea pig and human skin.[113] Each of these differences contributed to the decreased drug permeation in the human tissue as compared to the previously-studied guinea pig skin. Additionally, light scatter was noticeably more pronounced in human tissue, as is evident from Figure 5.7. A NIR scan of human tissue and guinea pig tissue treated with drug vehicle demonstrates that while the spectral features are similar between the two models, the absolute signal was far smaller in human tissue. This did not, however, affect the predictive ability of the NIR.

Great Flexibility and Durability. This experiment applied NIR spectroscopic analysis in four different applications; drug powders, drug in solution, and tissue concentrations of drug after exposure to solutions and creams. Additionally, it successfully measured drug concentrations in both guinea pig skin and human skin. The strong correlation of NIR concentration prediction to the HPLC results suggests that NIR spectrometry is a flexible technique for analysis of tissue concentrations of drugs and other chemicals after topical exposure. Once the method was established, no NIR spectra had to be discarded from a sample set consisting of 44 tissue samples and 264 scans, suggesting that NIR is a very reliable method of analysis. For the duration of this experiment, there were no failures or erroneous spectra. When the NIR does fail, a short noise spike from the preamplifier causes a very distinct and easily identifiable spectral feature, making it a simple matter to locate and discard.

NIR is also a nondestructive and rapid method of analysis, taking less than two minutes to complete each scan. In this research, samples were simply placed on a microscope slide and scanned, thus no sample preparation is required. NIR also has the capability of being a noninvasive method of analysis. Using a fiber-optic probe it is possible to scan skin tissue *in vivo* in whole animal or clinical studies. Skin samples in this research were scanned inside a closed metal chamber, thus external light and room noise were not a factor. In the chamber, NIR spectra are essentially free from external noise and interference. The limit of detection can be approximated as the concentration at which the S/N ratio is equal to three.[20] The lowest drug concentrations measured in this research were 11.5 μg/g for EST and 26.4 μg/g for EN, and in both cases the S/N ratios were greater than three. The NIR dynamic range can be expressed as the region of the NIR that responds linearly to a change in concentration. Since this experiment

did not exceed the upper or lower limits of detection, the dynamic range is simply the range of concentrations used. For EST, the dynamic range was 0 - 625.4μg/g, and for EN the dynamic range was 0 - 125.8μg/g. These results demonstrate that EN and EST can be measured in human skin at clinically relevant detection levels, because we have applied clinically relevant doses of EN cream and EST solution in the diffusion studies.

Conclusion

NIR spectrometry NAS multivariate regression demonstrated the ability to measure dermal absorption from estradiol in ethanol, econazole nitrate in propylene glycol, and a 1% econazole nitrate cream in human tissue samples with a very high degree of success. This suggests that NIR could serve as a noninvasive, rapid, and accurate alternative to the other topical bioavailability/bioequivalence analytical methods in use today.

Chapter Five Figures

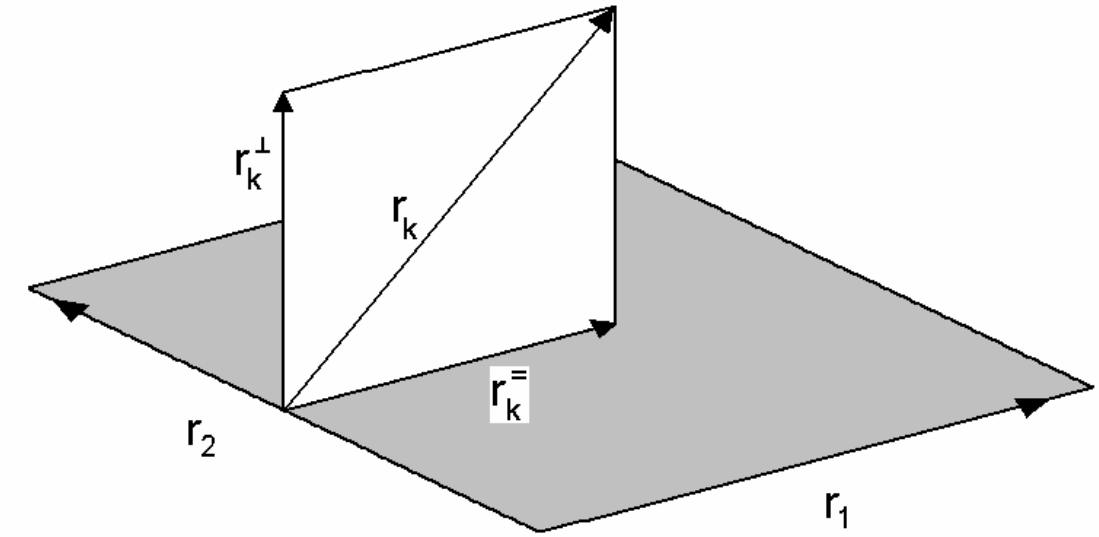

Figure 5.1: Figure depicting the NAS for any given analyte. Marked in gray is the space spanned by interferences $\mathbf{r}_1$ and $\mathbf{r}_2$. The spectrum of the analyte $\mathbf{r}_k$ is the sum of the NAS vector $\mathbf{r}_k^{\perp}$ and the interferences.

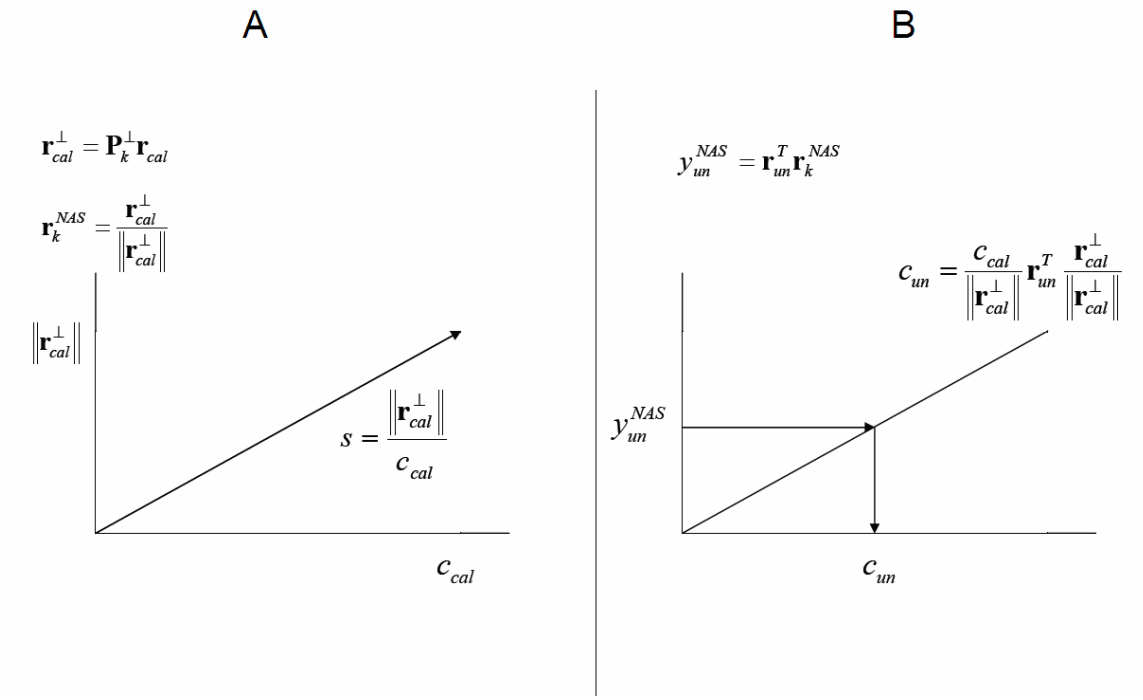

Figure 5.2: This figure summarizes the steps for multivariate calibration using the NAS. **(a)** The first step is to find the NAS direction and calculate the distance of the NAS calibration vector, and **(b)** the second step is to project the unknown spectrum in the NAS direction and calculate the unknown concentration.

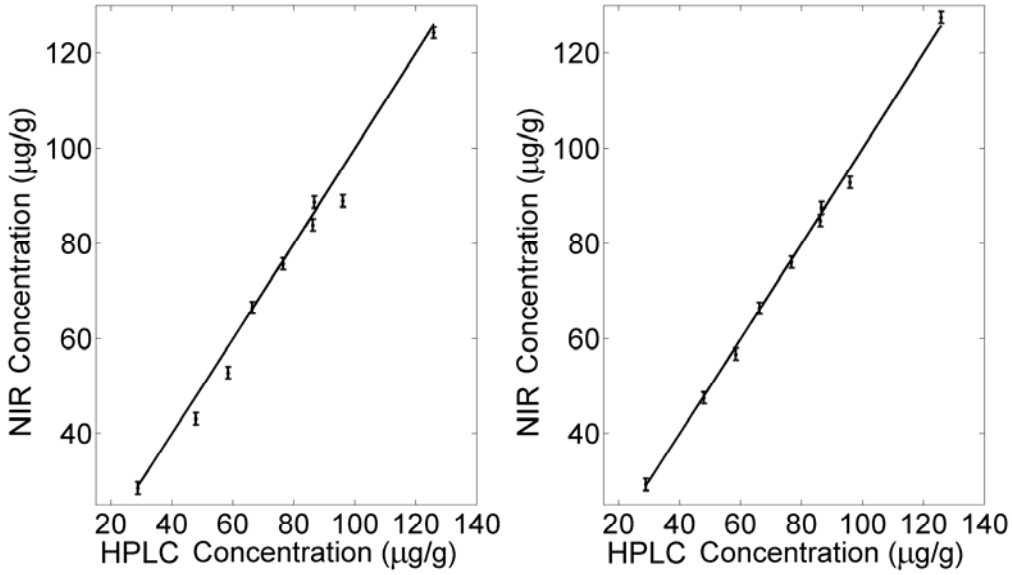

Figure 5.3: NIR calibration line from human skin samples treated with 200 µL of saturated econazole nitrate in propylene glycol, (left) epidermis calibration, r^2=0.988, SEE=2.46%, SEP=2.86%, (right) dermis calibration, r^2=0.996, SEE=1.98%, SEP=2.12%. %. The diagonal line illustrates the perfect correlation between HPLC and NIR, HPLC concentrations are shown on the x-axis, and the error bars are the means and standard errors of the NIR measurements.

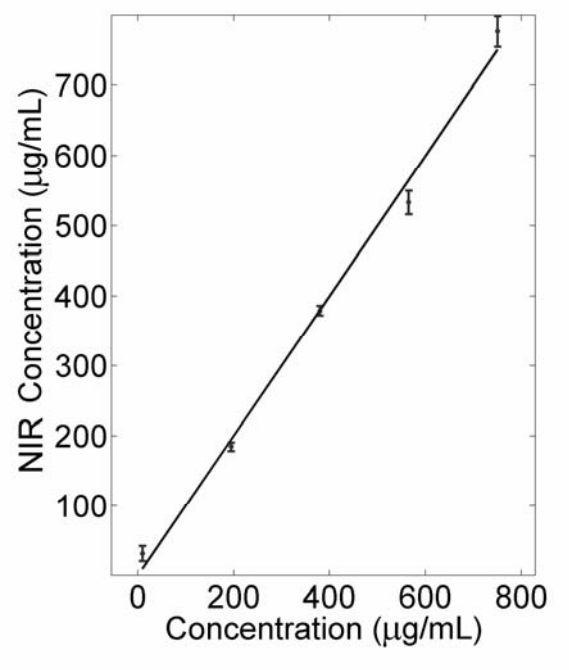

Figure 5.4: NIR calibration line for estradiol in ethanol using net analyte signal for the prediction of solution concentrations, r^2=0.988, SEE=3.46%, SEP=4.01%.

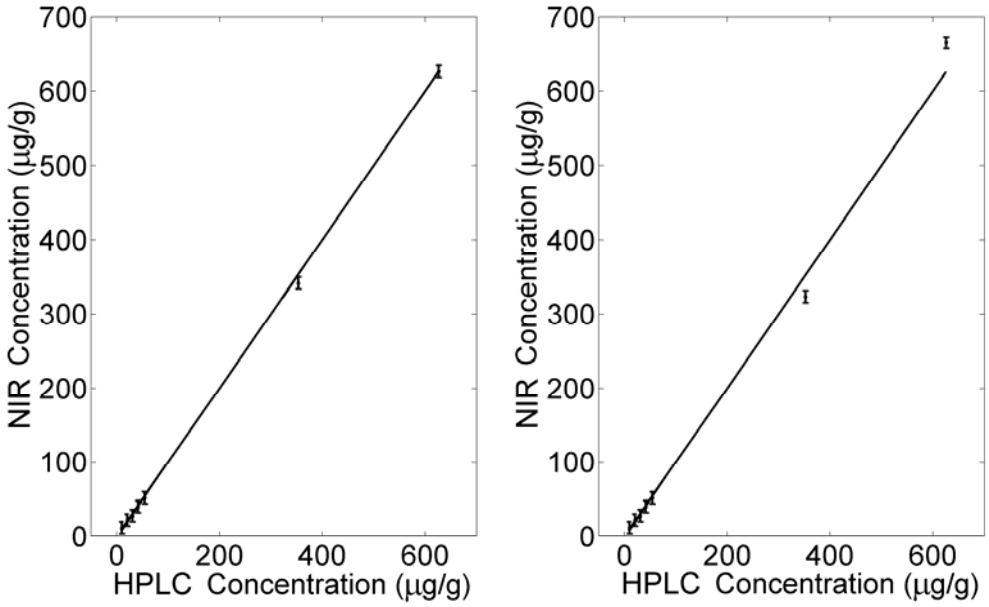

Figure 5.5: NIR calibration line from human skin samples treated with an applied dose of estradiol in ethanol, (left) epidermis calibration, r^2=0.987, SEE=3.30%, SEP=5.66%, (right) dermis calibration, r^2=0.967, SEE=5.53%, SEP=6.83%. The diagonal line illustrates the perfect correlation between HPLC and NIR, HPLC concentrations are shown on the x-axis, and the error bars are the means and standard errors of the NIR measurements.

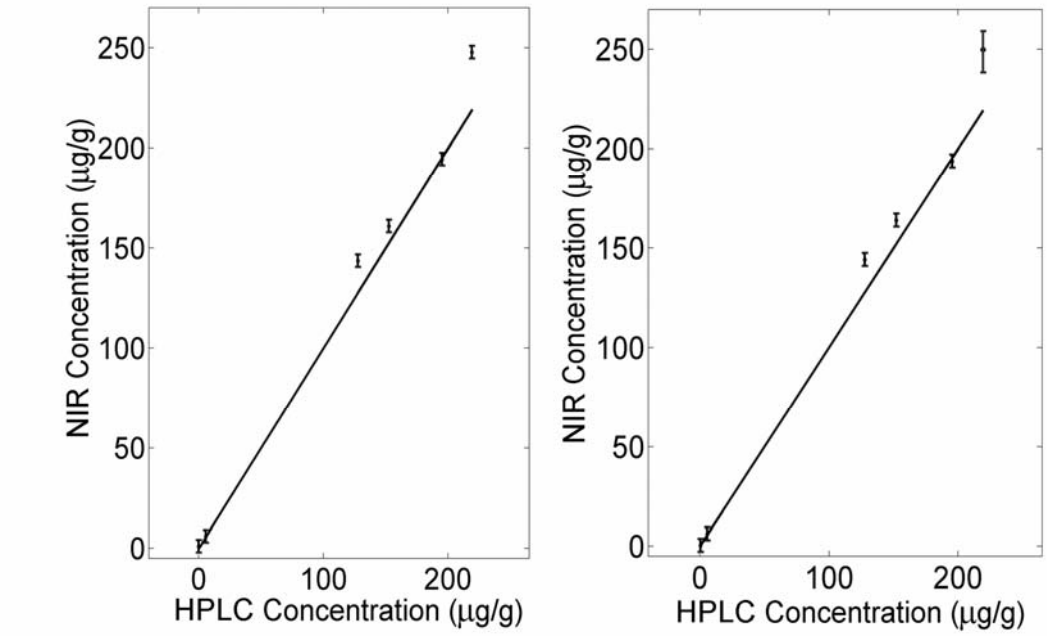

Figure 5.6: NIR calibration line from human skin samples treated with an applied dose of either 50 mg of placebo cream or 1% econazole nitrate cream, (left) dermis calibration, r^2=0.981, SEE=5.67%, SEP=6.07%, (right) epidermis calibration, r^2=0.980, SEE=5.93%, SEP=6.71%. The outlier is not included in this model.

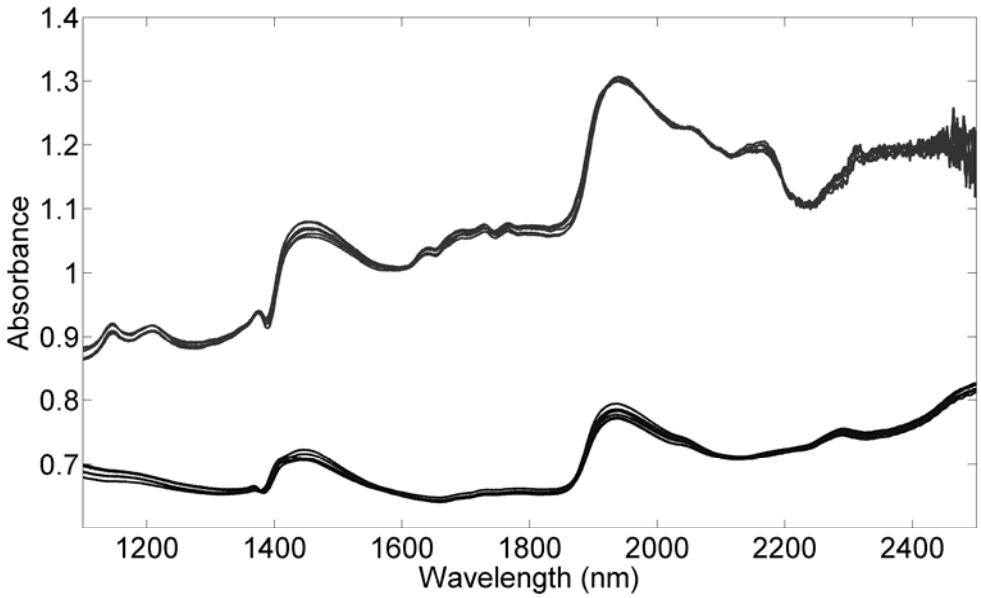

Figure 5.7: Raw NIR spectra from guinea pig skin (top) and human skin (bottom) treated with drug vehicle only (propylene glycol).

Copyright Statement

Joseph Medendorp, Kalpana Paudel, Robert A. Lodder, Audra L. Stinchcomb. *Pharmaceutical Research*. Accepted August 2006.

Chapter Six - Rapid near-infrared qualification of microcrystalline cellulose and sodium caprate minitablets through intact enteric coated capsules

Introduction

In this study we show the application of a nondestructive process analytical technology to separate two populations of enteric coated capsules that contained minitablets. In the course of a small, phase 1, clinical trial investigation it was deemed necessary to have two different placebo capsule formulations that were matched in appearance. These capsule batches were manually filled wherein subsequent processing required that they be specifically identified and accordingly classified (i.e., sorted). The sorting was based on a near-infrared (NIR) spectral differentiation between the capsule fill of the two formulations. Both capsule lots appear identical as enteric coated, white (titanium dioxide) hard gelatin capsules having an orange band. One lot contains pure microcrystalline cellulose (MCC) minitablets while the other lot contains minitablets composed principally of sodium caprate (C10) with PEG 3350 added as a binder. The enteric coating used on both lots is composed of methacrylate polymers with plasticizers at a cured coating level of approximately 14 mg solids per cm^2 of capsule surface area. In order to facilitate a streamlined manufacturing schedule, it was necessary to rapidly develop and GMP qualify a technically feasible, nondestructive capsule identification and sorting process.

Near-infrared (NIR) spectrometry has successfully been applied to the noninvasive and nondestructive differentiation of various capsules, tablets, and drugs in solution.[32,45,104,114-115] While current scholarship suggests that NIR is fairly well established for this purpose, previous studies focus on the utility of NIR for the quantification and identification of capsules filled with powders and solutions. In the current case, the formulations in the capsules (methacrylate coated; orange-banded, hard gelatin) were not powders or solutions, but were minitablets. Due to the large particle size (2 mm) and random orientation of the minitablets in the capsules, the NIR absorbance spectra demonstrated more variation than a uniformly filled powder capsule.

The combined effect of the enteric-coated hard gelatin exterior, and the random distribution of the minitablet interior, made analysis of these capsules potentially more challenging than those in previous experiments. We undertook a combination of selective wavelength analysis along with chemometric principles to develop a non-destructive method for distinguishing between these two capsules. The method was subjected to validation testing using a calibration set and was ultimately used to sort through two populations of capsules totaling 3407 units using only two near-IR wavelengths provided by interference filters from a tungsten-halogen source. Such a simple instrument can be easily manufactured and widely deployed as a process sensor.

NIR Spectrometry and Chemometrics. The near-infrared region of the electromagnetic (EM) spectrum covers wavelengths from 750 nm to 3000 nm. This region consists of broad, overlapping peaks that result from overtone bands, combination bands, and difference bands from molecular vibration of CH, NH, OH, and SH bonds.[116] Many factors contribute to the variations in NIR spectra, such as detector noise, environmental conditions, and different sample preparation. Often the largest variations come from the constituent or concentration differences. Using this knowledge, chemometric techniques such as principal axis transformation can be used to interpret complex overlapping spectra by placing the original spectral variables into a new, smaller coordinate axis system.[9] Calculation of the principal components is accomplished by a singular value decomposition of matrix A according to $\mathbf{A} = \mathbf{USV}$, where *A* is the matrix of original spectra, U is the matrix of eigenvalues (scores), S is a diagonal matrix of singular values, and V is the matrix of eigenvectors (loadings). The first principal component may capture 80% or more of the total variance. Many of the eigenvalues model only noise, therefore, only those

that contain a significant proportion of the variation with analytical signal are used in calibration and evaluation of samples.

BEST Algorithm. Principal components can be used to illustrate the separation of two groups. For a more quantitative result, the bootstrap error-adjusted single sample technique (BEST) was applied to the calibration data. The BEST algorithm begins by encoding the intensity on each wavelength as a separate dimension, thus reducing each spectrum to a single point in multidimensional space. Population **P** is an *m x n* matrix in hyperspace **R** whose rows are the individual samples and the columns are the frequencies.[30] BEST considers each wavelength from a spectrum of *n* wavelengths to be taken as a separate dimension, such that each spectrum is reduced to a single point in *n*-dimensional hyperspace.[32] **P*** is a discrete realization of **P** based on a calibration set **T** of the same dimensions as **P***. This realization is chosen one time from **P** to approximate all possible sample variations present in **P**. **P*** has parameters **B** and **C**, where **C** = E**(P)** and **B** is the Monte Carlo approximation to the bootstrap distribution. The expectation value, E**(P)**, is the center of **P**, and **C** is a row vector with the same number of rows as there are columns in vector **P**. New test spectra **X** are projected into hyperspace **R** containing **B**, rows of **B** are mapped onto a vector connecting **C** and **X**. **C** and **X** have the same dimensions. The integral over **R** is calculated from the center of **P** in all directions. A skew-adjusted standard deviation (SD) is based on the comparison of the expectation value **C**=E**(P)** and **C**=med**(T)**, the median of **T** in hyperspace projected on the hyperline connecting **C** and **X**. The result is an asymmetric SD that provides two measures of the SD along the hyperline connecting C and X. Equation 6.1 defines the SD in the direction of X, and Equation 6.2 defines the SD in the

opposite direction. Skew adjusted SDs can be used to calculate mean distances between spectra of different samples.

$$+\vec{\sigma}\left|\frac{\int_{0}^{+\sigma}\left(\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)\right.}{\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)}\right| = 0.34 \qquad \textbf{6.1}$$

$$-\vec{\sigma}\left|\frac{\int_{0}^{-\sigma}\left(\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)\right.}{\int_{R} P^{*} \rightarrow \left(\overrightarrow{CX}\right)}\right| = 0.34 \qquad \textbf{6.2}$$

Methods

Separation of MCC and C10 capsules with full NIR spectra. The first objective was to prove that full NIR spectra from 1100-2500 nm (in 2 nm steps with 10 nm bandpass) could be used to separate the MCC and C10 capsules. Initial experiments involved only three MCC capsules and three C10 capsules. Spectra were collected with a NIR spectrometer (Technicon InfraAlyzer 500, Tarrytown, NY) interfaced to a computer (OptiPlex GXM 5166, Dell, Round Rock, TX, USA) running SESAME 3.1 (Bran+Luebbe, Norderstedt, Germany). Each capsule was scanned three times for a total of 18 spectra (Figure 6.1). The capsules were scanned in random order and rotated following each scan to average possible sample variations due to inconsistencies in the enteric coating, gelatin layer and the orientation of the minitablets. Scans were collected inside of the instrument drawer to eliminate room noise and external interferences. All data were exported to Matlab 7.0.1 (The Mathworks Company, Natick, MA) for processing and analysis.

Data were multiplicative scatter-corrected to eliminate baseline variations,[85] and second derivatives[49] were plotted to find the regions in the spectrum where the most variation was apparent (Figure 6.2). Principal components (PCs) were calculated and the PCs with the largest contribution to variation were plotted in two and three dimensions to visualize the differences between capsule groups (Figures 6-3 and 6-4).[9] BEST standard deviations and cross-validation standard deviations (CV-SD) provided a quantitative measure of group separability.[30,32]

In order to estimate the limits of detection for the NIR, a cluster translation procedure was performed.[117] NIR spectra from the two capsule populations, $\mathbf{P_1}$ and $\mathbf{P_2}$ are expressed in *m x n* matrices, where n is the number of wavelengths and m is the number of spectra. The columns of the matrices are averaged by Equations 6.3 and 6.4, giving two *1 x n* vectors.

$$P_{1ave} = \frac{1}{m}\sum_{i=1}^{i=m} P_{1i} \qquad \textbf{6.3}$$

$$P_{2ave} = \frac{1}{m}\sum_{i=1}^{i=m} P_{2i} \qquad \textbf{6.4}$$

A difference spectrum **X** is calculated from $\mathbf{P_{2ave} - P_{1ave}}$. One population was spatially translated toward the other, $\mathbf{P_{Adjusted} = y*X+P_2}$, where **y** starts at zero, increasing in increments of 0.01 until $\mathbf{P_1}$ and $\mathbf{P_{Adjusted}}$ are inseparable. It is assumed that the two capsule groups represent the possible variations in the pure component spectra (MCC and C10), and that all points on the hyperline connecting the centers of the two population distributions correspond to mixtures of the two components because the Beer-Lambert Law holds. For example, when one population is translated one-half the distance toward the other, that population corresponds to a 50/50 mixture

of MCC and C10. The maximum distance the two groups can move toward each other while maintaining statistical separation determines the minimum quantities of MCC and C10 that can be detected in each other.

Rapid separation of MCC and C10 capsules with fewer wavelengths. Although the overall goal of this project was to develop a process analytical method sensitive enough to assign the MCC and C10 capsules to their respective groups accurately, it was also imperative that the project be completed as quickly as possible. This time constraint required a modification of the full spectrum approach. Experiments were conducted to accomplish the separation with as few wavelengths as possible yet still with adequate statistical assurance of specificity. Using the same six capsules scanned above at 701 wavelengths, new spectra were collected with a 19-wavelength NIR filter InfraAlyzer FLEX spectrometer (Bran+Luebbe, Elmsford, NY) between 1445 and 2348 nm. Capsules were placed in the conical reflective cup oriented with the thicker cap end down, body end up, and held in place with a steel rod (Figure 6.5). The conical reflecting cup is designed such that when a sample capsule is placed along the axis of radial symmetry of the cone, specular reflection at the detector is minimized while diffuse reflectance is maximized.[104] All radiation that follows a path parallel to the incident beam and perpendicular to the base of the conical reflector is reflected and collimated back toward the source. This radiation is predominantly specular reflectance and contains little information about the capsule fill. In the same fashion, a small amount of radiation that passes through the capsule but does not scatter is also returned to the source. Therefore, the majority of the radiation reaching the detector via the integrating sphere is scattered by the contents of the capsule. The amount of radiation that reaches any given location on the capsule is directly proportional to the curved

surface area of the frustum (the conic section defined by two parallel lines from the light source, and a plane parallel to the reflector base connecting the two lines) in which it lies. Therefore, more light reaches the top of the capsule than the bottom of the capsule because there is more curved surface area at the top of the reflector than at the bottom. The curved surface area is given by $\pi s(r_1+r_2)$ where r_1 and r_2 are the radii of the base and top of a circular frustum, s is the length between the top and bottom measured along the surface of the cone. If the detector collected scattered light from cross-sections of the capsule, there would be a different intensity value for each cross-section. However, the detector uses an integrating sphere, which collects all of the scattered light from the entire capsule. This configuration eliminates the concern of uneven illumination along the capsule.

In order to best accommodate the capsules, the spectrometer was inverted, and an instrument drawer was fashioned for consistent and reproducible sample loading. Each scan took approximately two minutes to collect the full 19 wavelengths. Capsules were scanned in random order and rotated to average sample positioning variations. Principal components, intercapsule BEST standard deviations (SDs), and intracapsule cross-validation standard deviations were calculated from the resulting spectra. Spectra for the two groups of capsules were very different, allowing a visual inspection of the spectra to sufficiently identify the more distinguishing wavelengths. Standard deviations and cross-validation-SDs were calculated from four selected wavelengths. To operate as quickly as possible while maintaining the highest level of capsule classification accuracy, the two most distinguishing wavelengths were selected by visual inspection. From these wavelengths, ratios were calculated and plotted to prove that the two wavelength approach was sufficient to justify exploration of a larger validation set.

Equipment Calibration. The FLEX spectrometer was installed in the GMP facility in the University of Kentucky Center for Pharmaceutical Science & Technology, and turned on for the remainder of the procedure to eliminate detector drift due to thermal variations. A polystyrene calibration film standard (Perkin Elmer Corp., Norwalk, CT, USA) fitted to a conical reflective cup was scanned 50 times to capture all possible sample variations. A ratio was calculated from the signal intensity at 1734 nm and 1445 nm for each scan of polystyrene standards, and confidence limits were constructed around the mean at ± six standard deviations, 1.0245 ± 0.0118. The choice of a 12 standard deviation acceptance range was made to ensure essentially 100% confidence limits on classifications for the set of 3407 capsules to be scanned. To prove that the instrument response was the same from day to day, these ratios were projected on top of the predefined constructed confidence limits. The same approach was used to identify the different capsules. A calibration set consisting of fifty MCC and fifty C10 capsules was scanned. Wavelength ratios were calculated and confidence limits were constructed at the mean ± six standard deviations for each of the two capsule formulations. These limits were defined as the acceptance criteria boundaries for the determination of which group the incoming capsule belonged (Figure 6.6).

Capsule identification screening using the reduced-wavelength approach and confidence limits. At the start of each day of data collection, the polystyrene calibration standard was scanned to prove the instrument was performing reproducibly. Capsules were sequentially scanned in groups of twenty; a total of 3407 capsules were scanned. An algorithm was written to automatically calculate and display the wavelength ratios on top of the predefined confidence

intervals (Figure 6.7). Capsules were sorted according to where their NIR ratios projected relative to the calibration experiments. Chemical structures for microcrystalline cellulose and sodium caprate can be found in Appendix B.

Results and Discussion

Figure 6.8 illustrates the two wavelengths that were selected to distinguish between the two groups of capsules. The inert MCC has a relatively flat spectrum while C10 has a very steep spectrum at these wavelengths. It is this spectral feature that allowed the simple calculation of a ratio from two wavelengths to distinguish between the two groups. Cross-validation results and the BEST standard deviations between capsules are reported in Table 6.1 for the full spectrum NIR measurements, for four wavelengths, and for two wavelengths. Note that the standard deviations are the same magnitude with two wavelengths as with the full spectra, suggesting that the basis of selectivity is improved by focusing on those wavelengths which can unequivocally distinguish between groups. This approach justifies and allows for a much smaller data set with which to base the classification upon. Illustrated in Figure 6.9 are the ratios (1734/1445 nm) calculated from all capsules projected into the confidence intervals determined during the calibration stage of the experiment. This experiment resulted in greater than 99.71% successful capsule identification. Of the 3407 capsules scanned, ten capsules projected directly on or just outside their respective decision boundaries. These outliers were scanned a second time at the end of the experiment, and their two-wavelength ratios projected inside their respective confidence intervals. These results proved that the unclassified capsules were a result of erroneous manual sample loading, and further demonstrated the consistency and integrity of both the method of analysis as well as the capsules themselves.

Detection limits were estimated by the spatial cluster translation experiment described in the Methods above. When using MCC capsule spectra as the calibration set and C10 capsule spectra as the test set, C10 capsules could be spatially translated 92% of the distance across space before the populations were inseparable. When using C10 as the calibration set and MCC as the test set, MCC was spatially translated 89% of the distance across space before the populations were inseparable. This demonstrates that NIR is capable of nondestructively identifying mixtures of C10/MCC in the capsules down to approximately ten percent of each one in the other. It is apparent that the NIR detection limits far exceed what is necessary to distinguish between MCC and C10 capsules.

Conclusion

This study presents an effective application of NIR spectrometry to the noninvasive and nondestructive classification of MCC and C10 minitablets contained in enteric-coated (methacrylate polymer) hard gelatin capsules. The experiment was conducted in a GMP facility and is relevant to PAT. It was a very rapid method with data collection times of only 10 seconds, and the two wavelength approach gave unequivocal separation between the two capsules. The entire process, from conception to completion, required only 20 days. Of the 3407 capsules scanned, greater than 99.71% projected into their respective confidence intervals. Of the ten capsules that initially failed to validate, it was proven that human sample loading errors were responsible, suggesting that an automated version of the same experiment could have classified 100% of the capsules correctly. One of the primary goals of the FDA PAT initiative is to increase automation to reduce human error. This experiment provides an example of the

benefits of automation, as well as of the utility of a method for real-time characterization and release of individual drug products.

New measurement, control and information technologies are needed in PAT to predict, control and assure product quality and performance. Using an appropriate PAT sensor, product quality attributes can be accurately and reliably predicted over the design space established for the materials used, the process parameters, and the environmental and other conditions. A two-wavelength near-IR sensor for PAT will be rugged, inexpensive, and simple to construct using two interference filters and a detector (e.g., PbS or InGaAs). Such a sensor would be a dynamic tool for process innovation and continuous quality improvement using risk-based models for inspection.

Chapter Six Tables

Table 6.1: BEST standard deviations reported for the broadband, 4 wavelength, and 2 wavelength calibration data. Using the BEST, a successful measure of statistical separation between two populations was defined as a standard deviation greater than three, and a measure of statistical inseparability was a standard deviation less than three. The table reports intercapsule standard deviations, all greater than three, and reports intracapsule standard deviations along the diagonal, all less than three.

Broadband	**MCC**	**C10**
MCC	**1.70**	**29.54**
C10		**1.85**
4 λ	**MCC**	**C10**
MCC	**1.52**	**25.56**
C10		**1.58**
2 λ	**MCC**	**C10**
MCC	**1.24**	**23.46**
C10		**1.27**

Chapter Six Figures

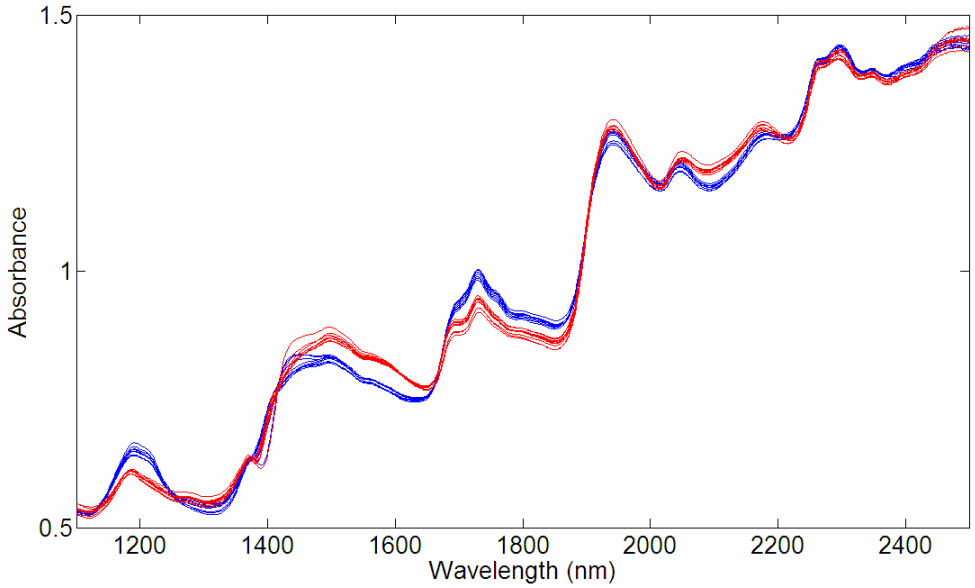

Figure 6.1: Broadband NIR spectra of C10 (blue) and MCC (red) capsules collected with an InfraAlyzer 500 scanning monochromator instrument.

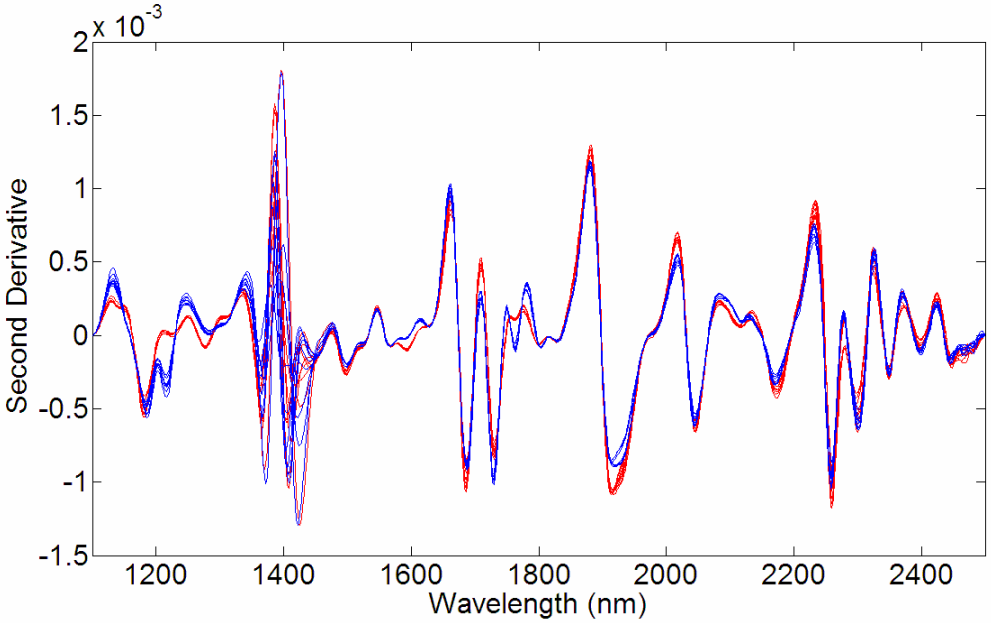

Figure 6.2: Second derivative spectra calculated from the broadband spectra of C10 (blue) and MCC (red) capsules.

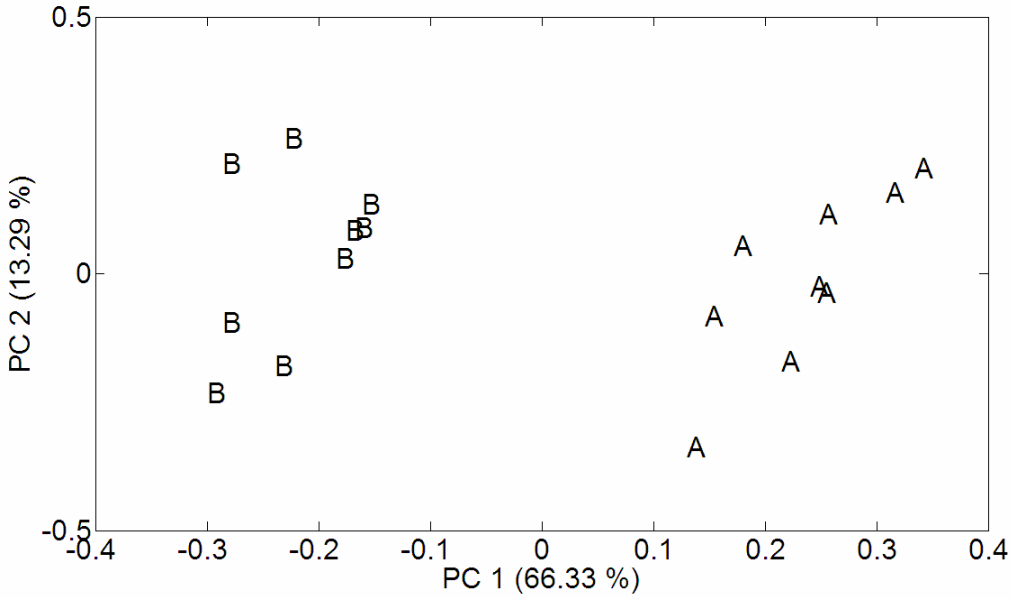

Figure 6.3: The two-dimensional principal component plot from broadband calibration data demonstrates how closely the capsules cluster in space, C10 capsules (marked with 'A') project on the right and MCC capsules (marked with 'B') project on the left.

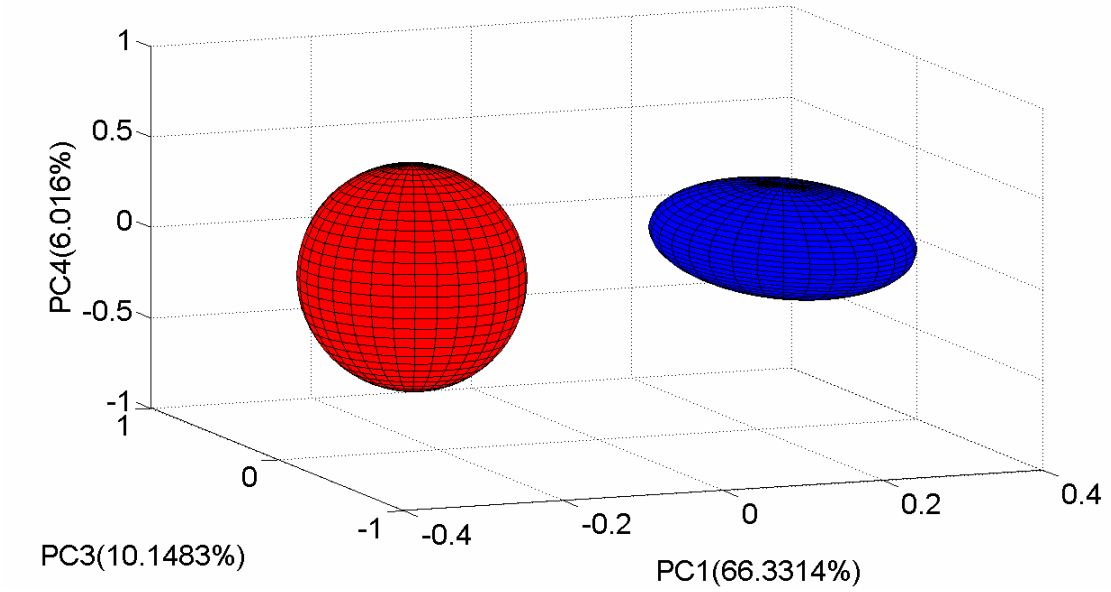

Figure 6.4: Three-dimensional principal component (PC) plot from broadband NIR spectra. The majority of the separation is seen along the PC 1 axis, which captures 66.33% of the total variation. This separation suggests that the difference in capsule constituents, C10 (blue) vs. MCC (red), is mainly responsible for the separation between the two groups of spectra.

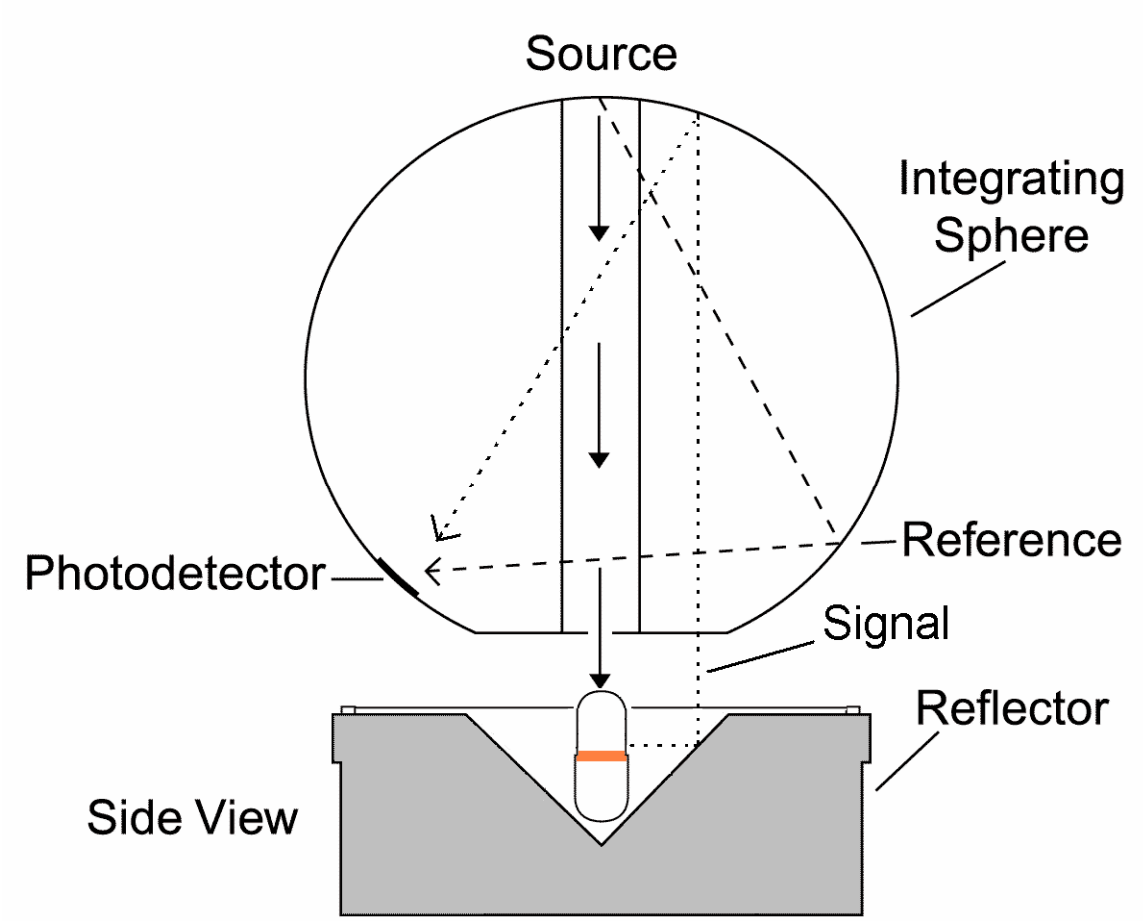

Figure 6.5: 90^{o} conical reflective cup used as a capsule holder. The cup is designed so that when a sample capsule is placed along the axis of radial symmetry of the cone, specular reflection at the integrating sphere detector is minimized while diffuse reflectance is maximized. As a result, the majority of the light reaching the detector has been scattered by the contents of the capsules.

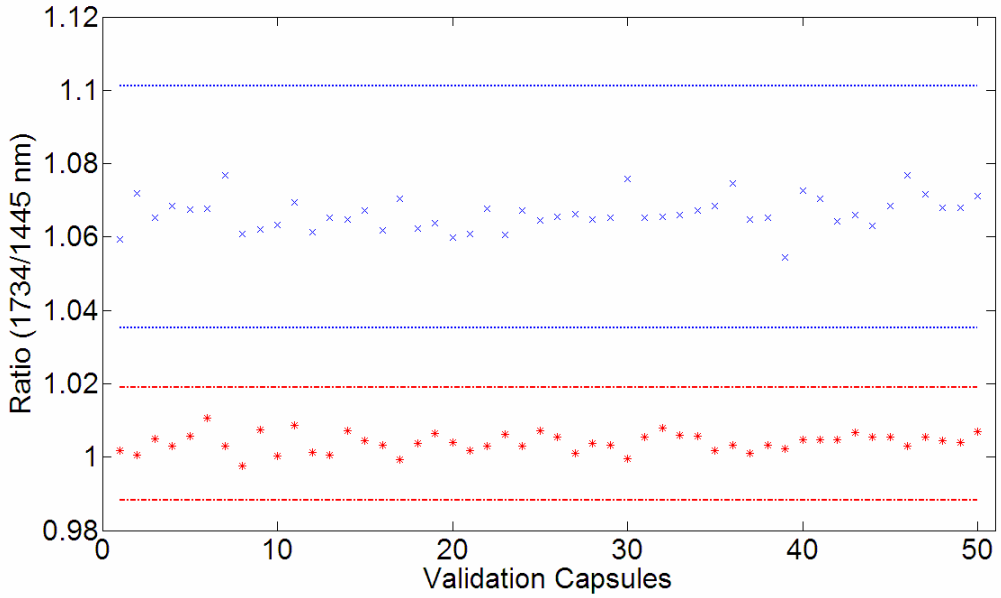

Figure 6.6: Ratio of the signals at 1734 and 1445 nm from the validation set of capsules. The steeper curve has the higher ratio, thus C10 projects on top (**x**), and the flatter curve has a lower ratio so the MCC capsules project on the bottom (*). The bars above and below the data sets are the confidence intervals, placed at six standard deviations above and below the mean.

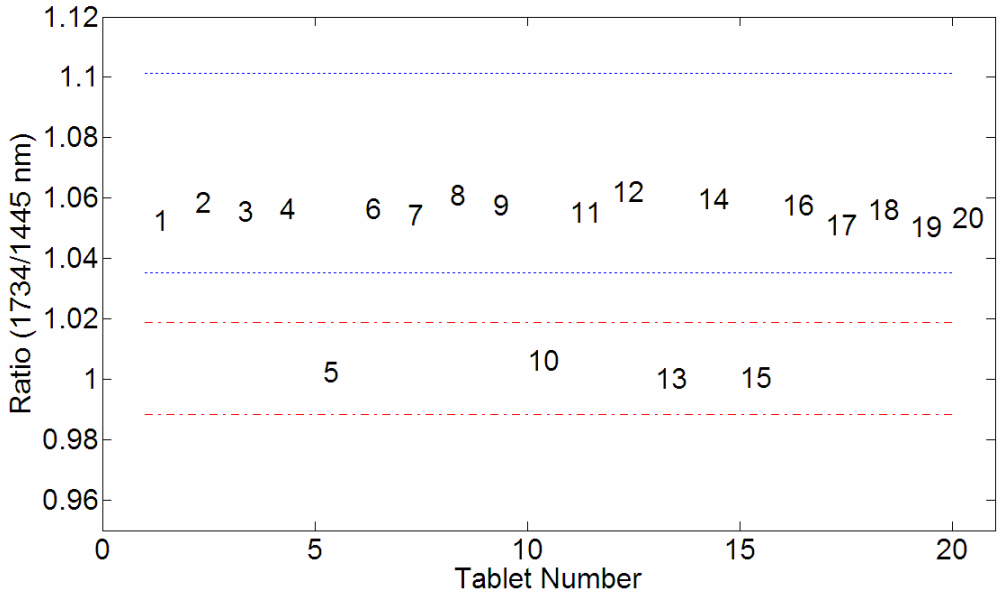

Figure 6.7: Real data sets were collected in groups of twenty capsules. An algorithm was written to output this plot to show the two-wavelength ratio for each capsule.

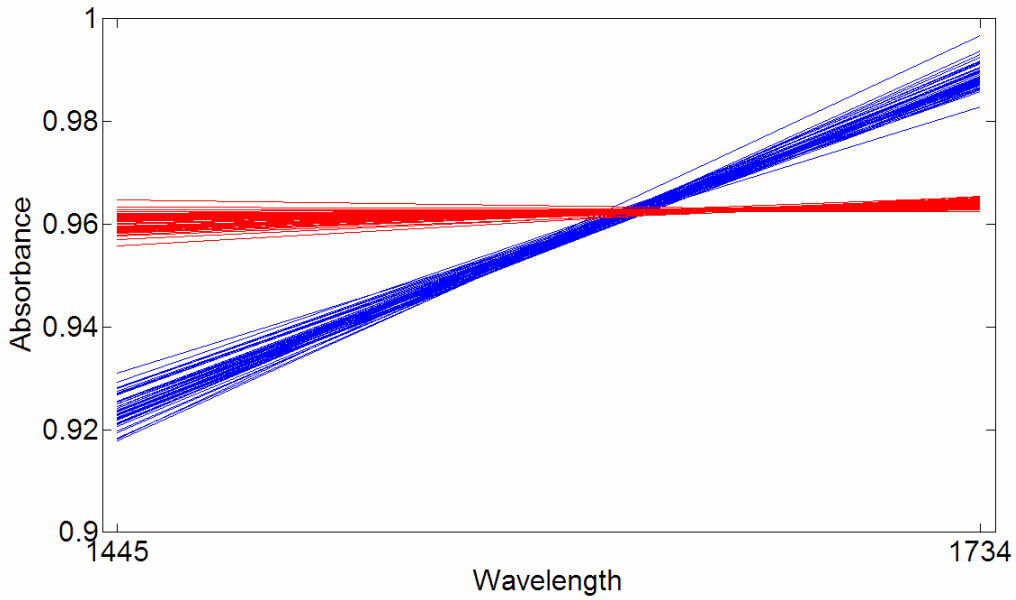

Figure 6.8: Data collected with an InfraAlyzer FLEX spectrometer at two NIR wavelengths (1734 nm and 1445 nm) collected from the two capsule validation set. MCC is the flat curve and C10 is the steeper curve. Of the 100 spectra displayed in this plot, there is obviously no overlap between the two different groups.

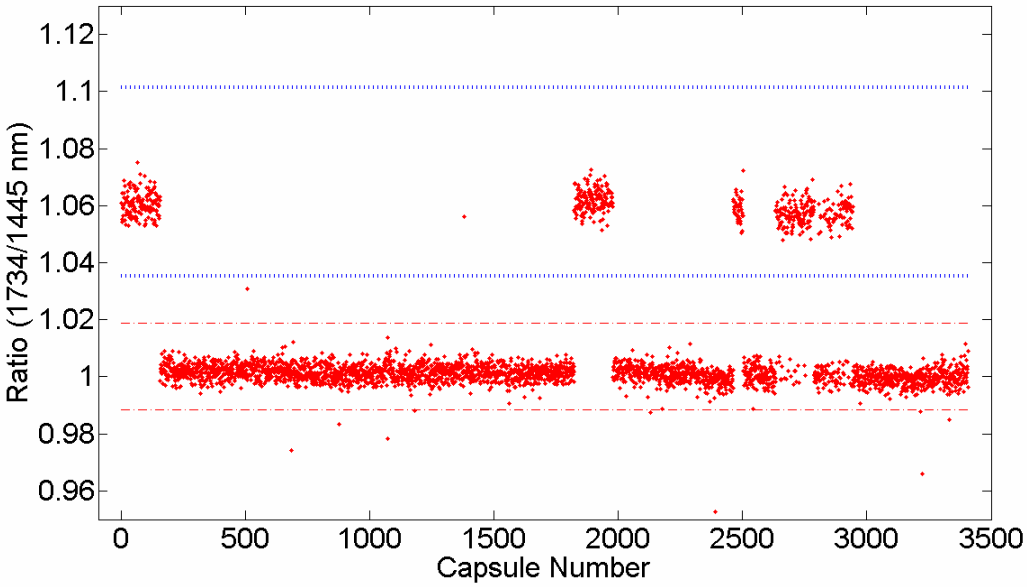

Figure 6.9: Two-wavelength ratios from all 3407 capsules projecting inside their respective confidence intervals, C10 (top clusters) and MCC (bottom clusters). This figure illustrates greater than 99% successful identification.

Copyright Statement

Joseph Medendorp, Joseph Wyse, Robert A. Lodder, Lloyd G. Tillman, Sujatha Dokka, Michael Jay. *The Journal of Process Analytical Technology.* **2006**; 3(3): 11-17

Chapter Seven - NIR spectrometry for the characterization of solid rocket booster fuel components in a novel tamper-resistant pill bottle

Introduction

OxyContin, the brand name for Purdue Pharma's narcotic pain reliever Oxycodone-HCl, is a Schedule II drug under the Controlled Substances Act due to its propensity for abuse and dependency.[118] It is an opium-based pain reliever prescribed for relief of moderate to severe pain; however, it exhibits heroin-like effects lasting up to twelve hours when abused. The illicit diversion of pharmaceuticals such as OxyContin is a pervasive problem across the country.[119] The chief source of diverted pharmaceuticals on the illicit market has historically been corrupt pharmacists and doctors.[120] However, as law enforcement agencies crack down on pharmacists, technicians, and fraudulent doctors, the burden of drug protection falls on drug suppliers and patients. Protecting drugs continues to cost millions of dollars for heightened security measures, including the redesign of pharmacies and drugstores so they bear resemblance to modern fortresses, where drugs are delivered in armored vehicles and stored in heavy vaults protected by guards. Repeated robberies at gunpoint suggest the times have long since passed when child-proof caps were sufficient to discourage drug abuse. Rather than redesigning architecture in pharmacies and fortifying drugstores, this manuscript proposes a more comprehensive and less expensive measure of additional security for patient's prescription drugs. The time is right for a well-secured, better-regulated pill dispensing system. Companies like e-pill, Inc. have attempted to respond with their Monitored Automatic Pill Dispenser (MD.2), which features voice alarms and reminders.[121] To the author's knowledge, the MD.2 is the only automated vault-like delivery system on the market. At a retail price of $899.00 US, dispensing new MD.2 bottles with each monthly refill is prohibited by cost. Therefore, the MD.2 comes with a lock and key, and it is the responsibility of the patient to refill their bottles. Potential thieves need only to obtain the key to pilfer the MD.2 contents. The MD.2 system offers only one line of defense against would-be

thieves. An opportunity lies in building an inexpensive and impenetrable container as a fail-safe second line of defense, capable of scheduling and dispensing medications like OxyContin, and deterring those interested in obtaining the drugs purely for abuse and illicit purposes.

Pill Safe. In response to the need for a better protected pill bottle, the Pill Safe precludes the possibility of theft and illegal dissemination of prescription drugs. The Pill Safe simply presents the patient with a button. When pressed, it dispenses the medication through a small aperture if and only if the prescribed dosing period has passed since the previous pill was dispensed. Meanwhile, the Pill Safe monitors its outer shell for tampering, destroying all of the pills rapidly upon tamper detection. Figure 7.1 gives a block diagram of the system. The mechanism shown in Figure 7.2 houses and delivers the pills and destroys them at the direction of the microcontroller. The pills are stored in columns adjacent to solid rocket booster (SRB) fuel. In this configuration hot exhaust gases from the burning fuel are directed toward the pills. The entire assembly is enveloped in a protective shell with a conductive loop printed on the interior using conductive ink. Breaching the shell breaks the loop, signaling the microcontroller to ignite the SRB fuel. The entire system is powered by 2 alkaline AA batteries, and constructed from inexpensive parts, costing less than $10 US.

NIR spectrometry can be used in the design of SRB fuel mixtures for the quantification of SRB fuel components and for the prediction of burn characteristics using NIR spectra of different fuel mixtures. When the fuel formula changes, the burn characteristics change; therefore, it is imperative that the fuel mixtures have the right constituents in the right amounts. For example, when the fuel mixture is deficient in ammonium perchlorate the burn duration increases

significantly, giving would-be thieves crucial extra seconds to break into the pill bottle. Therefore, NIR was chosen to monitor the component concentrations and provide quality control. NIR has previously been used with great success for the quantification of fuel components in both solid and liquid propellant mixes.[122-123] NIR, FT-IR, and transmission spectrometry have been used for accurate and precise quality control analysis of fuel pre-mixes, and have demonstrated the ability to monitor antioxidant depletion over time.[124] In-process reaction information such as intensity distribution, ignition processes, reaction temperatures, reaction species and their concentrations have been studied using the spectral range from UV, visible, NIR, and mid-IR.[125] Safety considerations for sample analysis using NIR for the study of rocket propellants have also been addressed, such as the thermal response to NIR exposure, and the effects of Raman spectrometry on the mixtures.[126]

Theory

Absorbance in the NIR region of the electromagnetic spectrum is primarily a result of overtones and combinations of the fundamental bands from the mid- and far- infrared regions. The bands are a result of anharmonic stretching and bending of functional groups such as N-H, O-H, C-H and C=O. In most cases, the molecular structures are sufficiently complex that the spectral features of interest are highly overlapping, and thus not directly usable without statistical analysis.[110] The formula for the pill safe rods used to destroy pills included aluminum dust with iron oxide catalyst (fuel), NH_4ClO_4 (oxidizer), bisphenol A/epichlorohydrin (casting agent), and a polyamide resin (curing agent). The latter three components are primarily the spectroscopically active constituents in this mixture; therefore, their concentrations were used for

building and testing the regression model. Rather than using absolute mass, the prediction model used the constituent percentages (by mass) of the total mixture.

Because NIR spectra are usually a linear combination of pure component spectra according to Beer's Law, it is theoretically possible to determine the concentration of each individual component in overlapping spectra. However, there are noise factors that tend to complicate this procedure, such as sample inhomogeneity, particle size differences, and temperature drift. For a linear calibration model, samples must be identified as outliers and removed from the calibration, or the noise and variation must be incorporated into the model. In this work, Hadi outlier detection was used to identify the spectra that belonged in the calibration model.[127] Outlier detection is generally approached by forming a clean subset of data **M** (free from outliers), followed by testing the fit of the remaining points relative to the clean subset. Consider the regression in Equation 7.1:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \textbf{7.1}$$

where **Y** is an *n x 1* vector of responses, **X** is an *n x k* matrix of responses and observations, $\boldsymbol{\beta}$ is a vector of estimated regression coefficients from fitting the model to **M**, and $\boldsymbol{\varepsilon}$ is a matrix of errors. The best clean subset **M** is found by the deletion of the variables that result in the largest reduction in $SSE_{M,}$ the residual sum of squares.

When building a calibration model from NIR spectra, it can be helpful to visualize the total instrument response, $\mathbf{r}_k$, as the sum of two orthogonal components, the interferences, $\mathbf{r}_k^{=}$, and the net analyte signal (NAS), $\mathbf{r}_k^{\perp}$, according to Equation 7.2.[22] The superscript '=' denotes the fact

that the interferences span the space occupied by the analyte, and the '⊥' denotes the fact that the NAS is orthogonal to the interfering species.

$$\mathbf{r}_k = \mathbf{r}_k^{=} + \mathbf{r}_k^{\perp} \qquad \textbf{7.2}$$

A linear combination of the interferences produces $\mathbf{r}_k^{=}$, therefore the signal orthogonal to $\mathbf{r}_k^{=}$ belongs exclusively to the analyte of interest.[32] A projection matrix is calculated according to Equation 7.3:

$$\mathbf{P}_k^{\perp} = (\mathbf{I} - \mathbf{R}_{-k}\mathbf{R}_{-k}^{+}) \qquad \textbf{7.3}$$

where $\mathbf{R}_{-k}$ is a matrix of samples without the analyte, **I** is the identity matrix, and the superscript "+" is the Moore-Penrose pseudoinverse. The NAS vector, $\mathbf{r}_{cal}^{\perp}$ can be calculated from a calibration spectrum, $\mathbf{r}_{cal}$ by projection of the spectrum onto the null space of the rows of **R** with Equation 7.4:

$$\mathbf{r}_{cal}^{\perp} = \mathbf{P}_k^{\perp}\mathbf{r}_{cal} \qquad \textbf{7.4}$$

The NAS vector is normalized to length one by Equation 7.5:

$$\mathbf{r}_k^{NAS} = \frac{\mathbf{r}_{cal}^{\perp}}{\left\|\mathbf{r}_{cal}^{\perp}\right\|} \qquad \textbf{7.5}$$

A linear regression is fit to this vector, and the regression coefficients are used for subsequent predictions.

One of the advantages of using the NAS approach is for the calculation of so-called figures of merit. In severely overlapping spectra, it has historically been difficult to quantify selectivity, sensitivity, and signal-to-noise (S/N) because of the inability to distinguish between interferences and the analyte of interest.[19-20] With the NAS, these quantities can be measured directly.

Selectivity can be calculated as the scalar degree of overlap, α, between the NAS vector and the calibration spectrum according to Equation 7.6:

$$\alpha = \frac{\left\| \mathbf{r}_{cal}^{\perp} \right\|}{\left\| \mathbf{r}_{cal} \right\|} \qquad \textbf{7.6}$$

The selectivity is a measure from 0 to 1 indicating how unique the analyte of interest is compared to the interferences. The sensitivity is a measure of how much the analyte varies in response to a change in concentration. This quantity can be expressed as Equation 7.7:

$$\mathbf{s}_k = \mathbf{r}_k^{NAS} / c_k \qquad \textbf{7.7}$$

where c_k is the concentration of the k-th analyte. Theoretically, sensitivity should be the same for each concentration and each NAS vector.[21] The S/N ratio can be expressed as Equation 7.8:

$$S/N = \frac{c_k \left\| \mathbf{r}_k^{NAS} \right\|}{\left\| \boldsymbol{\varepsilon} \right\|} \qquad \textbf{7.8}$$

where $\boldsymbol{\varepsilon}$ is the random instrumental error.

To prove the relationship between NIR spectra and the SRB fuel components is not a product of correlated initial constituent concentrations, the mixture concentrations were calculated by the following orthogonalization procedure. A random matrix **A** corresponding to I mixtures and J components per mixture was constructed with a random number generator. Singular value decomposition of **A** according to Equation 7.9 yields orthogonal principal component scores, **U**, between -1 and +1, **S** is a diagonal matrix of singular values, and **V** is the matrix of eigenvectors (loadings).

$$\mathbf{A} = \mathbf{USV} \qquad \textbf{7.9}$$

A coefficient matrix, K, was constructed according to Equation 7.10 such that each row contains a zero and a one.

$$\mathbf{K} = \frac{\min[\mathbf{U}_i] + [\mathbf{U}_i]_{j=1}^{j}}{\max[\mathbf{U}_i]} \qquad \textbf{7.10}$$

A final matrix of concentrations is constructed by multiplication of **K** with the fuel ratios from the accepted SRB fuel recipe.

In search for the most descriptive multivariate model, the present research compared NAS regression with principal component regression (PCR),[9] interval PCR (iPCR), and PCR-uninformative variable elimination (PCR-UVE).[101-102] For each of these models, principal components were calculated by a singular value decomposition of the raw spectra according to Equation 7.9.

The regression of **U** in Equation 7.11 indicates which of the components have the strongest correlation to a change in constituent concentration **c**, where **a** is the y-intercept, **b** is a vector of regression coefficients, and **ε** is the residual.

$$\mathbf{c} = \mathbf{a} + \mathbf{b}\mathbf{U} + \boldsymbol{\varepsilon} \qquad \textbf{7.11}$$

Equation 7.12 demonstrates how a leave-one-out cross validation was used to predict the concentration of fuel components, where σ^2 is the variance, $f_i(\mathbf{U}_i)$ is the prediction of the model for the *i*-th pattern *m* in the training set, after it has been trained on the *m*-1 other patterns.

$$\sigma_{LOO}^2 = \frac{1}{m}\sum_{i=1}^{i=m}(\mathbf{c}_i - f_i(\mathbf{U}_i))^2 \qquad \textbf{7.12}$$

In the case of a simple two component system it is a simple matter to observe a linear change in the analyte concentration. In the presence of two system components, theoretically the

concentration change can be modeled by one principal component. The loading corresponding to this principal component accurately reflects the contribution of each wavelength to the overall classification. However, when multiple system components change simultaneously, multiple principal components are needed for the prediction model.

Interval PCR performs the same analysis as above, but rather than using the full spectrum, it uses smaller subsets of variables.[102] For example, when the experimenter specifies an interval (I_n) of 100 wavelengths, the algorithm performs PCR followed by principal component selection and cross-validation on intervals of 100. With a moving boxcar, all wavelengths are paired with all other wavelengths inside of $\pm I_n$. For example, after the algorithm analyzes 2101-2200nm, the next iteration analyzes 2102-2201nm, and so on. At the final wavelength, the first I_n wavelengths are added to the end for the final iterations. In this manner, each wavelength is included in a new model $2*I_n$ times. The goal of interval selection is the minimization of the standard error of performance in Equation 7.13, which indicates the interval with the highest correlation to the change in drug concentration:

$$SEP = \frac{\left(\frac{\boldsymbol{\varepsilon}^2}{N-1}\right)^{1/2}}{\max(\mathbf{c}) - \min(\mathbf{c})} \qquad \textbf{7.13}$$

where $\boldsymbol{\varepsilon}$ is the residual, N is the number of spectra, and $\mathbf{c}$ is a concentration vector.

Experimental

Fuel preparation. Thirty-three different mixtures of ammonium perchlorate (NH_4ClO_4), aluminum dust (Al), iron oxide (Fe_2O_3), casting resin (Bisphenol A/Epichlorohydrin), and

polyamide curing agent (Versamid 140) (Firefox Enterprises, Pocatello, ID) were constructed such that the percentage of each component was uncorrelated to the other components. Epoxy and curing agents were heated to 35^{o}C on a hot plate. Powders were mixed together and stirred to ensure a homogenous mixture. Epoxy and curing agent were added to the powder mixtures and stirred vigorously for three minutes. Fuel was poured into a ceramic mold so it could easily be inserted into the Pill Safe model. Chemical structures for iron oxide, ammonium perchlorate, bisphenol A, epichlorohydrin, polyamide resin, and Oxycodone can be found in Appendix B.

NIR Data Collection. NIR spectra were collected from 1100-2500 nm in 2 nm steps with a scanning spectrometer (Technicon InfraAlyzer 500, Tarrytown, NY) interfaced to a computer (OptiPlex GXM 5166, Dell, Round Rock, TX, USA) running SESAME 3.1 (Bran+Luebbe, Norderstedt, Germany). To eliminate room noise, samples were scanned inside of the instrument drawer. Each of the mixtures was scanned six times, all in random order to eliminate the effects of drift. All NIR data were exported to Matlab 7.0.1 (The Mathworks Company, Natick, MA) for processing and analysis.

Ignition Tests. To identify the burn characteristics, fuel samples were ignited quickly with a propane torch. The torch was adjusted so its flame extended two inches, and samples were placed at the very tip of the two-inch flame. The torch and flame were held in place by a series of clamps to ensure reproducibility. Ignitions were recorded on digital video and ignition start times and durations were noted.

Analysis of NIR spectra. Algorithms for net analyte signal (NAS), Hadi outlier detection, principal component regression (PCR), and interval principal component regression (iPCR) were all written by the authors for Matlab 7.0.1. NAS, PCR, and iPCR were all used to predict the concentrations of each fuel component from the collected NIR spectra. These methods allowed for the identification of the statistically significant regions of the NIR spectrum for the measurement of each component. Using the same methods, NIR spectra were then used to predict the ignition times and burn durations.

Extent of Incineration. The extent of OxyContin incineration was determined by high-pressure liquid chromatography (HPLC). Burn residue from the pill bottle loaded with OxyContin was ground with a mortar and pestle, washed with 200 mL of mobile phase, filtered two times through a 0.2 μm filter, and a chromatogram was collected to measure the remaining concentration of OxyContin. The HPLC assay was done with a Waters 717plus Autosampler, Waters 1525 Pump, and a Waters 2487 Dual Wavelength Absorbance Detector with Waters Breeze v3.30 Software. A Waters μBondapak C18 column (300 x 3.9 mm) was used with the UV/VIS Detector set at a wavelength of 206 nm for OxyContin standards. The mobile phase used for OxyContin was 0.005 M 1-hexanesulfonate, methanol, phosphoric acid, and triethylamine (900:100:5:2). The flow rate of the mobile phase was at 1.5 mL/min with 10μL injections of the sample. Standards analyzed exhibited excellent linearity over the concentration range employed. The retention time for OxyContin was 22.22 ± 0.211 min. The sensitivity of the assay was 25 ng/mL.

Results and Discussion

HPLC analysis indicated that following tablet incineration, the burn residue contained 5.58% by mass of the initial OxyContin. It must be noted that this 5.58% was recovered by grinding, washing, filtering, purifying, and chromatographic separation of the burn residue. With such an extensive extraction procedure and such a small yield, it is likely that would-be thieves would be sufficiently discouraged from a recovery attempt. The best fuel mixture required approximately ten seconds to consume the contents of the safe, making it nearly impossible to break into the bottle and remove the contents before the burn was complete. The Pill Safe is designed so that the flame is contained, and there is no danger of igniting external fires. The fuel formula contains an oxidizer; therefore, there is no need for atmospheric oxygen to sustain the reaction that destroys the tablets.

The most effective chemometric method for the measurement of SRB fuel components and prediction of ignition characteristics from the NIR spectra was the NAS regression. The fuel components had no significant correlation to each other; therefore, the predictive ability of the NIR was not a product of correlated concentrations. Additionally, the NAS found the portion of the signal unique to the analyte of interest, thus there were two measures of certainty that the strong correlation was not an artifact. The NIR measurement of NH_4ClO_4 resulted in an r^2=0.983, RMSEP=4.54%, the measurement of epoxy resulted in an r^2=0.997, RMSEP=2.31%, and the measurement of curing agent resulted in an r^2=0.996, RMSEP=4.27%. Pure component and second derivative NIR spectra from NH_4ClO_4, epoxy, and curing agent are shown in Figure 7.3 and Figure 7.4. It is apparent from these figures that the spectra were sufficiently distinct from each other, making it relatively easy to identify the NAS for each component. Iron oxide and aluminum had no distinguishing features in the NIR; therefore, they were not quantified in

this experiment. Of course, other fuel recipes can be used in place of the SRB fuel recipe. A mixture with five components, each with its own unique NIR chromophore, might make the most sense from a quality control standpoint. On the other hand, iron oxide can be determined in mixtures of aluminum dust using visible light spectrometry. The figures of merit calculated from the NAS vector for NH_4ClO_4 resulted in a selectivity=0.014, sensitivity=5.874, and S/N=39.941, for the casting agent, selectivity=0.005, sensitivity=6.467, and S/N=50.76, and for the curing agent, selectivity=0.022, sensitivity=4.097, and S/N=35.32. Selectivity is a unitless quantity, and sensitivity is given in units of signal/concentration. The limit of detection (LOD) is defined as the concentration at which S/N=3.[19] The S/N ratios were calculated separately for each individual concentration, and the theoretical LOD was extrapolated from the linear plot of concentration versus S/N. The LOD was 0.93% for NH_4ClO_4, 1.37% for casting agent, and 1.36% for curing agent. Precision was measured by calculating the relative standard deviation (RSD) of the predicted concentrations from the NAS vector. The RSD for NH_4ClO_4 was 0.38%, casting agent was 0.32%, and curing agent was 0.70%. There was no net bias for any of these measurements.

Figure 7.5 is a plot of principal component ellipses calculated from the fuel mixtures, demonstrating how well NIR was able to separate them. Ellipses are drawn six standard deviations from their cluster centers. NH_4ClO_4 demonstrated a high correlation to both ignition time and burn duration, suggesting that it was largely responsible for the reaction rates. Principal component 1 (PC1) demonstrated a high correlation to the concentration of curing agent, PC2 to casting agent, and PC3 to NH_4ClO_4. Mixture 1 demonstrated the best burn characteristics for the purpose of this research (i.e. fastest ignition, longest burn). The cluster location from mixture 1

can quickly be compared to other fuel mixtures, allowing this figure to act as a map to identify those components in which a particular mixture might be deficient. This figure can also be used to predict the ignition time and burn duration of each mixture depending on its constituent concentrations.

NIR spectrometry is a very versatile, nondestructive, and rapid method of analysis. It has been applied to both liquids and solid propellants with no sample preparation required.[122-125] As is evident by the PC clusters in Figure 7.5, it is also highly reproducible. Even when samples are very close to each other in concentration, they still cluster in distinctly different regions of a PC plot. This effect can be quantified by the bootstrap error-adjusted single sample technique (BEST).[32] According to the BEST, multidimensional standard deviations (MSD) greater than three indicate that clusters belong to different populations. When MSDs are less than three, clusters are considered inseparable. The average MSD separation between fuel mixtures was 431.36, and the average MSD separation calculated between repeat scans of the same mixture is 1.69, demonstrating the high degree of reproducibility from scan to scan. NIR analysis is also very rapid technique. NIR spectrometers can now collect thousands of spectra per second with high resolution.[128] NIR can effectively function as an in-process assay for the quantification of fuel mixtures as they are cast into the Pill Safe model. From the two hundred NIR spectra collected for this experiment, there were no aberrant scans, demonstrating that NIR is also a very rugged and durable assay.

Conclusion

This research presents a novel pill bottle design intended to provide a second line of defense against the theft of prescription drugs. NIR spectrometry was able to rapidly and accurately quantify the fuel components in fuel mixtures, and to identify burn characteristics of the different mixtures. This research suggests that Pill Safe manufacturing and SRB fuel analysis would be neither challenging nor expensive, thus making this a valuable product for use in the pharmaceutical industry.

Chapter Seven Figures

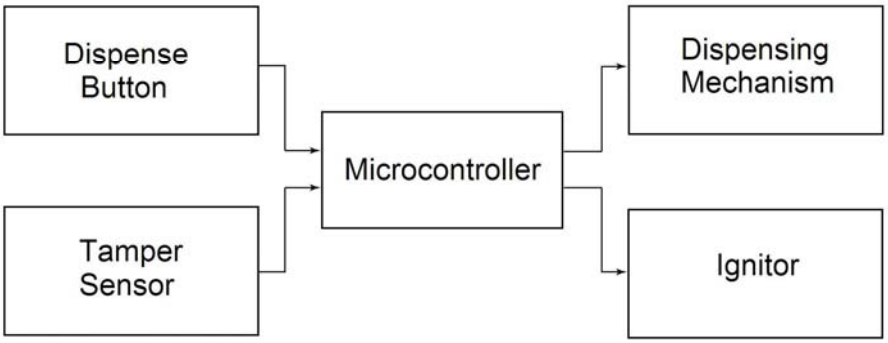

Figure 7.1: A simple schematic illustrating the main components of the Pill Safe. A microcontroller is responsible for the master coordinator. Based on the inputs and programmed instructions, the microcontroller dispenses the tablet as needed. If the sensor is perturbed, the microcontroller signals for neutralization, triggering the ignition of the fuel components.

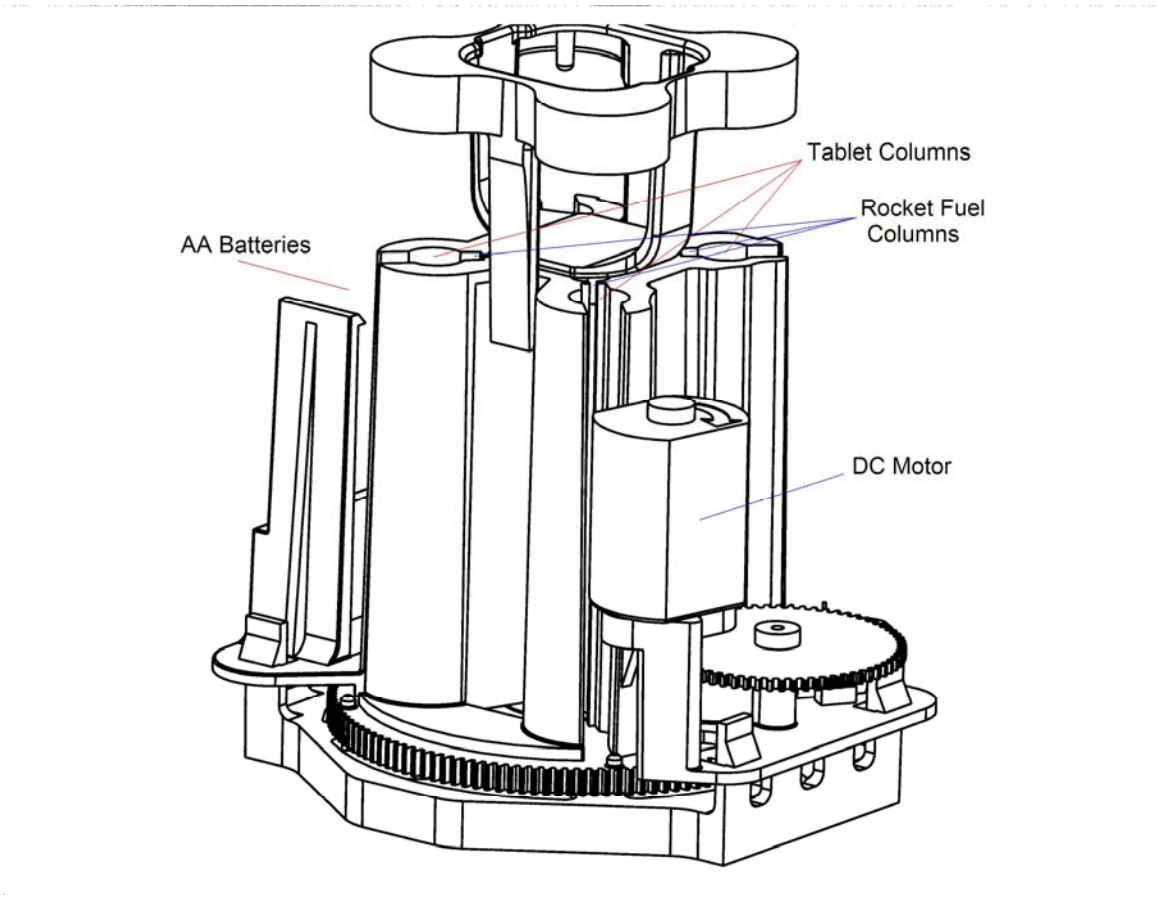

Figure 7.2: A cartoon depiction of the Pill Safe and the main components. Powered entirely by two AA batteries, the microcontroller dispenses a tablet once per dosing period. A conductive loop wraps around the assembly. Disruption of this loop causes the microcontroller to ignite the fuel rods and incinerate the tablets.

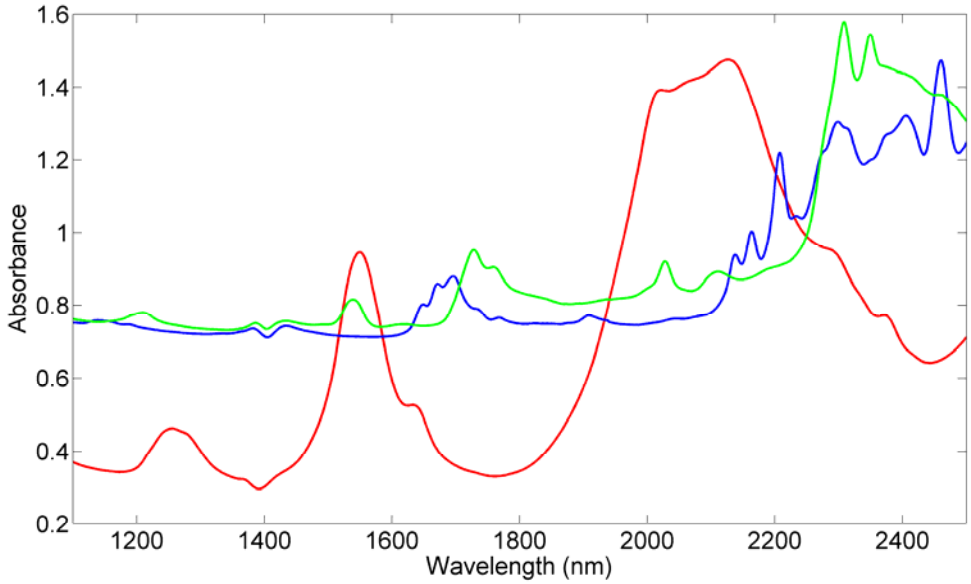

Figure 7.3: Pure component NIR spectra, epoxy resin is shown in blue, curing agent is shown in green, and NH_4CLO_4 is shown in red.

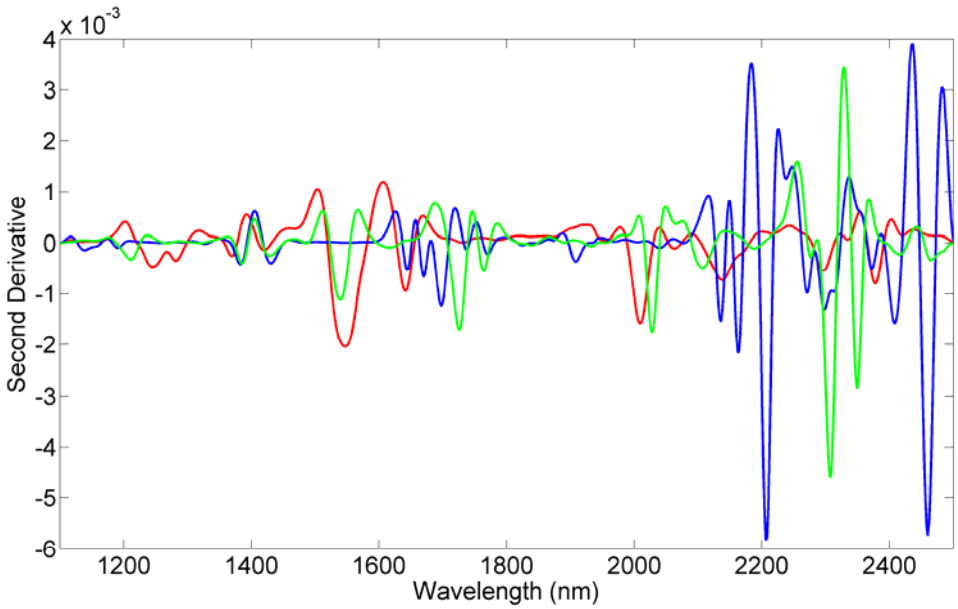

Figure 7.4: Second derivative NIR spectra, epoxy resin is shown in blue, curing agent is shown in green, and NH_4CLO_4 is shown in red.

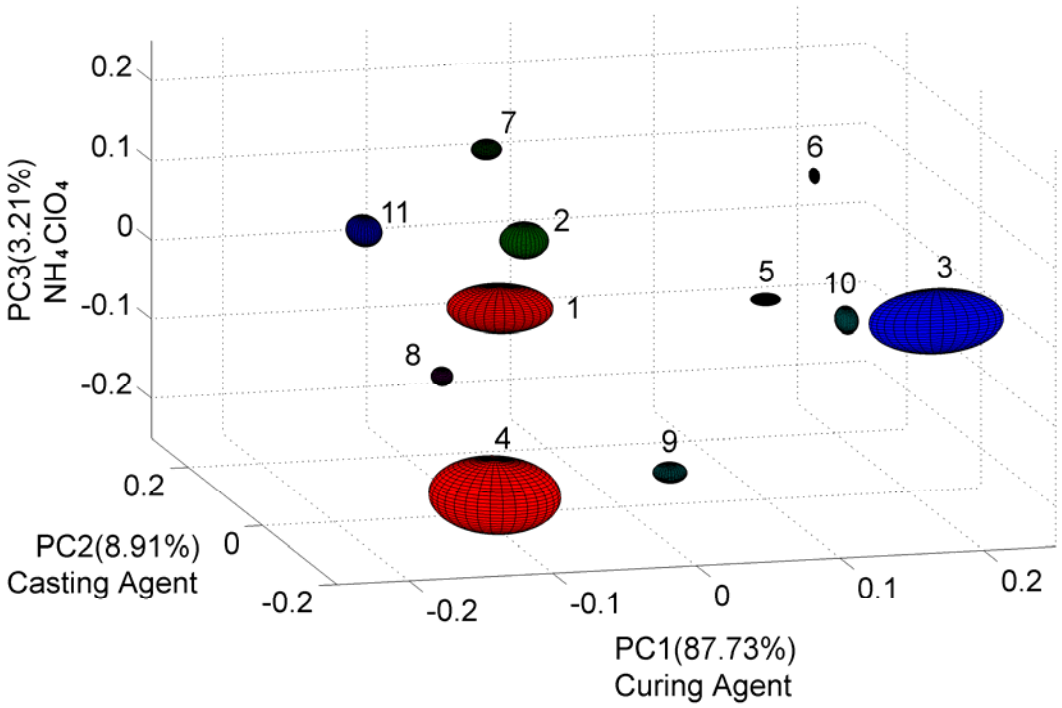

Figure 7.5: A plot of principal component scores indicates very rapidly if a mixture is deficient in a particular fuel component. PC score 1 corresponds very highly to curing agent, PC score 2 to epoxy, and PC score 3 to NH_4ClO_4. The ideal burn (fastest ignition and longest burn) was mixture number 1, therefore, projection of PC scores onto this plot can be used as a map for the prediction of SRB burn characteristics.

Copyright Statement

Joseph Medendorp, Jason A. Fackler, Tom Henninger, Bill Dieter, Robert A. Lodder. *Journal of Pharmaceutical Innovation.* Accepted February 2006.

Section IV: Sensors

Chapter Eight - Integrated Sensing and Processing Acoustic Resonance Spectrometry (ISP-ARS) for Sample Classification

Introduction

Integrated sensing and processing acoustic resonance spectrometry (ISP-ARS) can be used as an alternative to conventional Fourier transform ARS (FT-ARS) for the characterization of rocket fuel pre-mixes. ISP is a paradigm for the implementation of mathematics and chemometrics directly in the physical design of sensors for the purpose of data reduction.[1] Each new generation of computerized sensors and spectrometers contain more and more information obscured in data sets of higher and higher dimension. A typical data set from the FT-ARS used in this research contains *n* data points, where *n* is the product of sampling rate and duration. Even in low-frequency ultrasound applications, the resulting data vector contains a large number of variables (44.1KHz x 3s = 132,300 points per spectrum). Following data collection, a multivariate chemometric algorithm, such as classical/partial least squares, principal component regression, or multiple analysis of variance (MANOVA) must be applied to the data set to determine which frequencies correspond to the analyte of interest. For experiments requiring the collection of multiple spectra for multiple analytes, the computational burden of conventional FT-AR spectrometry is clear.

Conventional AR spectrometry has previously been used for the analysis of pharmaceutical tablets,[33,36] semi-solids and colloidal dispersions,[78,129] liquids,[41-42,75] and powders.[38-40] To the author's knowledge, this research presents the first application of ISP-ARS and comparison to full spectrum ARS. ISP-ARS is unique in that it does not rely exclusively on a scanning resonant waveform as with typical AR spectrometry, or the signal attenuation and time delay associated with active pulse acoustics. Instead, it relies on the excitation of key frequencies identified by MANOVA specifically for the maximization of inter-cluster distances and the

minimization of intra-cluster distances. The samples chosen for this research are different mixtures of ammonium perchlorate (NH_4ClO_4), aluminum dust (Al), iron oxide (Fe_2O_3), casting resin, and curing agent; the combination of which make up the fuel for the Pill Safe. Depending upon the concentrations of the individual components in fuel pre-mixes, different burn characteristics (i.e., burn duration, ignition time…) result, and there is a wide range of performance between mixes.[130] An analytical method capable of rapidly and accurately characterizing pre-mixes prior to burning can answer questions about burn performance, and optimize burns for different kinds and sizes of tablets in the Pill Safe.

The Pill Safe is a tamper-resistant pill bottle that uses fuel mixtures to destroy tablets when needed. Drug tablets are stacked next to fuel. Attempts to force the mechanism off the prescribed time or to penetrate the bottle cause instant destruction of the Pill Safe contents. The Pill Safe reduces the need for heightened security measures in pharmacies and drugstores due to theft of prescription drugs.

This study focused on the ISP waveform excitation. A detailed description of the basic AR spectrometer components has appeared elsewhere.[33,129] The ISP-ARS used in this research encoded the excitation signal (ISP waveform) so the detector automatically outputs the classification of the fuel pre-mix as a voltage. Calibration begins with acoustic excitation across a spectral range (0-22.05 KHz) using a broadband white noise source, as used with conventional FT-AR spectrometry. With the 22 KHz spectral range, the broadband excitation signal can be loaded into an inexpensive MP3 player, and the received signal can be collected with a computer sound card, thus sparing the expenses of a function generator and an A/D converter. As

illustrated in Figure 8.1, the AR spectrometer demonstrated two large resonant peak structures at 10-12 KHz and 14-17 KHz, making these useful regions to study for classification purposes.

In ARS, the resonance peak locations and heights shift in response to physical properties of samples such as pressure, density, viscosity, concentration, analyte size, and particle distribution. These properties are what the experimenter correlates to changes in instrumental response. When the goal of the experiment is quantification or classification, multivariate statistics can be used to determine a weight function by which the incoming spectrum can be projected onto for multivariate calibration. This projection can be accomplished in one of two ways; varying the excitation time at particular frequencies such that each frequency is weighted according to its importance in the regression (encoding the excitation signal), or by analog multiplication of the detected signal with an acoustic weight function prior to collection (post-collection).[83] As illustrated in Figure 8.2, conventional ARS consists of six steps. (1) A white noise signal source is used to excite all frequencies with equal amplitudes simultaneously. (2) The transmitting piezoelectric transducer (PZT) sends the signal through the quartz rod which is in mechanical contact with the fuel sample, which creates a standing wave resonant system. (3) The frequency spectrum from each sample is collected with a computer sound card or an analog/digital converter. (4) Using multivariate statistics, the weight function which accentuates the relationship between the most important frequencies or spectral features to the classification or analyte concentration can be calculated. (5) This weight function is either used as the excitation source (ISP-ARS), or is multiplied by the frequency spectrum, giving the (6) sample classification or quantification. Alternatively, ISP-ARS is effectively reduced to a two-step procedure: (1) excitation with the ISP waveform, and (2) data collection/analyte classification.

The data collection step is coupled with analyte classification by projection onto a three dimensional plot of the selected frequencies.

ISP-ARS offers the advantages over conventional AR spectrometry of reduced data analysis times and the reduction of computational burden with the instant output of analyte identity. In place of the white noise excitation signal, the weight function is used as the excitation source directly. Since the system resonant frequencies are a product of the physical components of the system, thus the weight function leads to analyte classification without the need for further computation. Therefore, ISP-ARS eliminates the entire digital component of the ARS.

Theory. Analytical methods are typically directed towards one of two things, analyte quantification, as in multivariate calibration, or analyte classification such as undirected data mining (i.e. cluster analysis). The intention of analyte classification, as is the focus of this research, is to determine the underlying structure of a data set. In addition, there are two types of data mining for analyte classification, supervised and unsupervised. In unsupervised data mining, no *a priori* information is given about the data structure as it relates to the classes of which it is comprised. Conversely, supervised data mining uses prior knowledge of classes to help identify the structure. For the purposes of ISP-ARS, class data is used to identify the key spectral features. MANOVA is used to identify which frequencies can be used to maximize between-cluster distances while minimizing same-cluster distances.[131] Once these frequencies are selected, they are encoded as the new excitation source, making possible the classification of incoming samples without the need for further computation.

Experimental Section

Mixture Preparation. Rocket fuel mixtures consisted of combinations of ammonium perchlorate (NH_4ClO_4), aluminum dust (Al), iron oxide (Fe_2O_3), casting resin (Bisphenol A/Epichlorohydrin), and polyamide curing agent (Versamid 140) (Firefox Enterprises, Pocatello, ID). The percentage of each component was chosen such that it was uncorrelated to the other components. The casting resin and curing agent were heated to 35^{o}C on a hot plate. Powders were hand blended prior to addition of casting and curing agents to ensure a homogenous mixture. The casting and curing agent were added to the powder mixtures and stirred vigorously for three minutes. Fuel mixtures were cast into mini plastic Petri dishes of equal size to ensure equal volumes between samples. Samples were dried for 48 hours prior to scanning. Chemical structures for iron oxide, ammonium perchlorate, bisphenol A, epichlorohydrin, polyamide resin, and Oxycodone can be found in Appendix B.

ARS Data Collection. Samples were placed on a scale (Model 3120, Health O Meter, Bridgeview, IL, USA) and kept in mechanical contact with the vertex of the quartz rod. Pressure was maintained at 200g between the rod and the sample for the duration of the scan. AR scans were first collected with a white noise excitation source. A battery powered MP3 player (Nike, Beaverton, OR) was used to supply the signal, thus keeping the 60Hz noise to a minimum. The signal was passed through a 741 dual supply operational amplifier prior to excitation of the PZT. The PZT detector output signal was collected via a 16-bit soundcard (Realtek AC97, Realtek Semiconductor Japan Corp., Yokohama, Kanagawa, Japan) for processing and analysis in Matlab 7.0.1 (The Mathworks Company, Natick, MA, USA). MANOVA was performed to identify the most significant frequencies, which were subsequently used for the construction of

the ISP excitation source. ISP-ARS data were collected with the weight function calculated with MANOVA. The signal consisted of three frequencies, one second of each, for a total of three seconds. The incoming voltage signal was summed for each frequency, and the total amplitude for each frequency was used as a single dimension in the classification procedure.

Results and Discussion

Three frequencies were retained for classification purposes. These three frequencies can be visualized as the three dimensions of an XYZ plot. As with all clustering algorithms, scans from similar samples tend to cluster together in hyperspace, while dissimilar samples tend to cluster in different regions of hyperspace. Therefore, projection of three frequencies onto a three dimensional plot quickly indicate to which class each sample belongs. Duplicate scans of the same analyte effectively draw probability contour plots, localized around the regions of highest probability. This approach is the electronic (analog) alternative to the digital calculation of probability densities for a given analyte.

ARS vs. ISP-ARS. Figure 8.3 illustrates the separation between clusters with conventional AR spectrometry. Principal components (PC) were calculated from the frequency range 14-17KHz, and the three PCs that captured the largest variation were plotted against each other. It is apparent from this figure that there was very little difficulty distinguishing between fuel samples. Illustrated in Figure 8.4 are the probability plots as calculated from the ISP-ARS data. In place of PC scores on the axes, the sum of the absolute value of each frequency selected from MANOVA makes up the XYZ coordinates. All clusters from both methods contain six points,

and the ellipses are drawn three standard deviations in all directions about the centroid of the clusters.

Comparison of Figures 8.3 and 8.4 indicate that there were notable differences between the ARS and ISP-ARS predictions. While both plots are drawn with three standard deviations about their centroid, it is clear that the clusters are smaller with ISP-ARS. This indicates that there is less variation and more reproducibility between scans than the full spectrum approach. This result is logical when one considers that the PC model was constructed with a window from 14-17KHz, whereas ISP-ARS uses only three frequencies. If the systematic error associated with each frequency is assumed to be the same, the inclusion of more variables has an additive effect on the total error included in the PC model. It must also be noted that the cluster sizes are all very similar to each other in Figure 8.3 (ARS), whereas there seems to be a larger variation in cluster sizes in Figure 8.4 (ISP-ARS). This is also a product of the sheer number of variables included in the PC model as compared with the ISP-ARS model. The calculation of PCs across a wide frequency range tends to suppress the overall effects of one erroneous frequency. In the ISP-ARS model however, one erroneous frequency can have an enormous effect on the overall projection point in hyperspace. These results suggest that ISP-ARS more effectively minimizes same-group cluster distances, however, is more sensitive to the presence of outliers than conventional ARS.

To illustrate the differences between ARS and ISP-ARS in a more quantitative fashion, the bootstrap error-adjusted single sample technique (BEST), a nonparametric cluster algorithm, was used to calculate multidimensional standard deviations (MSD) between clusters.[30] Reported in

Table 8.1 are the MSDs calculated from the ARS, using the region spanning 14-17KHz. Table 8.2 reports the MSDs calculated from the ISP-ARS clusters. According to this metric, standard deviations greater than three are considered separable. ARS resulted in an average BEST MSD of 1164.24, while ISP-ARS resulted in an average BEST MSD of 109.44. The average cross-validation MSD was 1.41 for the ARS and 1.58 for the ISP-ARS.

Versatility and Flexibility of ISP-ARS. ARS has previously been applied to the study of tablets,[33,36] semi-solids and colloidal dispersions,[78,129] liquids,[41-42,75] and powders.[38-40] Clearly, ARS can be used for the analysis of samples in nearly all phases. For any analyte that can be characterized or quantified by conventional ARS, an analogous ISP waveform can be programmed for the same samples. A database of ISP compatible waveforms can easily be built to accommodate any experimental need. For example, if an experiment requires a waveform to quantify the active pharmaceutical ingredient in a pharmaceutical tablet, they could simply download the ISP waveform from an online database to an MP3 player. Therefore, ISP-ARS is an extremely versatile method of analysis.

Speed of Method. The data collection process itself is very rapid for both ARS and ISP-ARS. No sample preparation is required, and all data files in this research were collected for three seconds. The data analysis for ARS, however, requires more time and more steps to process since it still requires the digital component of analysis. Therefore, the analysis times vary with the size of the data matrices for ARS. Alternatively, ISP-ARS automatically outputs the analyte classification without the need for further computation. So, only three seconds are necessary as data collection and analysis are performed simultaneously.

Figures of Merit. In multivariate data, where the concentrations of multiple species are changing simultaneously, calculation of standard figures of merit (FOM), such as dynamic range, limit of detection, selectivity, sensitivity, and signal-to-noise ratio is not a straightforward task.[83,130] The portion of the signal directly related to the analyte of interest is convoluted by the signals from all other species and interferences present in the signal. The only way to assess these quantities directly is by calculation of the net analyte signal (NAS), or the portion of the signal directly related to the analyte of interest, orthogonal to all interfering species.[19-22] Once the NAS is known, the relationship between system components and the analytical signal can be measured directly. However, calculation of the NAS requires pure component spectra from each of the analytes present in the system. For the purposes of this research, this is not a reasonable approach. For example, both the casting resin and the curing agent are liquids prior to addition of the other. As soon as they are mixed, they set and solidify almost immediately. Clearly, the acoustic signal of the individual liquids does not combine in a linear fashion to form the acoustic signal of the combined solid.

Calculations of accuracy, precision, and bias are typically associated with calibration experiments, where the analytical signal varies with a change in concentration in some quantifiable fashion. This research however, was for classification purposes, so calculation of these values was again not as straight forward. The following is a description of which FOM could be calculated and how each calculation was performed.

Detection limits. The detection limits were estimated by a mathematical translation of group populations $\mathbf{P_1}$ and $\mathbf{P_2}$ in hyperspace until they were no longer separable by the BEST metric.[33] For full spectrum ARS, the mean spectrum ($\overline{\mathbf{P}}$) between 14-17KHz is calculated for each population according to Equations 8.1 and 8.2.

$$\overline{\mathbf{P_1}} = \frac{1}{m}\sum_{i=1}^{i=m} \mathbf{P_{1i}} \qquad \mathbf{8.1}$$

$$\overline{\mathbf{P_2}} = \frac{1}{m}\sum_{i=1}^{i=m} \mathbf{P_{2i}} \qquad \mathbf{8.2}$$

A difference spectrum $\mathbf{X}$ is calculated by $\mathbf{X} = \overline{\mathbf{P_2}} - \overline{\mathbf{P_1}}$. The two populations are mathematically translated toward each other with $\mathbf{P_{ADJ}} = \mathbf{yX} + \overline{\mathbf{P}}_2$, where $\mathbf{y}$ is a coefficient matrix defined on the interval {0<y<1} until $\mathbf{P_{ADJ}}$ and $\mathbf{P_1}$ are inseparable by the BEST metric. To estimate the detection limits for each component, when the BEST MSD reaches three, the final value of $\mathbf{y}$ is multiplied by the difference in component concentrations. Mixtures 1 and 2 were very close in BEST MSD for both full spectrum ARS and ISP-ARS, and they also had the most similar chemical composition of all the mixtures, therefore, these samples were used for the calculation. The limits of detection were as follows (reported [ARS %, ISP-ARS %]): aluminum = [0.208, 0.176], NH_4ClO_4 = [0.161, 0.136], Fe_2O_3 = [0.004, 0.004], casting agent = [0.019, 0.016], and curing agent = [0.071, 0.060]. All values are reported in percent of the total mixture. ARS populations could be translated 93.5% of the way toward each other in hyperspace before $\mathbf{P_1}$ and $\mathbf{P_2}$ were inseparable, and ISP-ARS populations were translated 94.5% of the way before $\mathbf{P_1}$ and $\mathbf{P_2}$ were inseparable. Therefore ISP-ARS detection limits were slightly better than ARS for the measurement of each constituent.

Dynamic range. The lower end of the dynamic range is intimately related to the detection limit as described above. However, because the purpose of this research was not multivariate calibration, no upper-end limit could be explicitly calculated. The largest concentration differences were between mixtures 7 and 8. Mixture differences are reported in % difference of mixture 7 relative to 8 (aluminum = [62.3%], NH_4ClO_4 = [153.1%], Fe_2O_3 = [39.8%], casting agent = [73.2%], and curing agent = [73.3%]). In calculation of the BEST MSDs for ISP-ARS, mixtures 7 and 8 had the largest separation. However, for full spectrum ARS, the MSD separation between these two mixtures was the second highest (please refer to Tables 8.1 and 8.2). These results suggest that in ARS, when mixtures are exceedingly different, MSD separation is not a linear indicator of the degree of separation, and as such is not linearly related to changing concentrations. This conclusion is further supported by the ellipse plot in Figure 8.3. While mixtures 7 and 8 are clearly very far apart, they are not the most distant clusters in the representation. These results indicate that for estimation of dynamic range, BEST MSDs are useful for the low-end calculation. But for the high-end of the dynamic range, it is not necessarily the most accurate representation of the true chemical differences between AR spectra. Conversely, in ISP-ARS the BEST MSD between mixtures 7 and 8 is the largest degree of separation of all the mixtures, which agrees with the known chemical constituent differences, suggesting that BEST MSDs calculated from ISP-ARS clusters continue to vary linearly with changing concentration. In future studies, the ISP-ARS will be used for calibration in addition to sample differentiation; therefore, it will be a straightforward task to verify this claim.

Signal-to-noise, Precision, and Accuracy. The signal-to-noise ratio (S/N) was calculated as the mean amplitude of the signal divided by the standard deviation of repeated measurements. The

S/N ratios were [ARS = 24.87, ISP-ARS = 28.87]. Both S/N were of the same order, therefore, no appreciable differences were noted between the two methods. The precision is a measure of how repeatable each scan was for each mixture, expressed as the spread of repeated measurements relative to their total magnitude, or the inverse of the S/N ratio. The precisions were [ARS = 0.040, ISP-ARS = 0.035], thus the precision was slightly better for ISP-ARS. The measure of accuracy requires *a priori* knowledge or an expectation value, as in a calibration experiment. There was no known value to correlate the signal to; therefore, accuracy was meaningless in this context.

Group assignment. The set of fuel mixtures was scanned six more times each with the ISP-ARS as an external validation set. Projection onto the 3-dimensional plot of Figure 8.4 allowed for rapid group assignment. Each incoming spectrum was correctly assigned to its corresponding group one hundred percent of the time.

Freedom from Interferences. The main sources of potential interference in both ARS and ISP-ARS are (1) radio frequency (RF) cross-talk between transmitting and receiving PZTs, (2) sound waves propagating through the support structures of the instrument, and (3) erroneous mechanical and electrical noise bursts. (1) RF cross-talk was addressed by suspending two PZT transducers in mid-air 10cm apart from each other with no mechanical contact. Using one PZT in transmission mode and the other in collection mode, the non-contact mode signal strength was assessed. It was determined that the non-contact mode signal was negligible compared to the contact mode signal. Additionally, in its operational configuration, both the transmitting and receiving PZTs are fastened with epoxy to the quartz rod, therefore, they do not move unless

excited by mechanical contact. (2) The support structures for the ARS were made from wood, thus inhibiting aberrant waveforms from propagating through the beams. Wood is made up of a cellular network of pores that convert sound energy into heat by frictional and viscoelastic resistance.[68] The cellular pore network creates high internal friction, therefore, wood has more sound dampening capacity than most structural materials (e.g., steel, aluminum, or glass), thus, less sound traveled through the beams than if they had been made of metal. By coupling the quartz rod to the support structure through rubber grommets, an additional measure of dampening kept the sound interference to a minimum. Finally, the support structure was fixed in place and did not change during the course of the experiment, so interferences propagating through were the same for all sample and reference measurements. (3) All waveforms were programmed and generated synthetically in Matlab, therefore, they were tightly regulated for unwanted characteristics, such as spontaneous noise bursts, so the sound files for the broadband white noise source and the ISP waveform were not subject to reproach. However, the excitation signal was passed through a dual-supply 741 operational amplifier to drive the transmitting PZT to full capacity. Electronic interference from electrical outlets, lights, computers, or radios all emit unwanted radiation that was picked up and amplified by the amplifier. To address this issue, all electronic components associated with this research were battery powered. All scans were collected with the lights out and power sources unplugged from the outlets. In this manner, interferences were kept to a minimum.

Conclusion

Integrated sensing and processing acoustic resonance spectrometry has demonstrated the ability to perform favorably with conventional acoustic resonance spectrometry. With a specifically

tailored excitation signal, the ISP-ARS outputs the fuel mixture identity directly and requires no post-collection analysis, thus reducing data collection times and sparing the post-collection computational burden associated with ARS. An ISP waveform can easily be tailored for characterization or quantification of any type of analyte, therefore, this method is extremely attractive for further exploration.

Chapter Eight Figures

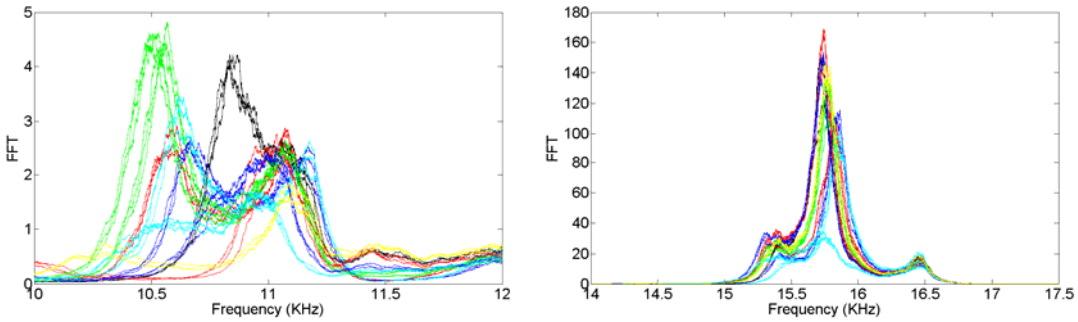

Figure 8.1: In its current configuration, the ARS has two large resonance peaks (10-12KHz and 14-17KHz). The spectral features shift in frequency and in amplitude as a response to changes in physical properties of the analyte. It is this property that allows for the quantification and classification of analytes using this resonant system.

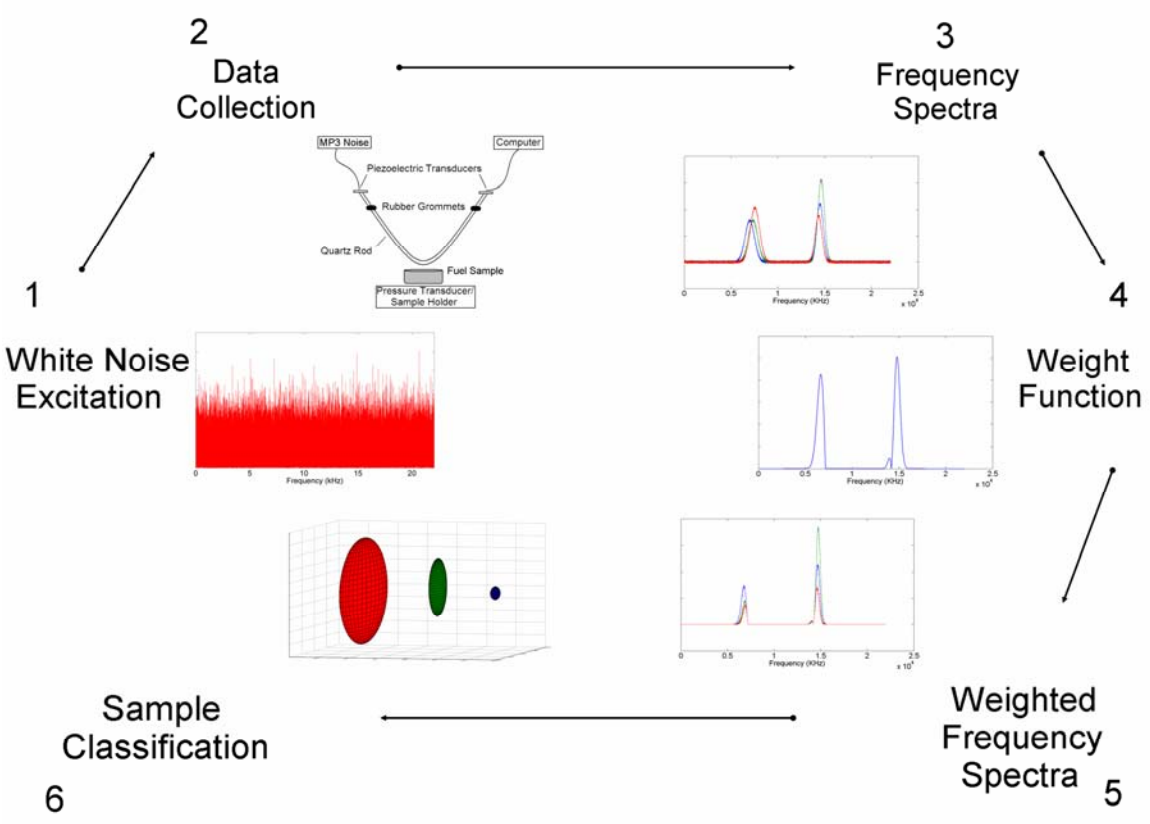

Figure 8.2: This figure illustrates conventional AR spectrometry with a broadband excitation source. When collecting full spectrum ARS, multivariate statistics are required to correlate the spectral features to the changing physical property of interest. Regression techniques, principal components, and multiple analysis of variance can be used to calculate a weight vector for use as the new ISP-ARS signal source.

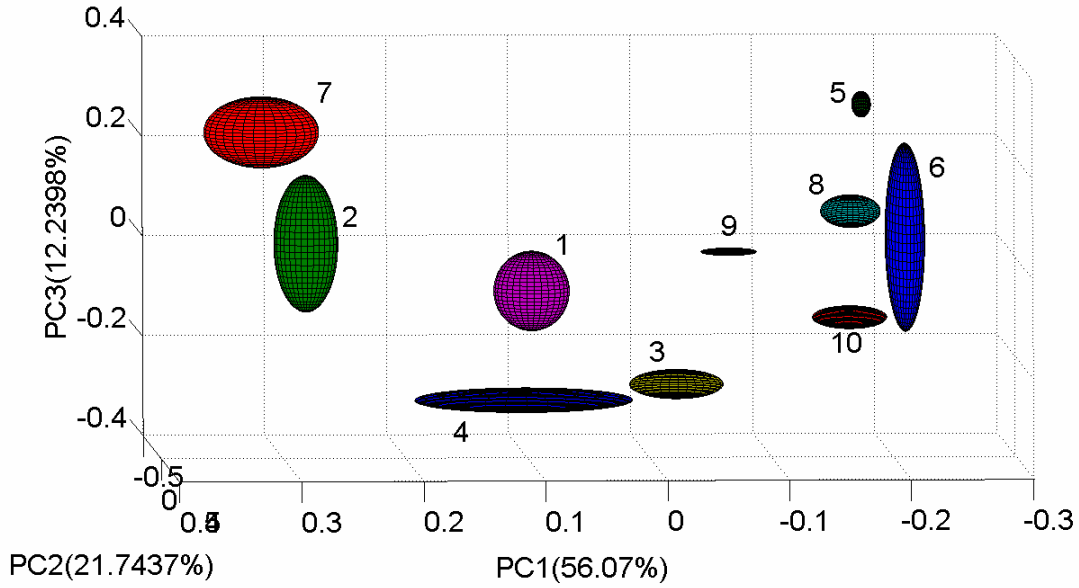

Figure 8.3: The three principal components which capture the largest amount of variance can be projected into three dimensional principal component space for easy visualization of the clusters. No *a priori* knowledge was included in the formation of this plot, therefore; the underlying data structure is clearly different for each fuel pre-mix.

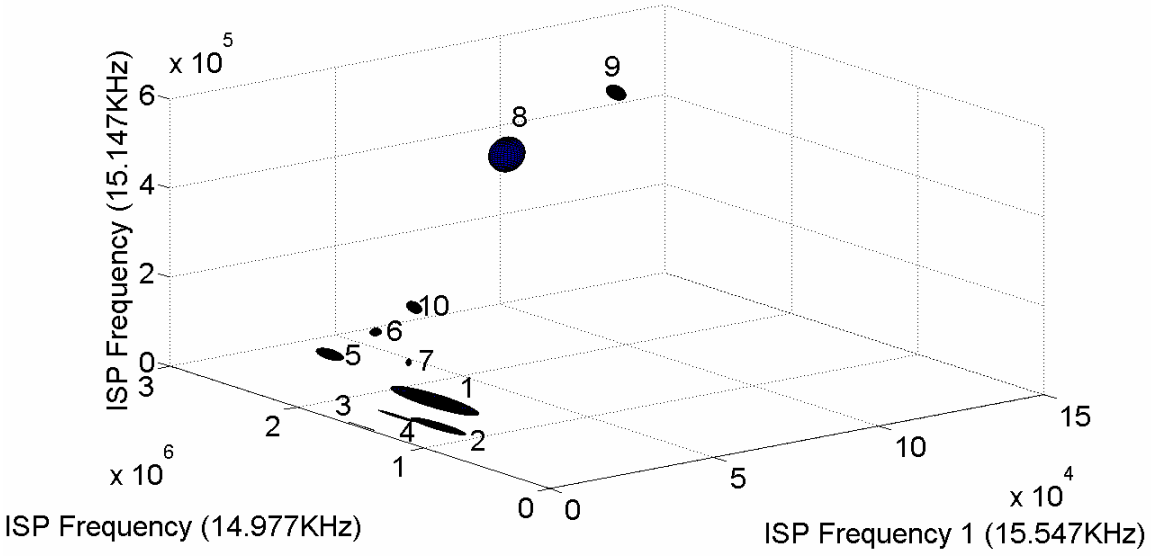

Figure 8.4: Full spectrum ARS spectra can be reduced to three dimensional data with the MANOVA selection of the most distinguishing frequencies, followed by excitation and data collection using those three frequencies. Contour plots drawn in three standard deviations indicate the separation of each mixture from all the others. Each ellipse contains six points, taken from repeat scans of the same analyte.

Chapter Eight Tables

Table 8.1: BEST distances calculated between mixture groups from 14-17 KHz. Same-cluster BEST distances are reported on the diagonal. Every one of the mixtures was greater than three standard deviations from the others, indicating that the data were separable by the BEST metric.

Mixture #	1	2	3	4	5	6	7	8	9	10
1	1.38	24.32	191.08	111.27	6429.50	28.96	224.42	713.40	12.85	80.22
2	-	1.47	146.72	289.11	9472.70	45.32	138.94	3433.40	94.00	112.53
3	-	-	1.33	172.79	186.34	150.29	185.14	185.17	205.29	200.36
4	-	-	-	1.58	507.43	18.63	92.03	51.25	321.79	57.69
5	-	-	-	-	1.31	980.15	783.20	11542.00	1367.20	530.10
6	-	-	-	-	-	1.40	76.34	145.72	100.52	32.22
7	-	-	-	-	-	-	1.43	7985.60	2045.00	51.18
8	-	-	-	-	-	-	-	1.46	1108.20	22.17
9	-	-	-	-	-	-	-	-	1.45	1738.20
10	-	-	-	-	-	-	-	-	-	1.29

Table 8.2: BEST distances calculated between mixture groups the ISP data. Same-cluster BEST distances are reported on the diagonal. Every one of the mixtures was greater than three standard deviations from the others, indicating that the data were separable by the BEST metric.

Mixture #	1	2	3	4	5	6	7	8	9	10
1	1.31	9.59	8.75	7.03	17.23	188.93	107.89	71.96	115.52	86.18
2	-	1.19	20.75	9.72	34.90	242.78	211.67	89.64	120.54	91.59
3	-	-	1.96	10.49	49.57	190.47	124.72	59.83	123.84	70.85
4	-	-	-	1.93	34.79	208.25	142.59	68.90	109.94	81.36
5	-	-	-	-	1.48	34.40	15.31	43.87	260.47	21.67
6	-	-	-	-	-	1.81	77.53	260.87	406.41	38.50
7	-	-	-	-	-	-	1.97	327.08	314.64	154.45
8	-	-	-	-	-	-	-	1.60	17.56	66.52
9	-	-	-	-	-	-	-	-	1.55	175.07
10	-	-	-	-	-	-	-	-	-	0.97

Copyright Statement

Joseph Medendorp. Jason. A. Fackler. Craig A. Douglas. Robert A. Lodder. *Analytical Chemistry*. Submitted May 2006.

Conclusion of Dissertation

The overall purpose of this dissertation is to illustrate the successful combination of chemometrics, spectrometry, and sensors, into one functional unit, termed 'integrated sensing and processing.' The utility of this fusion can be seen with the following three points: (1) optical spectrometry and acoustic spectrometry are preferred alternatives to slower and more invasive methods of analysis, (2) chemometrics can be implemented directly into the physical design of spectrometers, thus sparing the need for computationally demanding post-collection multivariate analyses, and (3) ISP sensors can be developed specifically for use in highly applied situations, making possible automatic analyte quantification or classification without the computational burden and extensive data analysis typically associated with conventional spectrometry. ISP and spectrometry comprise viable alternatives to many of the methods of analysis characteristic of those in pharmaceutical development today. For example, in place of complicated tissue extractions, chromatographic separations, and dissolution tests; NIR and ARS can reduce sample preparation, data collection, data analysis, and allow for high-throughput noninvasive testing. With the implementation of chemometrics directly into the design of spectrometers, ISP sensors are even more efficient than spectrometry alone. With these qualifications, ISP is the ideal solution for highly applied pharmaceutical challenges, such as process analytical technology. In a situation where an analytical instrument can be entirely dedicated to a single purpose, its functionality can be specifically tailored to suit that particular purpose. In this fashion, each step from inception to market; drug development, manufacturing, and testing can be monitored to minimize process variability. The goals of this research are therefore consistent with the advancement of process analytical technology, helping to ensure consistency from measurement

to measurement, batch to batch, and patient to patient, thereby reducing the risks associated with drug therapy.

Appendix A – Band Assignments

Irradiation with infrared light does not provide enough energy to excite an electron from its ground state configuration. It provides only enough energy to incite vibrational and rotational motion in the functional groups of the molecule. For a molecule with n-atoms, there are 3n-6 degrees of vibrational freedom (3n-5 for a linear molecule), allowing for motions such as stretching, bending, scissoring, rocking, and twisting. Each of these motions for a particular functional group has a characteristic frequency at which energy is absorbed. This property gives a characteristic absorbance response at each fundamental vibration mode in the mid-IR. The NIR region of the electromagnetic spectrum is a result of harmonic overtones and combinations from these mid-IR fundamental vibration modes. As such, it is more difficult to assign a response at particular frequencies to a specific functional group in the NIR. Often, a change in analyte concentration or identity can be seen across in multiple locations of the NIR spectrum, or in its entirety, rather than at specific bands. However, in order to approximate the region of the NIR spectrum where an analyst may find the appropriate analytical signal, a simple calculation can be performed. For example, a CH_2 asymmetric stretch frequency shows up as a sharp peak in the mid-IR around 2850cm^{-1}. To find the first harmonic, simply multiply the wavenumber by 2, and convert to nanometers. To find the second harmonic, multiply the wavenumber by 3 and convert to nanometers. With this calculation, an analyst would expect to see NIR absorbance peaks at 1169 and 1754 nm. Illustrated in Figure A1, these numbers are reasonably close approximations for the CH_2 asymmetric stretch frequency.

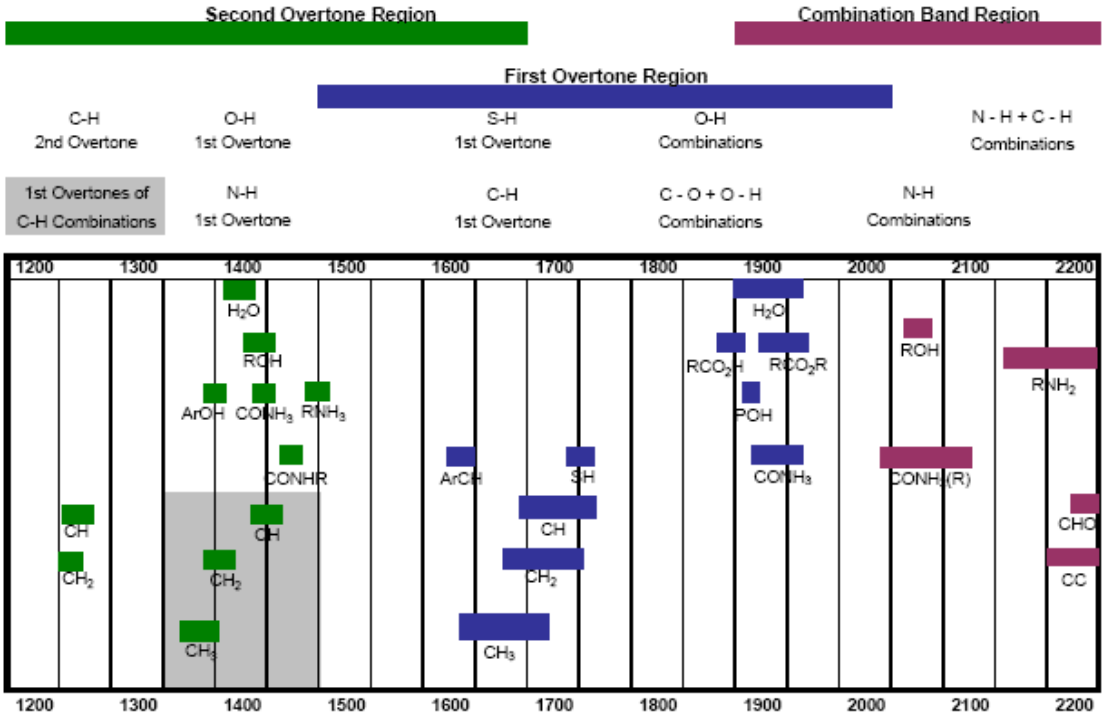

Figure A1: NIR functional group/spectrum correlation chart. This chart serves as an approximate guide for NIR band assignments by functional group.

Appendix B – Chemical Structures

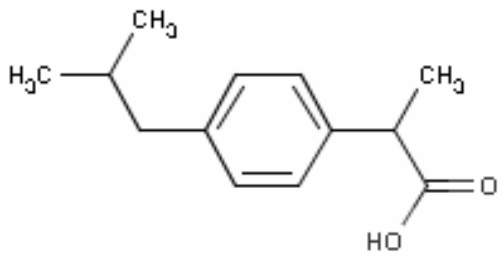

Figure B1: Ibuprofen.

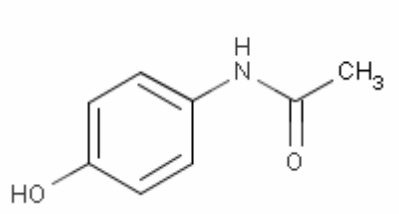

Figure B2: Acetaminophen.

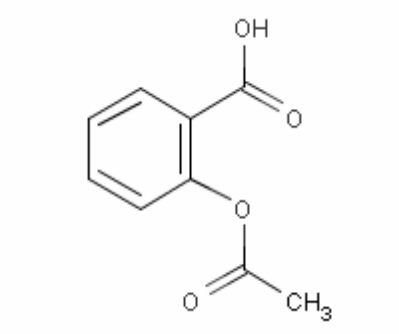

Figure B3: Aspirin.

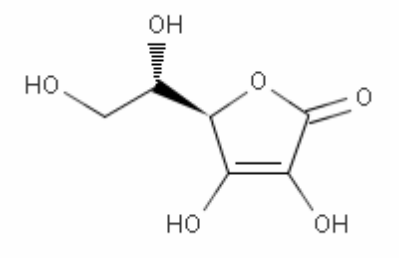

Figure B4: Ascorbic Acid.

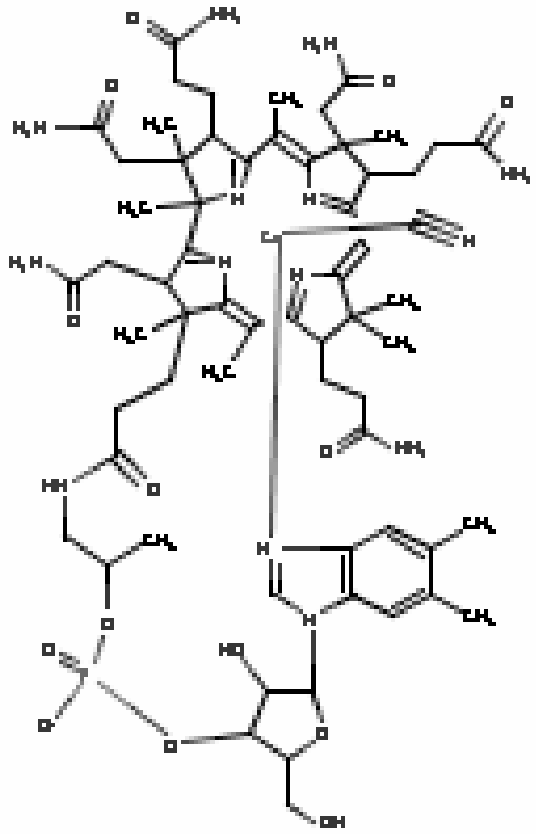

Figure B5: Vitamin B12.

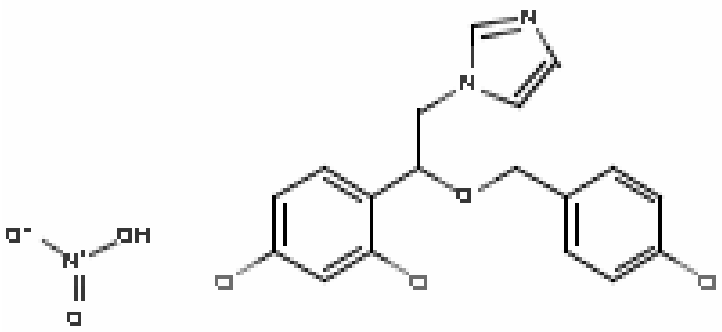

Figure B6: Econazole Nitrate.

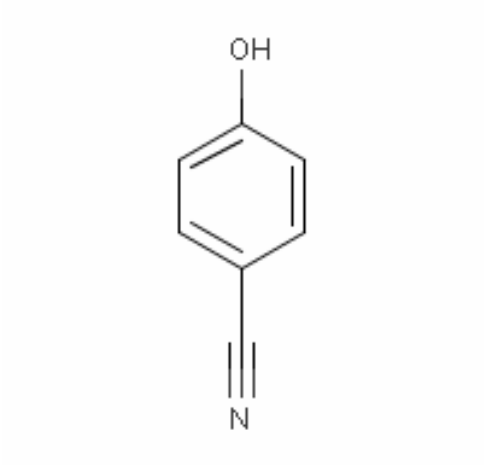

Figure B7: 4-Cyanophenol.

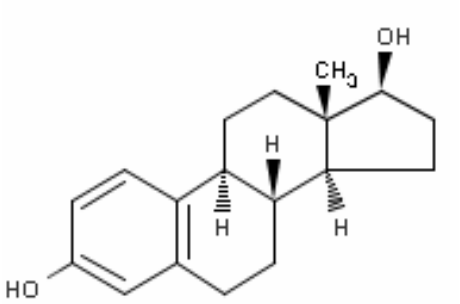

Figure B8: Estradiol.

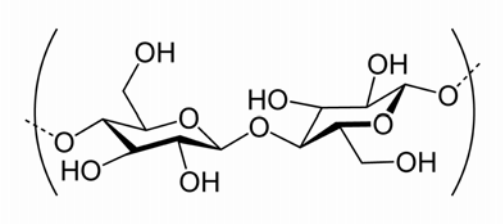

Figure B9: Microcrystalline Cellulose.

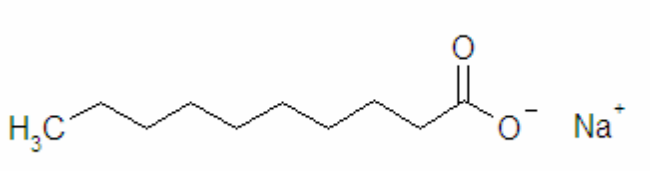

Figure B10: Sodium Caprate.

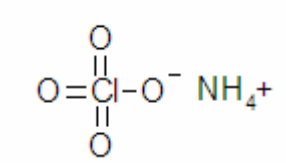

Figure B11: Ammonium Perchlorate.

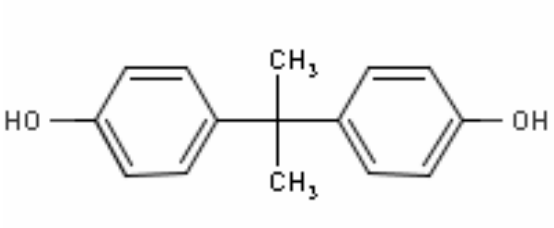

Figure B12: Bisphenol A.

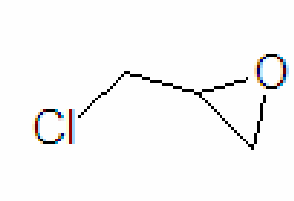

Figure B13: Epichlorohydrin.

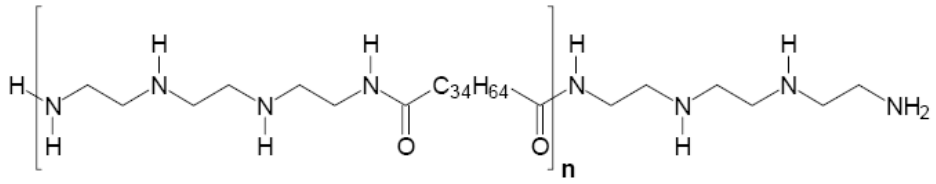

Figure B14: Polyamide Resin.

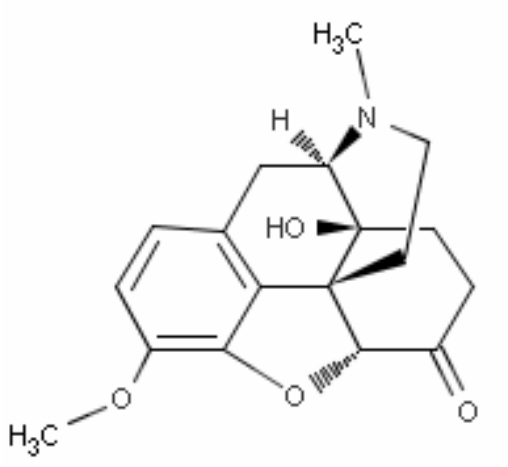

Figure B15: Oxycodone.

References

1. DARPA Defense Sciences Office. "Integrated Sensing and Processing." Available at: http://www.darpa.mil/dso/thrust/math/isp.htm/. Accessed May 21 2005.
2. Pantelic D. Optical Computation of Determinants. *Optics Communications*. **1987**; 64:421-424.
3. US Food and Drug Administration. Process Analytical Technology (PAT) Initiative. Available at: http://www.fda.gov/cder/OPS/PAT.htm/. Accessed April 1 2005.
4. Myrick M, Soyemi O, Haibach F, Zhang L, Greer A, Li H, Priore R, Schiza M, Farr J. Application of multivariate optical computing to near-infrared imaging. *Proceedings of SPIE*. **2002**; 4577:148-157.
5. Myrick M, Soyemi O, Schiza M, Farr J, Haibach F, Greer A, Li H, Priore R. Application of multivariate optical computing to simple near-infrared point measurements. *Proceedings of SPIE*. **2002**; 4574:208-215.
6. Haibach F, Myrick M. Precision in multivariate optical computing. *Applied Optics*. **2004**; 43(10):2130-2140.
7. Medendorp, J.; Lodder, RA. "Integrated Sensing and Processing and a Novel Acoustic-Resonance Spectrometer." American Association of Pharmaceutical Sciences, Baltimore, MD: 2004.
8. Medendorp, J.; Lodder, RA. "Integrated Sensing and Processing with a Tailored Excitation Signal for Acoustic-Resonance Spectrometry." The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, Orlando, FL: 2005.
9. Jolliffe, I. T. Principal Component Analysis. New York: Springer, 2002.

10. Belousov A, Verzakov S, von Frese J. Applicational aspects of support vector machines. *J. Chemometrics*. **2002**; 16:482-489.

11. Gestel T, Suykens J, Lanckriet G, Lambrechts A, De Moor B, Vandewalle J. Bayesian Framework for Least-Squares Support Vector Machine Classifiers, Gaussian Processes, and Kernel Fisher Discriminant Analysis. *Neural Computation.* **2002**; 14:1115-1147.

12. Thissen U, Ustun B, Melssen W, Buydens L. Multivariate Calibration with Least-Squares Support Vector Machines. *Anal. Chem*. **2004**; 76:3099-3015.

13. Chauchard F, Cogdill R, Roussel S, Roger J, Bellon-Maurel V. Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemometrics and Intelligent Laboratory Systems*. **2004**; 71:141-150.

14. Suykens J, Vandewalle J, De Moor B. Optimal control by least squares support vector machines. *Neural Networks.* **2001**; 14:23-35.

15. Muller K, Mika S, Ratsch G, Tsuda K, Scholkopf B. An Introduction to Kernel-Based Learning Algorithms. *IEEE Transactions on Neural Networks.* **2001**; 12(2):181-202.

16. Scholkopf B, Smola A, Muller K. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Computation*. **1998**; 10:1299-1319.

17. Wu W, Massart D, Jong S. Kernel-PCA algorithms for wide data Part II: Fast cross-validation and application in classification of NIR data. *Chemometrics and Intelligent Laboratory Systems*. **1997**; 37:271-280.

18. Bialkowski S. Species Discrimination and Quantitative Estimation Using Incoherent Linear Optical Signal Processing of Emission Signals. *Anal. Chem*. **1986**; 58:2563-2567.

19. Lorber A. Error Propagation and Figures of Merit for Quantification by Solving Matrix Equations. *Anal. Chem.***1986**; 58:1167-1172.

20. Booksh K, Kowalski B. Theory of Analytical Chemistry. *Anal. Chem.* **1994**; 66(15):782-791.

21. Lorber A. Net Analyte Signal Calculation in Multivariate Calibration. *Anal. Chem.* **1997**; 69(8):1620-1626.

22. Boelens H, Kok W, Noord O, Smilde A. Performance Optimization of Spectroscopic Process Analyzers. *Anal. Chem.* **2004**; 76(9):2656-2663.

23. Prakash A, Stellman C, Booksh K. Optical regression: a method for improving quantitative precision of multivariate prediction with single channel spectrometers. *Chemometrics and Intelligent Laboratory Systems*. **1999**; 46:265-274.

24. Cassis L, Urbas A, Lodder R. Hyperspectral integrated computational imaging. Analytical and Bioanalytical Chemistry. In Press.

25. Haibach F, Greer A, Schiza M, Priore R, Soyemi O, Myrick M. On-line reoptimization of filter designs for multivariate optical elements. *Applied Optics.* **2003**; 42(10):1833-1838.

26. Myrick M, Soyemi O, Li H, Zhang L. Spectral tolerance determination for multivariate optical element design. *Fresenius J Anal Chem*. **2001**; 369:351-355.

27. Nelson M, Aust J, Dobrowolski J, Verly P, Myrick M. Multivariate Optical Computation for Predictive Spectroscopy. *Anal. Chem*. **1998**; 70:73-82.

28. Soyemi O, Eastwood D, Zhang L, Li H, Karunamuni J, Gemperline P, Synowicki R, Myrick M. Design and Testing of a Multivariate Optical Element: The First Demonstration of Multivariate Optical Computing for Predictive Spectroscopy. *Anal. Chem.* **2001**; 73:1069-1079.

29. Weisstein E. "Multiplex Advantage -- from Eric Weisstein's World of Physics." Available at: http://scienceworld.wolfram.com/physics/MultiplexAdvantage.html/. Accessed May 1 2005.

30. Hamilton S, Lodder R. Hyperspectral Imaging Technology for Pharmaceutical Analysis. *Proceedings of Society of Photo-Optical Instrumentation Engineers*. **2002**; 4626:136-147.

31. Nahorniak M, Booksh K. Optimizing the implementation of the PARAFAC method for near-real time calibration of excitation-emission fluorescence analysis. *J Chemometrics.* **2003**; 17:608-617.

32. Lodder R, Hieftje G. Detection of subpopulations in near-infrared reflectance analysis. *App. Spec*. **1988**; 42(8):1500-1512.

33. Medendorp J, Lodder RA. Acoustic-Resonance Spectrometry as a Process Analytical Technology for Rapid and Accurate Tablet Identification. *AAPS PharmSciTech*. **2006**; 7(1): Article 25.

34. US Food and Drug Administration. Magno-Humphries, Inc. voluntarily recalls one lot of Dixon's®, Acetaminophen 325 mg Analgesic, 100 Tablet bottles. Available at: http://www.fda.gov/oc/po/firmrecalls/magno12_03.html/. Accessed April 1, 2005.

35. Lu Z, Hickey C, Sabatier J. Effects of Compaction on the Acoustic Velocity in Soils. Soil Science of America Journal. **2004**; 68:7-16.

36. Buice R, Pinkston P, Lodder R. Optimization of Acoustic-Resonance Spectrometry for Analysis of Intact Tablets and Prediction of Dissolution Rate. *App. Spec.* **1994**; 48(4):517-524.

37. Medendorp J, Lodder RA. "Acoustic Resonance Spectrometry." The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, Chicago, IL: 2004.

38. Serris E, Camby-Perier L, Thomas G, Desfontaines M, Fantozzi G. Acoustic emission of pharmaceutical powders during compaction. *Powder Technology.* **2002**; 128(2-3):296-299.

39. Reynaud P, Dubois J, Rouby D, Fantozzi G. Acoustic emission monitoring of uniaxial pressing of ceramic powders. *Ceramics International.* **1992**; 18(6):391-397.

40. Martin L, Poret J, Danon A, Rosen M. Effect of adsorbed water on the ultrasonic velocity in alumina powder compacts. *Materials Science and Engineering*. **1998**; 252:27-35.

41. Kaatze U, Wehrmann B, Pottel R. Acoustical absorption spectroscopy of liquids between 0.15 and 3000 MHz: I. High resolution ultrasonic resonator method. *J. Phys. E: Sci. Instrum*. **1987**; 20:1025-1030.

42. Bolotnikov M, Neruchev Y. Speed of Sound of Hexane + 1-Chlorohexane, Hexane + 1-Iodohexane, and 1-Chlorohexane + 1-Iodohexane at Saturation Condition. *J. Chem. Eng. Data.* **2003**; 48:411-415.

43. Shah V, Tsong Y, Sathe P, Liu J. In vitro dissolution profile comparison-statistics and analysis of the similarity factor f2. *Pharmaceutical Research.* **1998**; 15(6):889-896.

44. Shah V, Williams R. Roles of Dissolution Testing: Regulatory, Industry, and Academic Perspectives. Available at: http://www.dissolutiontech.com/DTresour/899Art/Roles_Pt1.html/. Accessed March 15 2005.

45. Drennen J, Lodder RA. Nondestructive near-infrared analysis of intact tablets for determination of degradation products. *Journal of Pharmaceutical Sciences.* **1990**; 79(7):622-627.

46. Zannikos P, Li W, Drennen J, Lodder R. Spectrophotometric prediction of the dissolution rate of carbamazepine tablets. *Pharmaceutical Research.* **1991**; 8(8):974-978.

47. Dousseau F, Pezolet M. On the spectral subtraction of water from the FTIR spectra of aqueous solutions of proteins. *App. Spec.* **1989**; 43(3):538-543.

48. Dousseau F, Pezolet M. Determination of the secondary structure content of proteins in aqueous solutions from their amide I and amide II infrared bands. *Biochem*. **1990**; 29:8771-8779.

49. Maddams W, Southon M. The measurement of derivative IR spectra. III. The effect of band width and band shape on resolution enhancement by derivative spectroscopy. *Spectrochim. Acta*. **1982**; 38A:459-466.

50. Fleissner G, Hage W, Hallbrucker A, Mayer E. Improved curve resolution of highly overlapping bands by comparison of fourth-derivative curves. *App. Spec*. **1996**; 50(10):1235-1245.

51. Cassis LA, Dai B, Urbas A, Lodder RA. In vivo applications of a molecular computing-based high-throughput NIR spectrometer. *Proc. SPIE Int. Soc. Opt. Eng.* **2004**; 239:5329-5344.

52. Kirmse D, Westerberg A. Resolution enhancement of chromatograph peaks. *Anal. Chem.* **1971**; 43:1035-1039.

53. Maldacker T, Davis J, Rogers L. Applications of Fourier transform techniques to steric-exclusion Chromatography. *Anal. Chem*. **1974**; 46:637-642.

54. Buice Jr. RG, Lodder RA. Determination of Cholesterol Using a Novel Magnetohydrodynamic Acoustic-Resonance Near-IR Spectrometer. *Appl. Spectrosc.* **1993**; 47:887-890.

55. Ultrasonic Scientific. High Resolution Ultrasonic Spectroscopy for Material Characterization. Available at: http://www.ultrasonic-scientific.com/. Accessed April 1, 2005.

56. Halliday D, Resnick R, Walker J. *Fundamentals of Physics*, 6th ed.; Wiley: New York, 2001.

57. Mills T, Jones A, Lodder R. Identification of Wood Species by Acoustic-Resonance Spectrometry Using Multivariate Subpopulation Analysis. *App. Spec.* **1993**; 47(11):1880-1886.

58. Ono Y, Jen CK, Cheng CC, Kobayashi M. Real time monitoring of injection molding for microfluidic devices using ultrasound. *Poly. Eng. Sci.* **2005**; 45(4):606-612.

59. Kowalska B. Ultrasonic measurements in polymeric processing. *Polimery.* **2005**; 50(9):646-650.

60. Moisan JF, Jen CK, Liaw JW, Zheng CQ, Chen TF, Sun Z, Loong CA. Ultrasonic sensor and technique for in-line monitoring of die casting process. Meas. Sci. Tech. **2001**; 12(11):1956-1963.

61. Kilian & Co Inc. Product showcase for pharmaceutical manufacturing, pharmaceutical process technology engineering. Available at: http://www.pharmaceuticalonline.com/Content/ProductShowcase/product.asp?DocID={E21A91E3-9BB6-11D4-8C69-009027DE0829}&VNETCOOKIE=NO/. Accessed November 25, 2005.

62. Lodder R. "CD/MP3 Acoustic Resonance Spectrometer." Available at: http://www.pharm.uky.edu/. Accessed April 5, 2004.

63. Alves M, Oliveira M. Interpolative biplots applied to principal component analysis and canonical correlation analysis. *J. Chemometrics*. **2003**; 17:594-602.

64. Lowell A, Kah-Siew Ho, Lodder RA. Remote Hyperspectral Imaging of Endolithic Biofilms Using a Robotic Probe. *Contact in Context.* **2002**; v1i1/planetprobe, 1(1):1-10.

65. Buice RG, Lodder RA. Detection of Subpopulations in Near-Infrared Reflectance Analysis. *App. Spec.*. **1993**; 47(7):887 – 890.

66. Lai E, Chan B, Chen S. Ultrasonic Resonance Spectroscopic Analysis of Microliters of Liquids. *App. Spec.* **1988**; 42(3):526-529.

67. Ritter T. "Random Noise Sources." Available at: http://www.ciphersbyritter.com/NOISE/NOISRC.HTM/. Accessed March 15, 2003.

68. "Western Red Cedar Physical Properties." Bear Creek Lumber. Available at: http://www.bearcreeklumber.com/generalinfo/onlineliterature/technicalinfohtml/wrcphysical properties.html/. Accessed March 1, 2004.

69. National Library of Medicine. Food and Drug Administration Actions: Recalls and Field Corrections 2002-May 12, 2004. Available at: http://householdproducts.nlm.nih.gov/NLM_FDARecalls.htm#rayblock1/. Accessed April 21, 2005.

70. Mariani E, Villa C, Neuhoff C, Dorato S. Derivatization Procedure and HPLC Determination of 2-Ethoxyethanol is Cosmetic Samples. *International Journal of Cosmetic Science.* **1999**; 21(3):199-205.

71. Westgate E, Sherma J. Determination of the sunscreen oxybenzone in lotions by reversed-phase HPTLC with ultraviolet absorption densitometry. *J. Liq. Chrom. & Rel. Technol.*, **2000**; 23(4):609-615.

72. Sabo M, Gross J, Rosenberg I. Quantitation of dimethicone in lotions using Fourier transform infrared spectral subtraction. *J. Soc. Cosmet. Chem.*, **1984**; 35:273-281.

73. Grunewald H, Kurowski C, Timm D, Grummisch U, Meyhack U. Rapid non-destructive raw material identification in the cosmetic industry with near-infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, **1998**; 6A:215-222.

74. Alltech Associates, Inc. Dimethicone. Available at: http://www.alltechweb.com/productinfo/technical/app/0048E.pdf/. Accessed April 15, 2005.

75. Leveque G, Ferrandis J, Van Est J, Cros B. An acoustic sensor for simultaneous density and viscosity measurements in liquids. *Review of Scientific Instruments*. **2000**; 71(3):1433-1440.

76. Ferrandis JY, Leveque G. In situ measurement of elastic properties of cement by an ultrasonic resonant sensor. *Cement and Concrete Research*. **2003**; 33: 1183-1187.

77. Dukhin AS, Goetz PJ. Acoustic Spectroscopy for Concentrated Polydisperse Colloids with high Density Contrast. *Langmuir*. **1996**; 12: 4987-4997.

78. Dukhin AS, Goetz PJ. Characterization of aggregation phenomena by means of acoustic and electroacoustic spectroscopy. *Colloids and Surfaces A: Phsicochemical and Engineering Aspects.* **1998**; 144: 49-58.

79. Ramdani A, Cros B, Sidki M, Ferrandis J. Acoustic near field technique for characterization of liquids, bitumen and cement setting. *European Physical Journal: Applied Physics*. **2001**; 15(1):69-76.

80. Patois R, Vairac P, Cretin B. Near-field acoustic densimeter and viscosimeter. *Review of Scientific Instruments*. **2000**; 71(10):3860-3863.

81. Prugne C, Van Est J, Cros B, Leveque G, Attal J. Measurement of the viscosity of liquids by near-field acoustics. *Measurement Science & Technology*. **1998**; 9(11), 1894-1898.

82. Medendorp J, Yedluri, J, Hammell DC, Ji T, Lodder RA, Stinchcomb AL. Near Infrared Spectrometry for the Quantification of Dermal Absorption of Econazole Nitrate and 4-Cyanophenol. *Pharm. Res*. **2006**; 23(4): 835-844.

83. Medendorp J, Lodder RA. Applications of integrated sensing and processing in spectroscopic imaging and sensing. *J Chemometrics*. **2005**; 19: 533-542.

84. Fountain W, Dumstorf K, Lowell AE, Lodder RA, Mumper RJ. Near-Infrared Spectroscopy for the Determination of Testosterone in Thin-Film Composites. *J Pharm Biomed Anal.* **2003**; 33(2):181-189.

85. Geladi P, MacDougall D, Martens H. Linearization and scatter-correction for NIR reflectance spectra of meat. *App. Spec*. **1985**; 39:491-500.

86. Shah VP, Hare D, Dighe SV, Williams RL. Bioequivalence of topical dermatological products. *Top. Drug Bioavailability, Bioequivalence, Penetration*. **1993**; 393-413.

87. Traulsen J. Bioavailability of betamethasone dipropionate when combined with calcipotriol. Int. J. Dermatol. **2004**; 43(8):611-617.

88. Pershing LK, Nelson JL, Corlett JL, Shrivastava SP, Hare DB, Shah VP. Assessment of dermatopharmacokinetic approach in the bioequivalence determination of topical tretinoin gel products. *J Am Acad Dermatol*. **2003**; 48(5):740-51.

89. Pirot F, Kalia YN, Stinchcomb AL, Keating G, Bung A, Guy RH. Characterization of the permeability barrier of human skin in vivo. *Proc. Natl. Acad. Sci. USA*. **1997**; 94:1562-1567.

90. Stinchcomb AL, Pirot F, Touraille GD, Bunge AL, Guy RH. Chemical Uptake into Human Stratum Corneum In Vivo from Volatile and Non-Volatile Solvents. *Pharm. Res*. **1999**; 16(8):1288-1293.

91. Reddy MB, Stinchcomb AL, Guy RH, Bunge AL. Determining Dermal Absorption Parameters in Vivo from Tape Strip Data. *Pharm. Res*. **2002**; 19(3):292-298.

92. Kreilgaard M, Kemme MB, Burggraaf J, Schoemaker RC, Cohen AF. Influence of a microemulsion vehicle on cutaneous bioequivalence of a lipophilic model drug assessed by microdialysis and pharmacodynamics. *Pharm. Res.* **2001**; 18(5): 593-599.

93. Xiao C, Moore DJ, Rerek ME, Flach CR, Mendelsohn R. Feasibility of tracking phospholipid permeation into skin using infrared and Raman microscopic imaging. *J Invest Dermatol*. **2005**; 124(3):622-32.

94. Cope M, Delpy DT. System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near-infrared transillumination. *Med. Biol. Eng. Comp*. **1988**; 26:289-294.

95. Chance B, Leigh JS, Miyake H, Smith DS, Nioka S, Greenfeld R, Finander M, Kaufmann K, Levy W, Young M, Cohen P, Yoshioka H, Boretsky R. Comparison of time-resolved and -unresolved measurements of deoxyhemoglobin in brain. *Proc. Natl. Acad. Sci. USA*. **1988**; 85:4971-4975.

96. Gandjbakhche AH, Weiss GH. Random walk and diffusion-like model of photon migration in turbid media. *Prog. Opt.* **1995**; XXXIV: 333-402.

97. Weiss GH. Some applications of persistent random walks and the telegrapher's equation. *Physica A*. **2002**; 311:381-410.

98. Urbas A, Manning MW, Daugherty A, Cassis LA, Lodder RA. Near-Infrared Spectrometry of Abdominal Aortic Aneurysm in the ApoE-/- Mouse. *Anal. Chem.* **2003**; 75:3650-3655.

99. Dempsey RJ, Davis DG, Buice RG, Lodder RA. Biological and medical applications of near-infrared spectroscopy. *App. Spec.* **1996**; 50:18A-34A.

100. Cassis LA, Lodder RA. Near-IR Imaging of Atheromas in Living Arterial Tissue. *Anal. Chem*. **1993**; 65:1247-1256.

101. Leardi R, Norgaard L. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectra regions. Journal of Chemometrics. **2004**; 18:486-497.

102. Noord OE. Elimination of Uninformative Variables for Multivariate Calibration. *Anal. Chem*. **1996**; 68:3851-3858.

103. Ferry L, Argentieri G, Lochner D. The comparative histology of porcine and guinea pig skin with respect to iontophoretic drug delivery. *Pharm. Acta Helv.* **1995**; 70:43-56.

104. Lodder RA, Selby M, Hieftje G. Detection of capsule tampering by near-infrared reflectance analysis. *Anal. Chem.* **1987**; 59:1921-1930.

105. Material Safety Data Sheet. "The MSDS Hyperglossary: Evaporation Rate." Available at: http://www.ilpi.com/msds/ref/index.html. Accessed July 7 2005.

106. Medendorp J, Yedluri J, Stinchcomb A, Lodder R. "A portable near-infrared sensor for noninvasively measuring drug concentrations in the skin after topical application." *The AAPS Journal*. **2005**; 7(S2).

107. Cassis LA, Urbas A, Lodder RA. Hyperspectral Integrated Computational Imaging. *Anal. Bioanal. Chem*. **2005**; 382:868-872.

108. Wierenga D, Eaton CR. Phases of Product Development – Drug Development and Approval Process. Alliance Pharmaceutical Corporation. Available at: http://www.allp.com/drug_dev.htm. Accessed March 23. Accessed January 1, 2006.

109. Dai B, Urbas A, Lodder RA. Sensor Batteries: Implantable Sensor Batteries. *NIR News*. **2006**; 17(1):14-15.

110. Introduction to NIR technology. Analytical Spectral Devices, Inc. Available at: http://www.asdi.com/ASD-600510_NIR-Introduction.pdf. Accessed February 12, 2006.

111. Moon KC, Wester RC, Maibach HI. Diseased skin models in the hairless guinea pig skin: *In vivo* percutaneous absorption. *Dermatologica.* **1999**; 180:8–12.

112. Panchagnula R, Stemmer K, Ritschel WA. Animal models for transdermal drug delivery. *Methods Find. Exp. Clin. Pharmacol.* **1997**; 19:335–341.

113. Jia X, Zhu Z, Li C. Immune and biochemical analysis of protein of homogenates of three different kinds of skin. *Zhonghua Zheng Xing Shao Shang Wai Ke Za Zhi*. **1996**; 12(1):51-3.

114. Lodder R, Hieftje G. Analysis of intact tablets by near-infrared reflectance spectrometry. *App. Spec.* **1988**; 42:556-558.

115. Vergote GJ, Vervaet C, Remon JP, Haemers T, Verpoort F. Near-infrared FT-Raman spectroscopy as a rapid analytical tool for the determination of diltiazem hydrochloride in tablets. *Eur. J. Pharm. Sci*. **2002**; 16(1,2):63-67.

116. Burns D, Ciurczak E. Handbook of Near-Infrared Analysis (Practical Spectroscopy) – 2nd edition. Marcel Dekker, 2001.

117. Drennen JK, Lodder RA. Qualitative Analysis by Near-Infrared Spectrometry: A Comparison of Discriminant Methods in Dissolution Testing. *Spectroscopy*. **1991**; 6(8):34-39.

118. National Drug Intelligence Center. Information Brief: Prescription Drug Abuse and Youth. Available at: http://www.usdoj.gov/ndic/pubs1/1765. Accessed December 11, 2005.

119. The New York Times. Pill Thefts Alter the Look of Rural Drugstore. Available at: http://query.nytimes.com/gst/health/article-page.html?res=9C03E1D6163BF935A35754C0A9629C8B63. Accessed August 17, 2005.

120. Domestic Strategic Intelligence Unit. Drug Intelligence Brief, OxyContin: Pharmaceuctical Diversion. Available at: http://www.avitarinc.com/pdf/Drug-Intelligence-Brief-Oxycotine-Facts.pdf. Accessed December 11, 2005.

121. Automatic Monitored Pill Dispenser. Available at: http://www.epill.com/md2.html. Accessed August 20, 2005.

122. Wang J. Shen G. Xing Z. Application of near infrared spectroscopy in liquid propellant analysis. *Huaxue Tongbao*. **2004**; 67(3):w023/1-w023/4.

123. Judge MD. The application of near-infrared spectroscopy for the quality control analysis of rocket propellant fuel pre-mixes. *Talanta*. **2004**; 62(4):675-679.

124. Rohe T. Gruenblatt E. Eisenreich N. Near infrared-transmission spectroscopy on propellants and explosives. International Annual Conference of ICT. 27th (Energetic Materials). **1996**; 85.1-85.10.

125. Weiser V. Eisenreich N. Fast emission spectroscopy for a better understanding of pyrotechnic combustion behavior. *Propellants, Explosives, Pyrotechnics.* **2005**; 30(1):67-78.

126. Harvey SD, Peters TJ, Wright BW. Safety considerations for sample analysis using a near-infrared (785 nm) Raman laser source. **2003**; *App Spec*. 57(5):580-587.

127. Hadi AS, Simonoff JS. Procedures for the Identification of Multiple Outliers in Linear Models. *J Am Stat Assoc*. **1993**; 88(424):1264-1272.

128. Klahn T, Billeb G, Weiser V. A high speed NIR-spectroscope to investigate ignition and combustion of propellants and pyrotechnics. *International Annual Conference of ICT*. **2004**; 35th, 161/1-161/9.

129. Medendorp JP, Buice Jr. RG, Lodder RA. *AAPS PharmSciTech.* Acoustic-resonance spectrometry as a process analytical technology for the quantification of active pharmaceutical ingredient in semi-solids. Accepted Apr. 2006.

130. Medendorp JP, Fackler JA, Henninger T, Dieter W, Lodder RA. NIR spectrometry for the characterization of fuel components in a novel tamper-resistant pill bottle. *J. Pharm. Innov*. Submitted Feb. 2006. Accepted Apr. 2006.

131. Krzanowski WJ. Principles of Multivariate Analysis. Oxford University Press, 1988.

Vita

JOSEPH MEDENDORP

Born: March 25, 1980
Racine, Wisconsin

EDUCATION

- **Cornell University**, Bachelor of Arts, May 2002

PROFESSIONAL

- Research Assistant, University of Kentucky College of Pharmacy, 2002-2006
- Teaching Assistant, University of Kentucky College of Pharmacy, 2002-2003

HONORS AND AWARDS

- 2005-2006 Research Fellowship, UK Center for Pharmaceutical Science and Technology
- 2005 Chair, American Association of Pharmaceutical Scientists, UK Chapter
- 2005-2006 North American Correspondent, International Council for Near Infrared Spectroscopy
- 2004 Vice Chair, American Association of Pharmaceutical Scientists, UK Chapter

PUBLICATIONS

1. Joseph Medendorp, Kalpana S. Paudel, Robert A. Lodder, Audra L. Stinchcomb. Near Infrared Spectrometry for the Quantification of Human Dermal Absorption of Econazole Nitrate and Estradiol. *Pharmaceutical Research.* Accepted August 2006.
2. Joseph Medendorp, Jason A. Fackler, Robert A. Lodder. Integrated Sensing and Processing Acoustic Resonance Spectrometry (ISP-ARS) for Sample Classification. *Analytical Chemistry*. Submitted May 2006.
3. Joseph Medendorp, Jason A. Fackler, Tom Henninger, Bill Dieter, Robert A. Lodder. NIR spectrometry for the characterization of fuel components in a novel tamper-resistant pill bottle. *Journal of Pharmaceutical Innovation*. Submitted Feb 2006. Accepted Apr 2006.
4. Joseph Medendorp, Robert Buice Jr., Robert A. Lodder. Acoustic-resonance spectrometry as a process analytical technology for the quantification of active pharmaceutical ingredient in semi-solids. *AAPS PharmSciTech.* **2006;** 7(3): Article 59.
5. Joseph Medendorp, Jhansi Yedluri, Dana Hammell, Tao Ji, Robert A. Lodder and Audra Stinchcomb. Near-Infrared Spectrometry for the Quantification of Dermal Absorption of Econazole Nitrate and 4-Cyanophenol. *Pharmaceutical Research.* **2006**; 23(4): 835-844.

6. Joseph Medendorp, Robert A. Lodder. Applications of Integrated Sensing and Processing in Spectroscopic Imaging and Sensing. *Journal of Chemometrics*. **2005**; 19: 533-542.
7. Joseph Medendorp, Robert A. Lodder. Acoustic-Resonance Spectrometry as a Process Analytical Technology for Rapid and Accurate Tablet Identification. *AAPS PharmSciTech.* **2006**; 7(1): Article 25.
8. Joseph Medendorp, Joseph Wyse, Robert A. Lodder, Lloyd G. Tillman, Sujatha Dokka, Michael Jay. Rapid near-infrared qualification of microcrystalline cellulose and sodium caprate minitablets through intact enteric coated capsules. *The Journal of Process Analytical Technology*. **2006**; 3(3): 11-17.

PRESENTATIONS

1. A portable near-infrared sensor for noninvasively measuring drug concentrations in the skin after topical application. *The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy.* March 2006: Orlando, FL.
2. A portable near-infrared sensor for noninvasively measuring drug concentrations in the skin after topical application. *American Association of Pharmaceutical Scientists.* November 2005: Nashville, TN.
3. Applications of Integrated Sensing and Processing (ISP) in Acoustic and Optical Spectroscopy. *Pharmaceutical Graduate Students Research Meeting.* June 2005: Lawrence, KS.
4. Integrated Sensing and Processing with a Tailored Excitation Signal for Acoustic-Resonance Spectrometry. *The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy.* March 2005: Orlando, FL.
5. Integrated Sensing and Processing and a Novel Acoustic-Resonance Spectrometer. *American Association of Pharmaceutical Scientists.* November 2004: Baltimore, MD.
6. Acoustic Resonance Spectrometry and Analysis of Powder Drying. *The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy.* March 2004: Chicago, IL.
7. Real-Time Determination of Hemoglobin Oxygenation State in Neonatal Mice. *The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy.* March 2004: Chicago, IL.
8. Near-Field NIR Spectra of HEK Membrane Bound Class A Scavenger Receptor. *The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy.* March 2004: Chicago, IL.