The Analyst

Critical Review

Near-infrared spectroscopy in the pharmaceutical industry

M. Blanco*, J. Coello, H. Iturriaga, S. Maspoch and C. de la Pezuela

Departamento de Química, Unidad de Química Analítica, Universidad Autónoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain

Summary of contents

Keywords: *Near-infrared spectroscopy; pharmaceutical preparations; pharmaceutical analysis; quality control; qualitative analysis; multivariate calibration; spectral pre-treatment; review*

Marcelo Blanco is Professor of Analytical Chemistry at Universitat Autonoma of Barcelona and Head of a group working on applied chemometrics. His major research interest is focused on molecular analytical spectrometry, including UV/VIS, FTIR, NIR, circular dichroism and molecular fluorescence. Topics under study in this group are the application of these techniques to the development of rapid methods for analytical control in several industrial fields (textiles, leather, electroplating, *etc*.) with special attention to pharmaceutical analysis by NIR applying multivariate calibration techniques. Other topics are the use of multivariate techniques to multicomponent resolution by kinetic measurements and chiral determination by capillary electrophoresis and circular dichroism.

Introduction

Although the near-infrared (NIR) was the first non-visible region discovered in the absorption spectrum (by Herschel in 1800)[1–3] analytical chemists made little use of it until the 1950s. A review published in 1960[4] that reported a comprehensive compilation of band assignments to different functional groups included only about 40 references to analytical applications of the NIR region; this prompted Wetzel's comment[5] that NIR spectroscopy was 'a sleeper among spectroscopic techniques'. However, analytical applications of the NIR technique have grown dramatically in number since the 1960s, so much so that a review of the topic was published in 1994 under the suggestive title 'Near-Infrared Spectroscopy. The Giant is Running Strong'.[6] Over the last 25 years, NIR spectroscopy has been increasingly used as an analytical tool, particularly by the food and agricultural industries, but also, to some extent, by the textile and polymer industries. Reported applications have been the subject of a number of reviews and books.[7–12] Applications to process control have also been developed over this period.[13] Although most of the ensuing analytical methods use some chemometric technique to correlate spectral data with physical or chemical properties of the samples, there are also recent uses of NIR for identifying impurities and elucidating structures from band assignments in addition to the earliest reported applications.[14–16]

The growing interest aroused by NIR spectroscopy in the industrial sector is probably a direct result of its two major advantages as an analytical tool for quality control. Thus, the low molar absorptivity of NIR bands permits operation in the reflectance mode and hence recording of spectra of solid samples with minimal or no pre-treatment, thereby substantially increasing the throughput. Also, the dual dependence of the analytical signal on the physical and chemical nature of the sample facilitates both its identification and the determination of physical and chemical parameters.

Notwithstanding these advantages, the pharmaceutical industry has been slow to adopt the NIR technique as it lacks the ability of mid-infrared (MIR) spectroscopy to identify samples by the mere inspection of spectra. In addition, quantitative NIR analyses involve calibration by sophisticated mathematical techniques that have reached extensive use only recently with the advent of microcomputing and chemometrics.

Despite the initial reluctance, NIR spectroscopy has aroused great interest in the last few years as a result of both instrumental breakthroughs (*e.g.*, improved detectors, the development of fast-scan and Fourier transform instruments to replace filter instruments, the widespread use of fibre-optic probes and instruments for recording spectra of individual tablets, which minimize or avoid sample pre-treatment) and the incorporation into equipment-bundled software of mathematical procedures for processing NIR spectra—a review of chemometric methods for NIR spectroscopy has been published by Mark.[17] In addition, the dependence of the NIR signal on

both the chemical composition and some physical properties of the sample, which was formerly considered a hindrance, permits not only the identification of compounds but also the total characterization of samples and the determination of non-chemical parameters with precision comparable to that of conventional methodologies, all due to powerful mathematical treatments for complex signals.

This paper is intended to provide readers with an overview of NIR uses by the pharmaceutical industry. To this end, the authors have divided references published up to 1996 into six different categories. The first lists specialized journals devoted to the NIR technique, some Internet addresses that can be used as literature sources and brief descriptions of other, previously published reviews. The second part is a brief introduction to the theoretical fundamentals of the technique and the third discusses the mathematical treatments used to process recorded signals prior to qualitative or quantitative analysis. The fourth and fifth parts include references to qualitative (identification of raw materials and end products, homogeneity studies, polymorphism and optical isomers) and quantitative applications (determination of physical parameters, water contents, active compounds and excipients); references are briefly commented on and practical aspects to be considered in addressing each type of analysis are discussed. The last part describes some applications that are not strictly pharmaceutical but might be of interest to the pharmaceutical industry.

Background and literature sources

Specialized journals

Although virtually every analytical chemistry journal has published NIR applications, the growing interest they have aroused and the widespread use they have reached in recent years have promoted the appearance of two specialized journals, *viz.*, *NIR News* and the *Journal of Near Infrared Spectroscopy*.

NIR News is the official newsletter of the International Committee for Near Infrared Spectroscopy. It contains up-to-date information about meetings, conferences, *etc.*, on NIR spectroscopy and various regular features that contain short articles aimed at disseminating general aspects of the technique, a list of published NIR papers and an advice page that answers general and specific questions posed by users of the technique.

The *Journal of Near Infrared Spectroscopy*, which started early in 1993, is a conventional-format publication consisting of papers, short communications and reviews of theoretical aspects of the technique and its uses in the industrial, agricultural, nutritional, polymer, textile and pharmaceutical fields, among others.

Internet addresses

The dramatic expansion of the World Wide Web (Internet) in the last few years has made a vast amount of information available in electronic format to the public. The acronym 'NIR' stands for 'Networked Information Retrieval' in Web jargon, so using it as a search string in any of the popular Web searching engines is bound to generate non-spectroscopic 'hits' (references to addresses containing information relevant to the target topic). Since late 1995, *NIR News* has included a special section devoted to Internet resources for NIR and related topics.[18] In a popularizing spirit, its issues list newsgroup and mail-list addresses [also called 'uniform resource locators' (URLs)] for bodies and societies concerned with NIR spectroscopy, and describe the procedure for joining them.

Tables 1 and 2 list newsgroups and mail-lists, respectively, with an interest in the use of NIR spectroscopy in various fields.[19,20]

Also worthy of mention here are electronic magazines (e-zines in Web jargon), which are partly or fully published in electronic format. Some are electronic versions of conventional journals while others have no 'paper parent.'. Thus, the magazine *e-JNIRS* is the electronic version of the *Journal of Near Infrared Spectroscopy*, in Adobe Acrobat format. Table 3 gives the URLs for some e-zines that publish papers on NIR spectroscopy and its uses. Some e-zines are published and managed by non-professional groups. Such is the case with *Wave of the Future*, which includes articles on the pharmaceutical uses of NIR. Comments, criticisms and suggestions are welcome, so readers act as true 'on-line reviewers'; readers' comments and authors' replies are all recorded and accessible to all. The e-zine can be reached at http://kerouac.pharm.uky.edu/ARSG/wave/wavehp.html.

At http://kerouac.pharm.uky.edu/ARSG/wave/cnirs/Ir_spec-htm, Kramer and Lodder have compiled URLs and resources related to MIR and NIR spectroscopy that include laboratories, departments, researchers' personal pages, instrument manufacturers, courses, journals, societies, *etc.* The same authors used *Wave of the Future* to publish a paper on MIR and NIR resources on the Internet, the printed version of which can be found in refs. 21 and 22.

Many NIR spectroscopic societies have their own addresses. Table 4 gives some of these. Finally, of special interest in relation to the pharmaceutical uses of NIR is Derksen's home page (http://leden.tref.nl/mderksen). This is an excellent repository of news and pharmaceutical uses of NIR that contains many links to papers about NIR and pharmaceuticals, as well as a list of NIR instruments of service to the pharmaceutical industry.

Table 1 Internet newsgroups concerned with the uses of NIR spectroscopy in various industrial fields

news://comp.ai.neural-nets	news://sci.data.formats
news://comp.soft-sys.matlab	news://sci.environment
news://comp.soft-sys.sas	news://sci.optics.fiber
news://comp.sys.mac.scitech	news://sci.polymers
news://misc.industry.quality	news://sci.stat.math
news://sci.agriculture	news://sci.techniques.misc
news://sci.bio.food-science	news://sci.techniques.spectroscopy
news://sci.chem.analytical	

Table 2 Mail lists concerned with NIR spectroscopy

List	Administrator's URL	Instruction
(Analysis Group) analysis-l	maiser@fs4.in.umist.ac.uk	analysis-l name
(American Society for Testing and Materials) astmsrch	listserv@uga.cc.uga.edu	astmsrch
(International Chemometrics Society) ics-l	listserv@umdd.umd.edu	ics-l name
(Process Group) process-l	maiser@fs4.in.umist.ac.uk	process-l
(Society for Applied Spectroscopy) applspec	listserv@uga.cc.uga.edu	applspec
(Spectroscopy Group of the UK's Institute of Physics) spectroscopy-group	mailbase@mailbase.ac.uk	spectroscopy-group
(Statistics and Statistical Discussion) stat-l	listserv@vm1.mcgill.ca	stat-l

Previous reviews

There are a multitude of reviews on NIR spectroscopy; few, however, deal exclusively with its pharmaceutical applications.

The *Handbook of Near-Infrared Analysis*[23] and the book *Advances in Near-Infrared Measurements*[24] each include one chapter devoted exclusively to NIR analyses of pharmaceuticals.

Below are briefly described literature reviews published in analytical chemistry or pharmaceutical journals containing more or less extensive sections on the uses of NIR by the pharmaceutical industry.

In 1986, Stark *et al.*[25] published a comprehensive review of the NIR technique as a tool for qualitative and quantitative analysis. The four parts of the work provide readers with information about the fundamentals of the technique, the equipment it uses, chemometric treatments for processing NIR signals and performing quantitative analyses, its advantages and a list of applications in a variety of industrial fields (pharmaceutical, textile, nutritional, biomedical, chemical).

Pharmaceutical applications of NIR spectroscopy were first reviewed by Ciurczak in 1987.[26] The review cited about 50 papers published up to 1987 but made no mention of the technique's fundamentals or associated equipment.

The 1991 review by Drennen *et al.*[27] encompassed several previous reviews published up to 1988 and included an introduction to NIR equipment and its recent developments; it also described the theoretical foundation of qualitative and quantitative chemometric methods used in connection with NIR applications.

In 1992, reviews by Martin[28] and by Ciurczak and Drennen[29] were published. Of special interest in Martin's extensive work is the section dealing with quantitative analysis, with subsections devoted to sample selection, mathematical signal treatments, calibration techniques and their selection, and calibration transfer. Also worthy of note is the applications section, with references to NIR uses in food, cosmetic, polymer, textile and pharmaceutical analyses. In contrast, the review of Ciurczak and Drennen only covers, and rather briefly, 31 references to qualitative and quantitative applications of NIR spectroscopy in the pharmaceutical industry.

In 1993, Corti *et al.*[30] reviewed NIR diffuse reflectance spectroscopic uses in the pharmaceutical and biomedical fields reported up to 1992. In addition to about 70 references to specific applications, the review included a brief description of the more widely used chemometric methods for qualitative and quantitative analysis by NIR spectroscopy.

The reviews by Workman[13] and by MacDonald and Prebble[31] were also published in 1993. The former included about 145 applications of NIR spectroscopy in various industrial fields, described in a chronological sequence and classified into different categories (in-line, *in situ*, on-line, non-invasive, remote and rapid NIR analyses). The latter review was defined by its authors as a 'non-comprehensive overview of near-infrared reflectance analysis in the pharmaceutical industry' and illustrates the technical potential of this technique in this industrial field.

Recent advances in NIR spectroscopic equipment were compiled by McClure in 1994.[6] He described commercially available NIR spectrophotometers from various manufacturers. Other reviews of NIR pharmaceutical uses were published by Morisseau and Rhodes,[32] and by Kirsch and Drennen,[33] both in 1995. The former authors reviewed far fewer references than did Corti *et al.*;[30] however, they included a short, albeit interesting, introduction to NIR equipment and its manufacturers. Kirsch and Drennen placed special emphasis on direct analyses of solid pharmaceutical preparations (intact dosage forms) and included a brief introduction to chemometric methods used for qualitative NIR analysis and about 30 references to qualitative and quantitative studies in the pharmaceutical field.

Fundamentals of the technique

Properly using an instrumental technique for a specific analytical purpose entails knowing what is to be measured and in what way, since extracting the full informative potential from an instrument requires a sound knowledge of the physicochemical theories on which its measurements rely and of the instrumental principles involved. In order to introduce the readers to the technique, this section is intended to provide a basic knowledge of the theoretical foundation of NIR spectroscopy and diffuse reflectance measurements in the NIR region.

Table 3 Uniform resource locators (URLs) for selected journals publishing articles about NIR spectroscopy

Journal	URL
Analyst	http://www.rsc.org/analyst
Analytica Chimica Acta	http://www.elsevier.com:80/locate/issn/03654877
Analytical Chemistry	http://pubs.acs.org/journals/ancham/index.html
Applied Spectroscopy	http://esther.la.asu.edu/sas
Applied Spectroscopy Newsletter	http://esther.la.asu.edu/sas/epstein
Chemometrics and Intelligent Laboratory Systems	http://www.elsevier.com:80/locate/issn/01697439
e-JNIRS	http://www.nirpublications.com/electron.html
Food Testing and Analysis	http://www.worldsys.com/labinfo/journal/fta/fta/htm
Journal of Chemometrics	http://www.wiley.com/journals/cem
Journal of Near Infrared Spectroscopy	http://www.nirpublications.com/jnirs.html
Spectroscopy	http://www.techexpo.com/toc/spectros.html
Journal of Pharmaceutical Sciences	http://pubs.acs.org/journals-sci/jsfa/index.html
NIR News	http://www.nirpublications.com/nirn.html
Talanta	http://www.elsevier.com:80/locate/issn/00399140
Trends in Analytical Chemistry	http://www.elsevier.com:80/section/chemical/trac

Table 4 Uniform resource locators (URLs) for selected societies concerned with NIR spectroscopy

Society	URL
American Association of Cereal Chemists	http://www.scisoc.org/aacc/info.html
Council for Near Infrared Spectroscopy	http://kerouac.pharm.uky.edu/ASRG/cnirs/cnirs.htm
Society for Applied Spectroscopy	http://esther.la.asu.edu/sas

Interested readers can find much more extensive descriptions of both NIR theory and equipment elsewhere.[7,34,35]

Principles of NIR spectroscopy

The NIR lies between the visible and MIR regions of the electromagnetic spectrum and is defined by the American Society for Testing and Materials (ASTM) as the spectral region spanning the wavelength range 780–2526 nm (or the wavenumber range 12 820–3959 cm^{-1}). Light absorption in this region is primarily due to overtones and combinations of fundamental vibration bands occurring in the MIR region. For infrared light to be absorbed, its energy must be high enough to produce vibrational transitions in the molecules concerned, *i.e.*, the light frequency should be exactly the same as a fundamental vibration frequency for a specific molecule and the molecule should undergo a change in its dipole moment by virtue of its fundamental vibration.

The vibrational frequency f for a diatomic molecule can be determined on the assumption of the harmonic oscillator model, where an atom shifts from its equilibrium position with a strength proportional to the shift (Hooke's law):

$$f = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}}$$

where c is the speed of light, k the bonding force constant (a measure of the strength or rigidity of a chemical bond in its normal equilibrium position) and μ the reduced mass.

In this case, the variation of the potential energy with bond distance is a parabola centred about the equilibrium distance with evenly spaced vibrational energy levels. The energy E_v of each level will be given by

$$E_v = f(v + \tfrac{1}{2})$$

where f is the vibrational frequency and v the vibrational quantum number.

Because the selection rule for harmonic oscillator transitions is $\Delta v = \pm 1$ and energy levels are evenly spaced, the energy difference between two consecutive levels will always be $E_{(v+1)} - E_v = f$, which is called the 'fundamental frequency' of the band.

Vibrations in polyatomic molecules involve complex movements of their constituent atoms. The movements can be resolved into individual vibrations called 'normal vibrations'. The motion of each atom is the result of its movements in the normal vibrations; the energy of each normal frequency is independent of the others, so the vibrational energy of the molecule is the sum of the individual energies:

$$E_v = \sum_{i=0}^{N} f(v_i + \tfrac{1}{2})$$

In practice, molecular vibrations tend to be non-harmonic, *i.e.*, vibrations about the equilibrium position are non-symmetric. The potential energy curve for real bonds is only roughly parabolic, with slight deviations at the lower energy levels that are more marked at the upper energy levels. Also, spacings between energy levels are not identical but rather decrease with increasing energy.

One correction to the harmonic oscillator model that improves consistency between theoretical and experimental data involves additional terms of higher order than those used by Hooke's law. Thus, the energy E_v for each level will be given by

$$E_v = f_e(v + \tfrac{1}{2}) - f_e x_e(v + \tfrac{1}{2}) + \text{higher-order terms}$$

where v is the vibrational quantum number, x_e the non-harmonicity constant (which measures the deviation of the potential function from the parabola) and f_e the uniform spacing between levels corresponding to a parabola with its centre at the equilibrium distance and the same curvature as the real potential energy function.

If higher-order terms are neglected, then the frequency of a transition between adjacent energy levels ($v \rightarrow v + 1$) will depend on the vibrational quantum number:

$$f = f_e[1 - 2x_e(v + 1)]$$

One further consequence of introducing the quadratic term into Hooke's law is that the selection rule becomes $\Delta v = \pm 1, \pm 2$, *etc.*; hence, in addition to the fundamental band (+1), other, higher frequencies called overtones or harmonics appear at frequencies two, three, *etc.*, times higher than the fundamental frequency. The intensity of these bands decays abruptly since transition probability decreases markedly with increase in the vibrational quantum number and, in practice, each fundamental band only exhibits its first two or three overtones.

Polyatomic molecules possess several fundamental frequencies so they may exhibit simultaneous changes in the energies of two or more vibrational modes; the frequency observed will be the sum of ($f_1 + f_2$, $2f_1 + f_2$, *etc.*) or the difference between ($f_1 - f_2$, $2f_1 - f_2$, *etc.*) the individual fundamental frequencies. This results in very weak bands that are called 'combination' and 'subtraction' bands—the latter are possible but rarely observed in room temperature NIR spectra. Non-harmonicity results in combination bands that are slightly smaller than the combined fundamental frequencies involved.

Many NIR bands are overtones and combination bands for hydrogen bonds (C–H, N–H, O–H and S–H). The small mass and large force constants for hydrogen are the origin of the high fundamental frequencies in this atom; as a result, its first few overtones appear in the NIR region. C=O, C–C, C–F and C–Cl groups usually exhibit very weak or no bands in the NIR region; fundamental vibrations in these groups occur at low frequencies in the MIR region, where their first few overtones also appear as a result.

NIR diffuse reflectance spectroscopy

The low molar absorptivity of adsorption bands in the NIR region (typically between 0.01 and 0.1 l mol^{-1} cm^{-1}) severely restricts sensitivity; however, it permits operation in the reflectance mode and hence the recording of spectra for solid samples.

Reflectance spectroscopy measures the light reflected by the sample surface, which contains a specular component and a diffuse component. Specular reflectance, described by Fresnel's law, contains little information about composition; consequently, its contribution to measurements is minimized by adjusting the detector's position relative to the sample. On the other hand, diffuse reflectance, which is described by the Kubelka–Munk theory,[36] is the basis for measurements by this technique.

The Kubelka–Munk function, $f(R_\infty)$, is given by

$$f(R_\infty) = \frac{(1 - R_\infty)^2}{2R_\infty} = \frac{k}{s}$$

where R_∞ is the absolute reflectance of the sample (*viz.*, the fraction of light impinging on it that is reflected), k its absorption coefficient and s its dispersion coefficient. In practice, relative reflectance (R), which is the ratio of the intensity of the light reflected by the sample to that by a standard, is preferred to absolute reflectance. The standard is usually a stable material with a high and fairly constant absolute

reflectance (*e.g.*, Teflon, barium sulfate, magnesium oxide, high-purity alumina ceramics).

The Kubelka–Munk equation can be rewritten in terms of the relative reflectance and the concentration of the absorbing analyte (c):

$$f(R) = \frac{(1-R)^2}{2R} = \frac{k}{s} = \frac{\varepsilon c \ln 10}{s} = \frac{c}{a}$$

where ε is the molar absorptivity and $a = s/2.303\varepsilon$. Thus, a plot of $f(R)$ against c for samples conforming to this relationship will be a straight line of slope $1/a$. However, if the matrix absorbs or the analyte exhibits strong absorption bands, the diffuse reflectance of the sample will not fit the Kubelka–Munk equation and the $f(r)$ *versus* concentration plot will be non-linear. As with Beer's law, the Kubelka–Munk equation is acknowledged to be a boundary equation that is only applicable to weak absorption bands, or when the product of absorptivity times concentration is small. This is so in the NIR region; however, because the matrix frequently absorbs strongly at the same wavelength as the analyte, absorption by the latter cannot be resolved and deviations from the previous equation result.

One widely used practical alternative is a relationship between concentration and relative reflectance similar to Beer's law, namely:

$$A = \log \frac{1}{R} = a'c$$

where A is apparent absorbance, R relative reflectance, c concentration and a' a proportionality constant. Although this relationship has no theoretical basis on the Kubelka–Munk equation, it provides highly satisfactory results under the typical conditions used in many diffuse reflectance spectroscopic applications.

Operational procedures in the NIR

The procedures used in the NIR region are much less labour-intensive than those employed in the MIR and very similar to those used for liquid samples in the UV and visible regions.

The absorbance of a liquid or solution can be readily measured by using quartz or sapphire cuvettes of variable pathlength or fibre-optic probes. No special precautions need be exercised since the absorption of NIR radiation obeys Beer's law. The most suitable solvents in this context are those not containing O–H, N–H and C–H groups, which exhibit little or no absorption in this spectral region.

The NIR spectrum of a solid sample can be obtained by using various types of device. The most frequent choices when the sample is in powder or grain form are reflectance cuvettes with a transparent window material (*e.g.*, quartz) and fibre-optic probes. The latter considerably facilitate recording of spectra; however, light losses resulting from transport along the fibre result in increased signal noise. The spectra of samples in tablet form requiring no pre-treatment (*e.g.*, powdering, sieving, homogenization) can be recorded by using three different types of equipment, namely: (*a*) specially designed reflectance cuvettes for tablets, which, however, provide spectra subject to marked scattering arising from dead spaces between tablets placed in the cuvette; (*b*) a commercially available instrument for recording reflectance spectra for individual tablets, inspired by the double-reflecting sample holder developed by Lodder and Hieftje in 1988;[37] or (*c*) recently introduced instruments that allow the transmission spectra for individual tablets to be recorded.

At this point, it is worth noting that the type of standard to be used for reflectance measurement remains a subject of strong debate. According to ASTM, the perfect standard for this purpose is a material that absorbs no light at any wavelength and reflects light at an angle identical with the incidence angle. Because no single material meets these requirements, the standards used in this context are stable, homogeneous, non-transparent, non-fluorescent materials of high, fairly constant relative reflectance. Springsteen and Ricker[38,39] discussed the merits and pitfalls of materials such as barium sulfate, magnesium oxide, Teflon and ceramic plates as standards for reflectance measurements.

Mathematical processing of signals

The analytical signal obtained in NIR spectroscopy is a complex function that depends on both the physical and chemical properties of the sample; also, it is non-linear owing to scatter, stray light and inconsistency in the instrument response. This entails converting recorded data into apparent absorbance values [$A = \log(1/R)$] or Kubelka–Munk (KM) units when measurements are made in the reflectance mode, and into absorbance units [$A = \log(1/T)$] when made in the transmission mode.

Osborne[40] compared the ability to obtain linear calibrations from raw data, apparent absorbance values [$\log(1/R)$] and KM units, and found that the last two provided calibrations that were not necessarily more linear than those obtained from raw data. He also found that the transformation choice was dictated by the particular data set.[40]

Griffiths[41] claims that the choice of KM or $\log(1/R)$ depends on both the type of sample and the spectral region. For any type of diffuse reflectance measurement where the baseline is irreproducible, band intensities change with it when KM units are used but not when $\log(1/R)$ is employed.

Dahm and Dahm[42] re-assessed the pitfalls of KM units noted by Olinger and Griffiths.[43] They explained why they did not share Griffiths' view that, as a rule, $\log(1/R)$ *versus* analyte concentration plots are more linear in practice than KM graphs; however, they also claimed that the use of $\log(1/R)$ with powdered samples was effective.

Converted values obtained from recorded data are markedly affected by scatter. This becomes apparent when NIR diffuse reflectance spectroscopy is used to analyse solid samples; in fact, the reflectance depends on the degree of scatter of incident light: the more marked the scatter is, the less deep light will penetrate into the sample and hence the smaller will be the absorption.

Light scatter depends essentially on the physical properties of the sample (particle size, crystalline environment) and has a multiplicative effect on the amount of light that is absorbed by the sample, which combines with other additive effects such as baseline shifts or chemical absorption.

The dependence of the signal on the physical properties of the sample, which is highly useful with a view to its characterization and makes the NIR technique a highly suitable tool for determining physical parameters, is a severe hindrance to qualitative analyses for identifying a product where physical appearance is not important, for detecting chemical deviations in the manufacturing process (*e.g.*, heterogeneity, moisture, omission of some component of a preparation), or for the quantitative determination of chemical components—in fact, scatter may be largely responsible for variability between samples, which leads to high correlation among measurements at different wavelengths. These situations call for the prior mathematical processing of spectra in order to minimize the effects of those physical properties of the sample that influence an NIR spectrum and introduce variability that provides irrelevant chemical information.[44–47]

Some of the more widely used mathematical treatments for scatter in NIR spectra[48] include normalization,[49,50] derivation,[51–53] multiplicative scatter correction (MSC),[54–56] piece-

wise multiplicative scatter correction (PMSC),[57] extended multiplicative signal correction (EMSC),[58] optimized scaling (OS),[59,60] standard normal variate (SNV),[61] de-trending (DT),[61] SNV followed by DT (SDT)[62] and DT followed by SNV (DTS).[62] Interested readers are referred to the cited references for information about each specific choice, a detailed description of which is obviously beyond the scope of this review.

Papers concerned with available mathematical treatments for scatter in NIR can be classified into two broad categories, namely: (*a*) those that make empirical comparisons of various treatments or establish relationships among them; (*b*) those that examine the effects of different treatments on NIR spectral quantification.

Prominent in the former group is the paper where Barnes *et al.*[61] demonstrate the efficiency of SNV and DT treatments by application to sucrose samples of different particle sizes, and those where Dhanoa and co-workers demonstrate a linear relationship between SNV and MSC on the one hand,[63] and DTS and SDT on the other.[62]

The latter group includes a large number of references, several of which are concerned with pharmaceuticals. Thus, Jacobsson *et al.*[64] determined sulfasalazine by using a fibre-optic probe, partial least-squares regression and different mathematical treatments (derivative, MSC, PMSC). They found that MSC and PMSC provided the lowest errors of prediction; however, implementation of the latter required a preliminary study in order to optimize window size. Blanco *et al.*[65] used the *V*/*M* ratio of Aucott *et al.*[44] to assess and compare the efficiency of derivative, normalization, MSC, SNV, DT and DTS treatments with a view to reducing the effects of scatter on mixed-phase spectra for a pharmaceutical preparation. They found that the derivative and SNV treatments provided the best results, and normalization the worst, for the case considered.

Qualitative analysis

Identification and qualification of raw materials and pharmaceutical preparations

Quality control involves implementing suitable procedures or measurements in order to ensure the identity of the materials at each stage of the manufacturing process, from the time the raw materials are received to that when the end products are released. NIR spectroscopy is an advantageous alternative to wet chemical methods and other instrumental techniques such as MIR spectrophotometry and NMR spectroscopy for this purpose.

Similarly to MIR, the earliest NIR studies aimed at the identification of substances were concerned with structural elucidation. In 1954, Kaye[66] assigned the bands in the spectra of bromoform, chloroform, benzene, methanol and *m*-toluidine, and studied the factors that influenced the position and intensity of the bands (inter- and intramolecular interactions, temperature, physical state of the sample). Several years later, Sinsheimer and Keuhnelian[67] stated that NIR spectra were among the most effective means for distinguishing dissolved primary, secondary and tertiary amines; thus, primary amines differ from secondary and tertiary amines by the presence of a band at 2180 nm and the absence of another at 2050 nm.

However, identifying a substance from the mere inspection of its NIR spectrum is usually difficult since this consists of very broad, usually overlapped bands that call for pattern recognition procedures (statistical treatments used to characterize spectra).[68] Essentially, the identification process involves two steps, *viz.*, recording a series of analytical signals for the product and generating a so-called 'spectral library', and recording the sample signal and comparing it with those in the previously compiled spectral library on the basis of mathematical criteria for parametrizing spectral similarity. If the similarity level exceeds a pre-set threshold, then the spectra are considered identical and the sample is identified with the corresponding product in the library.

Reliable identification of a product relies on correct choice of spectra for inclusion in the library. The spectra compiled for each product should contain every possible source of variability associated with spectral recording and the product's manufacturing process. Spectral variability is considered by including spectra for the same sample recorded by different operators on different days; manufacturing variability is considered by including spectra for samples from different production batches. It is difficult to anticipate the exact number of spectra to be included in a 'comprehensive' library. For a product that is manufactured in a highly reproducible manner, manufacturing variability can be spanned by samples from 5 to 10 different batches and a total of 20–40 spectra. If manufacturing reproducibility is poor, the number of spectra required can easily double.

One other important consideration in building a spectral library is checking that all the spectra included are correct. Uncontrolled factors (*e.g.*, incompletely filled cuvettes, voltage drops at the time of recording) may result in spectral differences not ascribable to natural variability; any such spectra should be discarded.

One widely used NIR mathematical treatment for expressing similarity is the correlation coefficient,[69,70] which is defined as the cosine of the angle between the vectors for the sample spectrum and the average spectrum for each product included in the library:

$$\rho_{jk} = \frac{\sum_{i=1}^{p}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^{p}(x_{ij} - \bar{x}_j)^2}\sqrt{\sum_{i=1}^{p}(x_{ik} - \bar{x}_k)^2}}$$

where p is the number of wavelengths; subscripts k and j denote the sample and reference product, respectively; x_i is the measured value at wavelength i; $\bar{x}_j$ is the average spectrum of the reference product j; and $\bar{x}_k$ is the average spectrum of the sample.

If the similarity coefficient exceeds a pre-set threshold, then the two spectra compared are considered identical and the sample is identified with the reference product. Theoretically, if the two spectra are coincident, the correlation coefficient should be unity; however, random noise associated with any type of spectral measurement precludes obtaining a coefficient of exactly 1.

This parameter has the advantage that it is independent of library size and concentration changes, which permits correct identifications by use of libraries consisting of a small number of spectra. On the other hand, it is calculated from second-derivative spectra; hence, samples of the same product in different grain sizes will have the same correlation coefficient since particle size only affects band intensity in second-derivative spectra.

Van der Vlies and co-workers used a correlation coefficient which they called the spectral match value (SMV) as a simple, expeditious and precise tool for identifying different types of cellulose[71] and ampicillin trihydrate.[72]

Blanco *et al.*[70] demonstrated the discriminating ability of a correlation coefficient that they called the match index (MI) in the identification of a pharmaceutical preparation by use of a library consisting of 163 substances including excipients, active compounds, amino acids and vitamins.

One especially interesting application in this context is the non-invasive NIR method of Galante *et al.*[73] for assessing microbiological contamination in injections. The method, based

on correlation measurements, is fast, detects contamination by various types of microbe (yeasts, mould and bacteria), avoids contamination by the analytical method itself and can be implemented on-line with the manufacturing process.

Replacing conventional identification techniques, while important, is not the sole advantage of NIR spectroscopy for qualitative analytical purposes. In fact, this technique also affords qualification. The pharmaceutical industry must guarantee the correct dosage, manufacture and stability of each product, so it must carefully control raw materials and each step of the manufacturing process for factors such as potency, moisture, density, viscosity and particle size in order to detect potential deviations and correct them in a timely manner. Controls can rely on numerical determinations of the target parameters by using qualitative methods of analysis or comparing the NIR spectrum for the sample with the body of spectra for samples complying with the specifications and encompassing every possible source of natural and manufacturing variability. This latter choice is known as 'qualification' and involves expressing similarity in distance terms in order to determine whether a sample falls within the normal variability range or is subject to manufacturing deviations that call for comprehensive analysis. Distance-based methods rely on a compromise between the maximum number of wavelengths that can be used—in fact, if the wavelengths are correlated, increasing their number will increase the distance without providing additional information—and the minimum number required to encompass all possible sources of manufacturing variability in the product.

One of the most widely used qualification methods is probably the wavelength distance method,[70] which assumes that measurements at each wavelength are distributed according to the normal law. It generally uses the second-derivative of spectra from a library that defines the accepted variability for the product to obtain an average spectrum and the standard deviation at each wavelength. The distance between the unknown sample and the average spectrum for the reference product at each wavelength is calculated and the most unfavourable situation (*viz.*, the wavelength that results in the maximum distance) is determined from the following equation:

$$d_{kj} = \max \frac{\left| x_{kp} - \bar{x}_{jp} \right|}{s_{ij}}$$

where subscripts k and j denote sample and reference product, respectively; x_{kp} is the measured sample value at wavelength p; $\bar{x}_{jp}$ is the average spectrum of reference product j at wavelength p; and s_{jp} is the standard deviation of the measured values for reference product j at wavelength p.

If the sample belongs to the same population as the reference product, then there will be a probability of 99.7% that the distance will be less than three times the standard deviation. If the maximum distance does not meet this criterion, then the sample must belong to a different population (*i.e.*, it will not meet the qualification criterion). The qualification criterion based on the expression $D_{\max} \leqslant 3\sigma$ is usually too conservative; it is often more practical to have users decide on the most suitable limit for their own problems and working methods.

Correct usage of this method entails exhaustive control of the instrument in order to ensure that noise remains roughly constant, since measurements at individual wavelengths and derivative spectra tend to introduce noise and wavelength shifts.

One shortcoming of this method is the risk of false-negatives at wavelengths coinciding with x-intercepts in second-derivative spectra (zero cross-over). If the standard deviation for the average spectrum at a given wavelength is very small, then the distance at that wavelength will be very large and a negative qualification will result. This may be the case when second-derivative values are very close to zero. This problem can be circumvented by using the wavelength library stabilization method,[74] where the average spectrum and its standard deviation behave as though each second-derivative spectrum had been shifted by a fraction of a nanometre to the left and right along the wavelength axis (stabilization constant) in such a way that the standard deviation at zero cross-over points will be increased and false-negative qualifications avoided.

Plugge and van der Vlies[72,75] used the wavelength distance method to determine what they called the conformity index (CI), which is seemingly highly sensitive to sample impurities and occasionally allows one to pinpoint the sources of the inability to qualify a raw material or product by using a C-PLOT (*viz.*, a plot of the absolute distance at each wavelength as a function of the wavelength itself). One constraint of the C-PLOT is that it does not provide a sign of manufacturing deviations. Based on the C-PLOT, González and Pous[76] developed DISPLOT (a plot of the distance, sign included, as a function of wavelength), which identifies the sign of small chemical and/or physical deviations introduced during the manufacture of the mixed phase of a product.

One alternative to the wavelength distance method is to use the whole information contained in the spectrum by calculating the Mahalanobis distance.[77–79] This parameter is useful for cluster analysis and can be calculated for multi-dimensional spaces. The distance between the sample and the centre of the cluster formed by the spectra of the reference product is defined as

$$D^2 = (\mathbf{X}_j - \bar{\mathbf{X}}_k)^{\mathbf{T}} \mathbf{C} (\mathbf{X}_j - \bar{\mathbf{X}}_k)$$

where $\mathbf{X}_j$ is the vector describing the spectrum of sample j, $\mathbf{X}_k$ is the vector for the average spectrum of reference k, $\mathbf{C}$ is the matrix that describes distance measurements in the multi-dimensional space studied and superscript $\mathbf{T}$ denotes transpose matrix.

Usually, if the distance thus obtained is less than three times the standard deviation, the sample meets the qualification criterion (*i.e.*, the manufacturer's specifications).

In the early 1990s, Corti and co-workers published several papers reporting on the use of the Mahalanobis distance for quality control in various pharmaceutical preparations. They qualified chloroform extracts of samples containing 0.05% estrogen and 0.25% progesterone or only one of them[80] and discriminated among creams containing the same active principle but different excipient proportions.[81] They showed that the Mahalanobis distance was a highly effective choice for qualifying antibiotics[82,83] as it allows one to discriminate between their crystalline and amorphous forms, as well as among mixtures containing variable concentrations of the same antibiotic. Finally, they obtained a high reproducibility in the classification of organic and inorganic raw materials for which spectra had been recorded by different operators under different conditions.[84]

Dreassi and co-workers used the Mahalanobis distance to discriminate among samples of the same active principle differing in some physical and/or chemical property,[85] as well as for quality control in the production of an antibiotic, where their method allows samples to be characterized at different stages of the process[86] and distinguishes them from other products manufactured in the same production area.[86,87]

Recently, van der Vlies *et al.*[88] developed a procedure for the qualification of pharmaceuticals based on the conversion of NIR spectra to polar coordinates and the subsequent calculation of the corresponding Mahalanobis distance (the polar qualification system). The method uses spectra, which facilitates relating the distribution of the products to their spectral features (*e.g.*,

the presence of an impurity absorbing in a specific region); also, graphs are two-dimensional and hence easy to interpret. Notwithstanding these advantages, it remains to be proved whether this method surpasses existing alternatives in practice and is applicable to extensive spectral libraries—in fact, its use is seemingly restricted to the discrimination of similar products with known spectral features. Plugge and van der Vlies[89] showed that the method allows one to discriminate chemically identical substances from different suppliers and also to detect differences in physical properties among samples from the same supplier.

One alternative to direct spectral computations, where correlations and distances are calculated in the wavelength space, is the use of principal component analysis (PCA)[90] to reduce variables (in PCA, correlations and distances are calculated from scores in the space bound by the principal components). Because the number of data involved is smaller—a large number of wavelengths is replaced with a few principal components (PCs)—library searches are much faster; however, all spectra in a library influence PC calculations, so every time a new spectrum is included in or an existing one is excluded from the library, PCs must be recalculated and the modified library validated. In addition, equipment-bundled software usually restricts the maximum number of PCs that can be used in the calculations, which in turn limits the number of different products that a spectral library can contain.

Lo and Brown[91] used the correlation coefficient in the PC space to identify components in mixtures of organic solvents. The ensuing method is selective, requires no prior knowledge of the mixture composition and avoids variability due to spectral noise. Wu *et al.*[46] identified tablet blisters containing different amounts of an active principle by using PCA and different mathematical treatments of the spectra. Second-derivative calculations proved to be the most effective transformation as regards discriminating power.

Shah and Gemperline[92] qualified different batches of Avicel PH101 microcrystalline cellulose by using the Mahalanobis distance in the PC space. Their criterion was to assume that a sample was qualified when the probability level for a χ^2 distribution fell in the range 1.0–0.05.

One other classification procedure used in some reported applications is the soft independent modelling of class analogy (SIMCA).[93] Each of the products in the sample is subjected to PCA and Fisher's test is subsequently applied in order to estimate the likelihood of a sample belonging to the class defined by the spectra of the reference product. The residual variance for a spectrum k to be identified (S_k^2) that is assumed to belong to class j (defined by the spectra of the reference product j) is divided into the total variance for the samples belonging to class j (S_0^2) in order to obtain the following variance relationship:

$$F = \frac{S_k^2}{S_0^2} \frac{n}{n-a-1}$$

where n is the number of spectra for the reference product and a the number of PCs used to construct the class model.

Gemperline *et al.*[94] used SIMCA to qualify 400 first-derivative spectra for six raw materials and found it to be sensitive to the presence of impurities such as production intermediates and degradation products, as well as to particle size. In subsequent work, Gemperline and Boyer[95] identified and qualified samples adulterated with small amounts of impurities by using libraries of variable size and the Mahalanobis distance, maximum distance and SIMCA method. The maximum distance proved to be the most suitable choice for identifying samples with small spectral libraries, but performed worse for qualification purposes as it was insensitive to impurities at levels below 2%.

Shah and Gemperline[96] used the Mahalanobis distance and the SIMCA method to classify different types of cellulose and detect impurities at levels of 0.1–2%. They found that temporal changes in instrument response influenced the limit of detection; consequently, reliable detection of impurities required considering such changes in the spectral libraries. Dempster *et al.*[97] used the maximum distance to confirm the identity of tablet blisters containing different concentrations of the active principle (5, 10 and 20% m/m). They compared three different spectral recording procedures, namely: (*a*) extracting tablets from their blisters prior to measurement; (*b*) making measurements through blisters, using the horizontal set-up presentation module; and (*c*) using a fibre-optic probe for measurements. The first procedure proved to be the most sensitive as it distinguished among the three concentration levels tested and the placebo; on the other hand, the other two failed to discriminate the 5% sample and the placebo but had the advantage that they were non-invasive. Subsequently, they used the fibre-optic probe in conjunction with the maximum distance, Mahalanobis distance and SIMCA to confirm the identity of coated and uncoated tablet blisters.[98] They used only those spectral zones where differences among products were maximum and the results were optimum, provided that the tablets and blisters to be qualified had been manufactured under the same conditions as those included in the library.

Ciurczak and Maldacker[99] compared the ability of cross-correlation spectral reconstruction methods and that of discriminant analysis based on the Mahalanobis distance to classify tablets in terms of their active principle. The spectral reconstruction method, developed by Honigs *et al.*,[100] allows one to obtain the individual spectrum for each mixture component and determine the nature of interactions among analytes; however, it is less suitable for classification purposes.

One qualification alternative similar to that using the Mahalanobis distance is the boostrap error-adjusted single-sample technique (BEAST),[101,102] which uses reflectance values obtained at pre-set wavelengths to obtain a multi-dimensional data distribution. The chief difference between the Mahalanobis distance and BEAST is that, in the latter, the confidence limits used to define the clustering limits consider asymmetry in the sample distribution rather than the symmetric distribution assumed in the Mahalanobis distance. The most severe shortcoming of BEAST is that it requires extensive data storage resources. Lodder and co-workers showed that the use of NIR spectroscopy in combination with BEAST provides a rapid method for detecting adulterants (Fe_2O_3, NaF, NaCN, KCN and As_2O_3) in capsules[103] and allows discrimination among aspirin tablets from different manufacturers[37] with no need for sample pre-treatment—and hence with minimal manipulation errors.

This overview of the qualitative applications of NIR spectroscopy in the pharmaceutical industry would be incomplete if no mention were made of the fact that this technique has been endorsed by several agencies in their official methods of analysis (Table 5).

Determination of homogeneity

One important operation in manufacturing solid pharmaceuticals is monitoring of the homogenization process, which determines the encapsulation or compression quality of the end product. Almost every application of NIR spectroscopy in this field has been reported recently and uses one of the qualification methods described in the previous section. Ciurczak[104] developed three different approaches to the monitoring of the homogenization of aspirin–vitamin B_{12} mixtures by use of fibre optics, namely: (*a*) visual comparison of second-derivative spectra recorded at different homogenization times; (*b*) calcula-

tion of the correlation coefficient; and (*c*) calculation of the maximum distance. The last proved to be the most reliable method for determining the end-point of the homogenization process as it discriminated between the penultimate and last mixtures.

The qualification concept was used by Wargo and Drennen[105] to verify the homogeneity of solid mixtures and determine the optimum homogenization time for a preparation containing hydrochlorothiazide as the active principle. Qualitative analytical algorithms based on BEAST proved to be more sensitive to variations in sample homogeneity than did a χ^2 test as the former uses the entire NIR spectrum.

van der Vlies and co-workers converted NIR spectra into polar coordinates and used these to calculate the Mahalanobis distance[88] in order to identify non-homogeneous samples.[89] They found that analysis of variance (ANOVA) was an effective choice for validating homogenization processes.

Hailey *et al.*[47] and Sekulic *et al.*[106] developed systems for monitoring the homogenization of solid mixtures based on measurements *via* a fibre-optic probe fitted to the mixer. The most salient advantage of these systems is that they permit the determination of the end-point of the homogenization process in real time and in a non-invasive manner. In both cases, mixture homogeneity is determined by plotting the standard deviation for several replicates against the homogenization time.

Polymorphism and optical isomers

NIR spectroscopy was used by Gimet and Luong[107] for the qualitative control of a dimorphic analgesic. The pure forms exhibit NIR spectra that are sufficiently different in their maximum wavelengths and absorbances to allow quantification. Mixed spectra confirm that quantitative analyses are possible even in the absence of qualitative differences between the spectra. The ensuing method is applicable to substantial amounts of product, which avoids errors arising from sampling and sample heterogeneity.

The ability of the Mahalanobis distance and SIMCA to identify and assess the polymorphic quality of a drug was compared by Aldridge *et al.*[108] The former proved to be the more effective choice since, in addition to discriminating between the polymorph of interest and other substances with highly similar spectra, it is more sensitive to the presence of low levels of impurities in the polymorph.

Norris *et al.*[109] used NIR spectroscopy to monitor polymorphic conversion. Their method subjects spectra recorded over the course of the reaction to PCA and allows the end-point of the process to be determined in real time.

Buchanan *et al.*[110] determined the enantiomeric purity of the optically active forms of valine. Mixtures containing D- and L-valine in different proportions exhibited identical spectra except for baseline shifts resulting from differences in particle size. However, when the mixtures were dissolved and recrystallized, the resulting spectra exhibited qualitative and quantitative differences that permitted the determination of enantiomeric purity in the starting product.

β-Cyclodextrin and silica gel were used as chiral selectors for distinguishing the (1*R*)-(+) and (1*S*)-(–) enantiomers of α-pinene by NIR transmission spectroscopy.[111] The bond between the (+) enantiomer and a reagent is different from that between the (−) enantiomer and the same reagent, which facilitates discrimination of the enantiomers by PCA.

Quantitative analysis

NIR spectra typically contain broad, overlapping bands that cannot always be ascribed to an individual sample component. As a result, whenever the NIR technique is used for quantitative purposes—whatever the physical or chemical property of the sample to be determined—a calibration must be performed by using an existing multivariate procedure.[69,90] Essentially, the procedure for quantification using multivariate calibration involves the following steps: (*a*) selecting a representative sample set; (*b*) acquiring the analytical signals and obtaining the reference values; (*c*) mathematical processing of the signals; (*d*) selecting the model that relates the property to be determined and the signals; and (*e*) validating the model. Each step is described in detail below, with special emphasis on the problems arising from NIR analyses of pharmaceuticals.

Halsey[112] devised a protocol for developing quantitative NIR methods for the pharmaceutical industry. Although the protocol is based on the NSAS software package, bundled with NIRSystems instruments, it can provide users of equipment from other manufacturers with basic concepts to be considered in developing an NIR analytical method.

Sample selection

The starting point for every calibration technique is a set of samples which have previously been analysed by a reference method, span the working concentration range and are representative of the manufacturing variability sources that are bound to influence the NIR spectra.

One of the problems encountered in using NIR spectroscopy for the quantitative analysis of pharmaceuticals is the need to obtain a sample set that can be used to establish a calibration model. As a rule, all available production samples contain the active principle and excipient in amounts very close to the nominal values; this precludes spanning a wide enough concentration range for calibration. One way of circumventing this shortcoming is by using a set consisting of production and laboratory-made samples; the former will introduce the variability sources typical of the production process while the latter will expand the narrow range spanned by the former. Therefore, the two essential questions that arise when developing an NIR quantification method are as follows: what concentration range is the sample set to span? and how can preparation of the laboratory samples be approached? Regarding the former question, Corti and co-workers[83,84,113] claim that a sample set spanning a concentration range about ±5% of the nominal value affords precisely and reproducibly sufficient calibration for quality control purposes. However, such a range may be too narrow if the manufacturer's tolerated limits are greater than ± 5% of the nominal value. In order to expand the concentration range without altering any physical properties potentially affecting NIR spectra, one can make the samples at a pilot plant,[114,115] prepare laboratory samples containing each compo-

Table 5 NIR methods endorsed by various official agencies

Agency	Method
Food and Drug Administration (FDA)	Identification, quantification and determination of moisture content in ampicillin trihydrate (Gist Brocades)
Health Protection Branch (HPB)	Identification of raw materials and packaging components (Merck Frosst Canada)
Norwegian Medicines Control Authority (SLK)	Identification and quantification of paracetamol, and determination of moisture content, in Paracet 500 mg (Wieders Farmasoytiske)
Medicinal Controls Agency (MCA) in UK	Identification of Zovirax 200 mg (Glaxo Wellcome)

nent of the pharmaceutical at concentrations over the manufacturer's specified ranges,[116] or over- and underdose samples from different production batches with small amounts of the active principle or excipient, respectively, to obtain the desired concentration range.[70,86,87] The first approach is probably that providing the samples that are closest in composition to production samples; however, it is also the least feasible in practice as it is rarely possible to manufacture production batches suited to particular needs, nor is it possible to ensure that a small-scale process will be comparable to the actual production process. The principal advantage of the second approach is the ease with which the different concentrations needed to span the required range can be obtained; however, the grinding, mixing and other miscellaneous processes used in the laboratory may differ substantially from those used in the manufacturing process and hence lead to samples differing markedly—NIR spectra included—from production samples. Over- and underdosing production sample make expanding the concentration range a labour-intensive, care-demanding task; provided that strict control is exercised, however, variability in the physical features of the samples can be much smaller than in the previous case.

Blanco *et al.*[117] found that the use of over- and underdosed samples did not alter the quality of the results for production samples; they quantified production samples by using calibration sets consisting of an increasing number of laboratory-made samples. The same authors[118] compared two different approaches, *viz.*, preparing samples by (*a*) weighing of all components and (*b*) over- and underdosing samples from different production batches, to quantify the active principle in the mixed phase of a commercially available preparation. Although the results obtained with both approaches were similar, the over- and underdosing approach resulted in simpler calibration models and in slightly smaller errors of prediction.

Once the calibration set has been established, it is split into two sub-sets, *viz.*, a calibration set consisting of a small number of samples that are representative of the entire set and allows the determinand to be related to the analytical measurement, and a prediction set composed of the remainder of samples that is used to assess the predictive ability of the model. In splitting the original set, the questions arise as to what the optimum number of samples to be included in the calibration set is and how such samples should be chosen.

The use of a small number of samples in the calibration set may result in some source of variability in the product being excluded and hence in spurious results in analysing new samples. Several workers claim that the optimum number of samples depends on their complexity, on the concentration range to be spanned and on the particular calibration method used.[113,119] Thus, when the aim is to quantify 1–4 components and the samples exhibit no large differences in their physical and chemical properties, a calibration set consisting of a minimum of 15–20 samples will be more than adequate.

As regards the second question, there are two types of approach, *viz.*, those focusing on general aspects of sample selection for NIR calibration[69,120] and those based on comparisons among available choices. Although, in general, the latter have been developed for and applied to food samples, they are worth mentioning here because they are also applicable to pharmaceuticals.

Honigs *et al.*[121] used a sample selection method similar to Gaussian elimination, They selected 'unique' samples in a sequential manner in order to identify that exhibiting the highest NIR absorbance. The selected sample was removed from the remaining set and the process was repeated until the desired number of samples was chosen or the absorbance values of the remaining spectra were below a pre-set limit.

Næs[122] and Isaksson and Næs[123] developed a method for selecting samples based on an unsupervised pattern recognition algorithm that is applicable to highly collinear data. They identified clusters of closely related samples by constructing a dendrogram based on the PCA scores for the sample spectra. From each cluster, the sample falling at the greatest distance from the cluster centre was chosen. A similar sample selection system was reported by Puchwein.[124] The sample with the greatest Mahalanobis distance from the cluster centre was selected first and those samples with a Mahalanobis distance similar to the selected sample were left out. Subsequent samples were chosen similarly from among the remainder.

The normalized Mahalanobis distance was used by Mark[125] to select samples on the basis of discrete wavelengths. One disadvantage of this method is the difficulty involved in selecting a suitable wavelength.

Ferré and Rius[126] reported a procedure for selecting the best calibration set for principal component regression (PCR) based on a D-optimum design and on instrumental responses alone. Calibration investment and effort are reduced if the reference method is applied to the selected samples only. This approach was criticized by Davies,[127] who stated that the prediction set provided by the method was strongly dependent on the calibration set and thus a poor choice for assessing the predictive ability of the model.

Blanco *et al.*[70] compared flat calibration,[128] which involves spanning the working concentration range with a large number of samples, and the sample selection subroutine included in the NSAS software package,[129] with a view to selecting calibration samples for the quantification of the active principle in a commercially available preparation. The flat calibration approach provided more robust calibration models. The same authors proposed using PCA to select the production batches best representing variability in the manufacturing process, which must be included in the calibration set in addition to laboratory-made samples.[116,118]

Multivariate calibration methods

The calibration methods most frequently used in NIR spectroscopy in order to relate the property to be measured to the analytical signals acquired are multiple linear regression (MLR),[130,131] PCR[90] and partial least-squares regression (PLSR).[90] Most of the earliest quantitative applications of NIR spectroscopy rely on MLR because spectra were then recorded on filter instruments, which afforded measurements at a relatively small number of wavelengths only. Applications involving PCR and PLSR have proliferated after the introduction of commercially available instruments that allow the whole NIR region to be scanned.

The choice of the calibration method is dictated by the nature of the sample, the number of components to be simultaneously determined, the *a priori* knowledge of the system studied and available data on it. Below are briefly described the features of the different calibration options.

The MLR technique is the usual choice with filter instruments and is also occasionally used with instruments that record whole spectra. It is an effective calibration approach when the analytical signal is linearly related to the concentration, spectral noise is low and the analyte does not interact with other sample components. The MLR technique also affords modelling some non-linear relationships as it assumes that modelling errors arise from concentrations. However, it can only be used at a small number of wavelengths, which, if incorrectly selected, may result in overfitting (*i.e.*, in modelling of noise or random errors). Also, if spectral data are highly collinear, then the precision of the results suffers appreciably. A detailed description of available procedures for determining how many and which wavelengths should be used can be found elsewhere.[90,132,133]

Whole-spectrum methodologies (*viz.*, PCR and PLSR) have the advantage that they use every single wavelength in a recorded spectrum with no prior selection. Also, they allow the simultaneous determination of several components in the same sample and avoid the problems associated with collinearity among spectral data and with noise-related variability. As noted earlier, non-linearity in NIR signals is ascribed to non-linear detector responses that result in curved signal–concentration plots, as well as to physical and/or chemical factors giving rise to shifts and width changes in spectral bands.[134,135] In some cases, non-linearity is so marked that a non-linear calibration methodology such as neural networks,[136–140] locally weighted regression,[141,142] projection pursuit regression[143,144] or quadratic versions of the PCR or PLSR algorithms[145,146] must be used.

Determination of physical parameters

Particle size determinations are of paramount importance to the pharmaceutical industry as incorrect grain size analyses can lead to altered properties such as coating power and colour, hinder subsequent mixing of powders (for tablet formulations) or powders and liquids (suspensions), and result in defective pressing of solid mixtures for making tablets. Because particle size is one of the physical parameters most markedly influencing NIR spectra, the NIR technique is an effective alternative to the traditional methods involving sieving, light scattering by suspensions, gas adsorption on solid surfaces or direct inspection under a microscope.

Ciurczak *et al.*[147] used the linear dependence of band intensity at a constant concentration on the average particle size at a pre-set wavelength to determine pure substances and granules. Absorbance *versus* particle size plots at different wavelengths exhibited two linear segments. These authors postulated that the sample's absorption coefficient in the Kubelka–Munk function was large below 85 μm and ascribed the abrupt decrease in absorbance from 250 to 85 μm to the influence of such a coefficient. Consequently, the effect of particle size on reflectance measurements was significantly reduced below 80 μm. Blanco *et al.*[148] determined the average particle size of Piracetam over the wavelength range 175–325 μm with an error of 15 μm, based on the assumption that an increase in particle size would produce an increase in absorbance that could be measured and used to quantify the former by MLR and PLSR calibration. They found that spectral reproducibility varied in an exponential manner with particle size and that sample compactness was the most influential factor on particle size.

Ilari *et al.*[149] investigated the feasibility of improving the determination of the average particle size of two highly reflecting inorganic compounds (*viz.*, crystalline and amorphous NaCl) and an NIR-absorbing species (amorphous sorbitol), using the intercept and slope obtained by subjecting spectra to MSC treatment as input parameters for PLSR. While particle size continues to be the physical property of samples most frequently determined by NIR spectroscopy, several other parameters such as the dissolution rate and the thickness and hardness of the ethylcellulose coating on theophylline tablets have also been determined, all with good errors of prediction.[150]

Determination of moisture content

The presence of crystallization or adsorbed water in pure substances and pharmaceutical preparations, whether during treatment of the sample or its storage, causes significant changes in those properties that influence chemical decay rates, crystal dimensions, solubility and compaction power, among others. NIR spectroscopy is an effective alternative to traditional analytical methods such as thermogravimetry and Karl–Fischer (KF) titration as water gives a characteristic absorption spectrum the mere visual inspection of which allows one to determine, for example, if different batches of a given substance contain also different amounts of moisture.[151]

The NIR spectrum of water exhibits five absorption maxima at 760, 970, 1190, 1450 and 1940 nm; the positions of these bands can be slightly shifted by temperature changes[152–154] or hydrogen bonding between the analyte and the matrix.[155,156] The bands at 760, 970 and 1450 nm correspond to the first three overtones of O–H stretching bands, whereas the other two arise from combinations of O–H oscillations and stretching. The specific band to be used for determining water depends on the desired sensitivity and selectivity levels.[157] As a rule, the overtone bands are appropriate for this purpose when using solutions in solvents containing no O–H groups; on the other hand, the band at 1940 nm provides increased sensitivity.

In Tables 6 and 7, available NIR methods for determining moisture are classified according to whether they rely on transmittance or reflectance measurements, respectively.

NIR transmittance methods are mainly used to determine water in solvents. Their earliest applications to solid preparations entailed dissolving the sample in a solvent with little or no absorption in this spectral region. All methods of this type use least-squares calibration to construct a straight line from absorbance values at the absorption maximum at about 1900 nm for solutions containing variable concentrations of the target species.

Table 6 Applications of NIR transmittance spectroscopy to the determination of moisture content

Sample type	Remarks	Ref.
Solid	The most suitable solvents for determining water in solid samples are pyridine and methanol, which exhibit no absorption band at 1900 nm; also, their mixtures with water obey Beer's law over a wide composition range	155
Solid	Determination of trace amounts of water (0.05%) in mono-, di- and triglycerides using chloroform as solvent. Free from interferences from triglyceride OH groups	158
Solid	Use of the NIR technique in conjunction with dimethyl sulfoxide as solvent provides an expeditious, accurate and precise alternative to the traditional method for the determination of water in starch, which is time-consuming and involves cumbersome manipulations of the viscous, sticky samples involved	159
Solid	Methanol is used to determine water in organic compounds and pharmaceutical preparations. Results are consistent with those provided by the conventional dehydration method and KF titration. The spectrophotometric method is more reproducible, simple and expeditious	156
Liquid	Determination of small amounts of water in solvents (acetonitrile, propionitrile, tetrahydrofuran and dimethylformamide). The NIR method is less sensitive than KF titration and its LOD is about 20 ppm. The former is more rapid and flexible, and involves less extensive sample manipulation	160
Liquid	Flow injection analysis method for the determination of water in dichloromethane and isobutyl methyl ketone, the LODs for which are 0.01 and 0.005% v/v, respectively. Free from sample contamination by environmental moisture	161

Determination of active compounds and excipients

The number of determinations of active compounds and excipients has grown enormously in recent years. Table 8 summarizes reported uses of the NIR technique for this purpose.

Calibration transfer

In previous sections, we discussed the most important considerations in developing a quantitative NIR method. However, this review would be incomplete if one of the major hindrances to the application of developed methods, *viz.*, transferability of calibration models among instruments,[177–179] were not mentioned. Because detector responses are not uniform, the signals recorded by two different instruments may differ owing to wavelength shifts and/or changes in measured intensities. This precludes the use of one instrument's calibration model by another. The problem can be overcome in three different ways, namely: (*a*) by recording spectra for calibration samples on each instrument and constructing a calibration model for each, which is not feasible when a large number of samples are to be processed, the instruments are very distant from each other or the samples are unstable; (*b*) by applying mathematical corrections to the spectra recorded by one instrument so that they can be used in the calibrations obtained with the another; (*c*) by transferring the calibration model from one instrument to another.

The principal calibration transfer methods.[180] are briefly commented on below. Some come from fields other than the pharmaceutical field but are indeed applicable in many sectors.

The method of Shenk and Westerhaus[181,182] uses a large number of stable samples for transfer and corrects the response of an instrument at each wavelength with reference to that of a primary instrument of identical resolution; after wavelengths have been corrected, the spectral intensity is adjusted. The applicability of this method has been assessed by several workers. Dardenne *et al.*[183] focused on the problems encountered in transferring calibrations among different types of instrument and on the sample set required for this purpose. Bouveresse and co-workers found that the method provided satisfactory results as long as the transfer samples span the same spectral intensity range and are of the same nature as those to be subsequently analysed;[184] they investigated whether altering the spectral intensity correction algorithm improved the quality of the results when some of the previous conditions were not fulfilled.[185]

Mark and Workman[186] developed a method suitable for MLR calibration that uses no transfer sample set; rather, calibration transfer relies on a model that is constructed from those wavelengths that remain unchanged relative to spectral shift as independent variables. The previous two methods are applicable at a relatively small number of wavelengths and are usually incompatible with multivariate calibration based on whole spectra. Although the method of Shenk and Westerhaus[181] affords whole-spectrum correction, its applicability to multivariate calibration approaches is restricted by the fact that corrections rely on a univariate scheme.

Wang *et al.*[187] developed four calibration transfer methods based on a multivariate scheme; all four use whole spectra and an unrestricted number of wavelengths in the calibration model. Two of them correct the calibration model (one with respect to a classical calibration model and the other relative to an inverse calibration model); the other two, based on direct standardization (DS) and piecewise direct standardization (PDS), correct the response of an instrument so that its spectra can be used by another. The DS method uses a PCA to obtain the transformation matrix that relates the responses of both instruments, but has the disadvantage that it requires a large number of samples. The PDS method[188,189] relies on the fact that spectral variations are usually restricted to a small region; hence it reconstructs each point in the spectrum from one instrument by using several measurements through a small window in the spectrum from another. As a result, this method requires a fairly small number of transfer samples that need not span the whole concentration range spanned by the calibration samples.

Bouveresse *et al.*[190] compared the PDS method with the slope/bias correction method used by Jones *et al.*;[191] when differences between instrument responses are small, the latter method provides good results, the quality of which can be assessed by Fisher's test if the number of transfer samples used exceeds five.

The method of Forina and co-workers[192,193] uses PLSR to establish the relationship between the transfer samples processed with both instruments and then the regression equation for the first instrument.

Blank *et al.*[194] used a calibration transfer method based on a finite response filter to relate the response of a spectrophotometer to that from a second instrument without the need to record spectra for transfer samples on the latter.

Miscellaneous applications

This section comments on some uses of NIR spectroscopy with their roots outside the pharmaceutical field but of potential interest to the pharmaceutical industry.

Table 7 Applications of NIR reflectance spectroscopy to the determination of moisture content

Sample type	Calibration	Remarks	Ref.
Injection	MLR PLSR	Comparison of calibration methods by using products containing different amounts of active compound (0.5 and 1.5 mg per vial). Spectra are recorded through vial bottoms, using the horizontal set-up sample presentation module	162
Solid	MLR	Determination of moisture content (11.5–15%) in ceftazidime, with errors of 0–5%. The calibration set encompasses ±10% of the nominal value and consists of samples from different production batches and laboratory-made samples in 1:1 ratios	163
Solid	MLR	NIR method for determining moisture in the antibiotic ampicillin trihydrate; approved by FDA in 1992	72
Solid	MLR PLSR	Determination of moisture in the active compound ferrous lactate dihydrate (11.1–14.6%), using a fibre-optic probe. The two calibration methods used provide similar results, with errors less than 1.5%	164
Solid Tablets	MLR	Determination of moisture at different production stages (mixed phase, cores and tablets), using a single calibration equation that provides prediction errors less than 4%. Moisture contents below 1% can seemingly not be detected by reflectance measurements	87
Tablets	MLR	Determination of moisture in tablets with a maximum certified content of 2%. The results obtained over a one year period reveal that the method is suitable for quality control analyses	165

Reeve studied changes in NIR spectra due to various factors (moisture, pH and state of aggregation of the sample, among others).[195,196] The presence of moisture in the samples was found to shift absorption bands to an extent dependent on the type of compound concerned–shifts are especially prominent with alcohols and ketones, and less marked with acids; also, the spectral features of solid samples vanish on dissolution, which destroys the crystal structure.[197] Reeve also investigated the interactions between monomers and polymers of carbohydrates (glucose and sucrose with amylose, and amylopectin and cellulose with starch) that affect NIR spectra and assessed their effects on calibration methods.[198]

In several papers, Lin and Brown[152,199–201] showed that MLR and PCR allow one to construct calibration models for determining NaCl in aqueous solutions as they afford measurements of the small intensity changes undergone by bands in the presence of the electrolyte. PCR was found to provide the smaller errors of prediction and models less markedly influenced by the temperature at which spectra were recorded.

NIR spectroscopy with fibre-optic probes and PLSR or PCR is an alternative to gas chromatography and mass spectrometry for the *in situ* analysis of solvent mixtures on account of its high responsiveness, low maintenance costs and the need for no sample treatment. There are references to the analysis of mixtures of ethanol, acetone, acetic acid and water in ethyl acetate;[202] methanol and water in hexane;[202] methanol, ethanol and propanol;[203] and ethanol, propan-1-ol and propan-2-ol in methanol.[204] Martens and Stark[58] demonstrated the significance of the prior mathematical treatment of signals as a means of suppressing multiplicative (pathlength variations) and additive changes (baseline shifts, spectral overlap) encountered in the determination of toluene in mixtures of benzene and xylene; use of a mathematical treatment led to simpler models of increased predictive capacity.

Table 8 Applications of NIR spectroscopy to the determination of active compounds and excipients

Analyte	Sample type	Calibration	Remarks	Ref.
Glycerol, ethanol, phenazone, sodium thiosulfate	Liquid	MLR	Errors less than 3% in major components (glycerol, ethanol and phenazone) and of 5–10% in minor components (lidocaine and sodium thiosulfate)	166
Glucose, fructose, maltose	Syrup		Comparison of NIR, FTIR and HPLC techniques. The accuracy of the spectrophotometric techniques is lower	167
Acetaminophen, codeine phosphate	Syrup	MLR	Comparison of NIR spectroscopy and HPLC in terms of accuracy and throughput. The NIR technique is recommended for components at contents of at least 1%	168
Meprobamate (200 and 400 mg)	Suspension Injection	MLR	Absorbance measurements at 1960 nm of the drug extract in chloroform	169
Cloxacillin benzathine (12.7%)	Cream	MLR	Quantification in creams containing variable proportions of the same excipients, using a single calibration. Errors less than 3.5%	81
Nicotinamide	Solid	MLR	Reproducible results (comparable to those of HPLC) obtained by using two different wavelengths	170
Ceftazidime (77%)	Solid	MLR	Determination of a major active compound with errors of 0–3%. Variability in the raw materials used over a one year period has no effect on the quality of the results	163
Streptomycin sulfate; Cloxacillin benzathine	Solid	MLR	Errors less than 4% in both components that change little on expanding the concentration range used for calibration	84
Ketoprofen (33%)	Solid	MLR	Prior extraction of the drug into chloroform. Errors less than 3.5%	113
Ranitidine hydrochloride	Solid	MLR	Errors less than ±2% and calibration transferability	87
Erythromycin ethylsuccinate (12.9, 19.9 and 34.3%)	Granules	MLR	Errors less than 2.5% with a single calibration for the three presentations	83
Cimetidine (71.8%)	Granules	MLR	The reproducibility of the NIR method (RSD = 0.16%) is comparable to that of a UV method (RSD = 0.15%) and is not influenced by particle size or grain colour	171
Vitamin C (16.7, 22.9 and 40%)	Granules Tablets	MLR PLSR	Errors of 1–2%. The PLSR method provides slightly smaller errors	117
Ranitidine hydrochloride (53.5%)	Tablets	MLR	Errors less than 5%. The precision is not operator-dependent	165
Pirisudanol dimaleate (88%)	Tablets	PLSR	Comparison of recording systems (spinning cuvette and fibre-optic probe). Errors less than 1% and similar in both cases	70
Metoprolol succinate (47.5%)	Tablets	PLSR	Comparison of NIR transmission and diffuse reflectance measurements. The former uses more favourable sample volumes but suffers from spectral noise above 1350 nm	172
Acetylsalicylic acid	Tablets	PCR	Correlation of NIR spectra with the amount of salicylic acid formed by hydrolysis of acetylsalicylic acid. Prediction errors of ±0.04% of the tablet mass	173
Active principle (0.6, 1.2 and 2.4%)	Tablets	PLSR	Automation of an NIR transmission spectroscopic method for determining content uniformity. The reproducibility of the autosampler used was studied	174
Aminodarone hydrochloride ketone (52%)	Tablets	MLR	Errors less than 0.5%. The reproducibility was studied at different temperatures	175
SB 216469-S (1.5, 3 and 6%)	Tablets	PLSR	Quantitative control of the active compound at the different production stages, with no sample pre-treatment	176
Cefuroxime acetyl (66.8%)	Tablets	MLR PLSR	Quantitative control of the active compound at the different production stages. The PLSR method provided lower errors in all cases	86

Conclusions

Recent breakthroughs in analytical instrumentation and available techniques for processing complex signals have fostered the development of new uses of NIR spectroscopy in various industrial fields, prominent among which is the pharmaceutical industry. The large number of references cited in this review testifies to the potential of NIR spectroscopy for qualitative and quantitative analysis of pharmaceutical preparations.

References

1 Herschel, W., *Philos. Trans. R. Soc. London*, 1800, **90**, 255.
2 Herschel, W., *Philos. Trans. R. Soc. London*, 1800, **90**, 284.
3 Herschel, W., *Philos. Trans. R. Soc. London*, 1800, **90**, 293.
4 Goddu, R. F., in *Advances in Analytical Chemistry and Instrumentation*, ed. Reilly, Ch. N., Interscience, New York, 1960, pp. 347–424.
5 Wetzel, D. L., *Anal. Chem.*, 1983, **55**, 1165A.
6 McClure, W. F., *Anal. Chem.*,1994, **66**, 43A.
7 Osborne, B. G., and Fearn, T., *Near Infrared Spectroscopy in Food Analysis*, Longman Scientific & Technical, Harlow, Essex, 1986.
8 Williams, P. C., and Norris, K. H., in *Near-Infrared Technology in the Agricultural and Food Industries*, ed. Williams, P. C., and Norris, K. H., American Association of Cereal Chemists, St. Paul, MN, 1987, ch. 15, pp. 241–246.
9 Ghosh, S., and Rodgers, J., in *Handbook of Near-Infrared Analysis*, Practical Spectroscopy Series, vol.13, ed. Burns, D. A., and Ciurczak, E. W., Marcel Dekker, New York, 1992, vol. 1, ch. 18, pp. 495–526.
10 Hammersley, M. J., in *Handbook of Near-infrared Analysis*, Practical Spectroscopy Series, ed. Burns, D. A., and Ciurczak, E. W., Marcel Dekker Inc., New York, 1992, vol. 13, ch. 17, pp. 475–494.
11 Kradjel, C., and McDermott, L., in *Handbook of Near-Infrared Analysis*, Practical Spectroscopy Series, ed. Burns, D. A., and Ciurczak, E. W., Marcel Dekker, New York, 1992, vol. 13, ch. 21, pp. 565–608.
12 Bunding Lee, K. A., *Appl. Spectrosc. Rev*., 1993, **28**, 231.
13 Workman, J., Jr., *J. Near Infrared Spectrosc.*, 1993, **1**, 221.
14 Workman J., Jr., *Appl. Spectrosc. Rev.*, 1996, **31**, 251.
15 Weyer, L. G., *Appl. Spectrosc. Rev*., 1985, **21**, 1.
16 Lipp, E. D., *Appl. Spectrosc. Rev.*, 1992, **27**, 385.
17 Mark, H., *Anal. Chim. Acta*, 1989, **223**, 75.
18 Michel, I., *NIR News*, 1995, **6**(6), 4.
19 Hermiller, J., *NIR News*, 1996, **7**(2), 4.
20 Hermiller, J., *NIR News*, 1996, **7**(1), 4.
21 Kraemer, E. G., and Lodder, R. A., *Spectroscopy*, 1996, **11**(7), 24.
22 Kraemer, E. G. and Lodder, R. A., *Spectroscopy*, 1996, **11**(8), 17.
23 Ciurczak, E. W., in *Handbook of Near-Infrared Analysis*, Practical Spectroscopy Series, ed. Burns, D. A., and Ciurczak, E. W., Marcel Dekker, New York, 1992, vol. 13, ch. 20, pp. 549–564.
24 Drennen, J. K., and Lodder, R. A., in *Advances in Near-Infrared Measurements*, ed. Patonay, G., Jai Press, Greenwich, CT, 1993, vol. 1, pp. 93–112.
25 Stark, E., Luchter, K., and Margoshes, M., *Appl. Spectrosc. Rev.*, 1986, **22**, 335.
26 Ciurczak, E. W., *Appl. Spectrosc. Rev.*, 1987, **23**, 147.
27 Drennen, J. K., Kraemer, E. G., and Lodder, R. A., *Crit. Rev. Anal. Chem.*, 1991, **22**, 443.
28 Martin, K. A., *Appl. Spectrosc. Rev.*, 1992, **27**, 325.
29 Ciurczak, E. W., and Drennen, J. K., *Spectroscopy*, 1992, **7**(6), 12.
30 Corti, P., Dreassi, E., and Lonardi, S., *Il Farmaco*, 1993, **48**, 3.
31 MacDonald, B. F., and Prebble, K. A., *J. Pharm. Biomed. Anal.*, 1993, **11**, 1077.
32 Morisseau, K. M., and Rhodes, C. T., *Drug Dev. Ind. Pharm.*, 1995, **21**, 1071.
33 Kirsch, J. D., and Drennen, J. K., *Appl. Spectrosc. Rev*., 1995, **30**, 139.
34 McClure, W. F., in *Near-Infrared Technology in the Agricultural and Food Industries*, ed. Williams, P., and Norris, K. H., American Association of Cereal Chemists, St. Paul, MN, 1987, ch. 5, pp. 89–105.
35 Williams, P. C., in *Near-infrared Technology in the Agricultural and Food Industries*, ed. Williams, P., and Norris, K. H., American Association of Cereal Chemists, St. Paul, MN, 1987, ch. 6, pp. 107–142.
36 Kubelka, P., and Munk, F., *Z. Tech. Phys*., 1931, **12**, 593.
37 Lodder, R. A., and Hieftje, G. M., *Appl. Spectrosc.*, 1988, **42**, 556.
38 Springsteen, A., and Ricker, T., *NIR News*, 1996, **7**(4), 6.
39 Springsteen, A., and Ricker, T., *NIR News*, 1996, **7**(5), 12.
40 Osborne, B. G., *Analyst*, 1988, **113**, 263.
41 Griffiths, P. J., *J. Near Infrared Spectrosc*., 1995, **3**, 60.
42 Dahm, D. J., and Dahm, K. D., *J. Near Infrared Spectrosc.*, 1995, **3**, 53.
43 Olinger, J. M., and Griffiths, P. R., in *Handbook of Near-Infrared Analysis*, ed. Burns, D. A., and Ciurczak, E. W., Marcel Dekker, New York, 1992, ch. 3, pp. 13–35.
44 Aucott, L. S., Garthwaite, P. H., and Buckland, S. T., *Analyst*, 1988, **113**, 1849.
45 Salamin, P. A., Cornelis, Y., and Bartels, H., *Chemom. Intell. Lab. Syst.*, 1988, **3**, 329.
46 Wu, W., Walczak, B., Massart, D. L., Prebble, K. A., and Last, I. R., *Anal. Chim. Acta*, 1995, **315**, 243.
47 Hailey, P. A., Doherty, P., Tapsell, P., Oliver, T., and Aldridge, P. K., *J. Pharm. Biomed. Anal*., 1996, **14**, 551.
48 Mobley, P. R., Kowalski, B. R., Workman, J. J., Jr., and Bro, R., *Appl. Spectrosc. Rev*., 1996, **31**, 347.
49 *The Unscrambler User's Guide*, Computer Aided Modelling A/S, Trondheim, Norway, 1993.
50 Downey, G., Robert, P., and Bertrand, D., *Anal. Proc*., 1992, **29**, 8.
51 McClure, W. F., *NIR News*, 1993, **4**(6), 12.
52 McClure, W. F., *NIR News*, 1994, **5**(1), 12.
53 McClure, W. F., *NIR News*, 1994, **5**(4), 14.
54 Isaksson, T., and Næs, T., *Appl. Spectrosc.*, 1988, **42**, 1273.
55 Geladi, P., MacDougall, D., and Martens, H., *Appl. Spectrosc.*, 1985, **39**, 491.
56 Næs, T., Isaksson, T., and Kowalski, B., *Anal. Chem.*, 1990, **62**, 664.
57 Isaksson, T., and Kowalski, B., *Appl. Spectrosc*., 1993, **47**, 702.
58 Martens, H., and Stark, E., *J. Pharm. Biomed. Anal*., 1991, **9**, 625.
59 Karstang, T. V., and Manne, R., *Chemom. Intell. Lab. Syst*., 1992, **14**, 165.
60 Isaksson, T., Wang, Z., and Kowalski, B., *J. Near Infrared Spectrosc.*, 1993, **1**, 85.
61 Barnes, R. J., Dhanoa, M. S., and Lister, S. J., *Appl. Spectrosc.*, 1989, **43**, 772.
62 Dhanoa, M. S., Lister, S. J., and Barnes, R. J., *Appl. Spectrosc.*, 1995, **49**, 765.
63 Dhanoa, M. S., Lister, S. J., Sanderson, R., and Barnes, R. J., *J. Near Infrared Spectrosc.*, 1994, **2**, 43.
64 Jacobsson, S. P., Carlsson, M., Jönsson, U., and Nilsson, G., *J. Pharm. Biomed. Anal*., 1995, **13**, 415.
65 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and de la Pezuela, C., *Appl. Spectrosc*., 1997, **51**, 240.
66 Kaye, W., *Spectrochim. Acta*, 1954, **6**, 257.
67 Sinsheimer, J. E., and Keuhnelian, A. M., *J. Pharm. Sci.*, 1966, **55**, 1240.
68 Massart, D. L., Vandegiste, B. G. M., Deming, S. N., Michotte, Y., and Kaufman, L., *Chemometrics: A Textbook*, Elsevier, Amsterdam, 1988.
69 Workman, J. J., Jr., Mobley, P. R., Kowalski, B. R., and Bro, R., *Appl. Spectrosc. Rev.*, 1996, **31**, 73.
70 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., de la Pezuela, C., and Russo, E., *Anal. Chim. Acta*, 1994, **298**, 183.
71 Brik, H., and van der Vlies, C., *Pharmaeuropa*, 1991, **3**(1), 3.
72 Plugge, W., and van der Vlies, C., *J. Pharm. Biomed. Anal.*, 1993, **11**, 435.
73 Galante, L. J., Brinkley, M. A., Drennen, J. K., and Lodder, R. A., *Anal. Chem.*, 1990, **62**, 2514.
74 Perstorp Analytical, Technical Note 0554.
75 Plugge, W., and van der Vlies, C., *J. Pharm. Biomed. Anal.*, 1992, **10**, 797.
76 González, F., and Pous, R., *J. Pharm. Biomed. Anal.*, 1995, **13**, 419.
77 Mark, H. L., and Tunnell, D., *Anal. Chem.*, 1985, **57**, 1449.
78 Mark, H. L., *Anal. Chem.*, 1987, **59**, 790.
79 Tunnell, D. A., *Anal. Proc.*, 1990, **27**, 59.
80 Corti, P., Dreassi, E., Corbini, G., Lonardi, S., and Gravina, S., *Analusis*, 1990, **18**, 112.
81 Corti, P., Dreassi, E., Corbini, G., Montecchi, L., and Paggi, J., *Analusis*, 1990, **18**, 117.

82 Corti, P., Dreassi, E., Ceramelli, G., Lonardi, S., Viviani, R., and Gravina, S., *Analusis*, 1991, **19**, 198.
83 Corti, P., Savini, L., Dreassi, E., Petriconi, S., Genga, R., Montecchi, L., and Lonardi S., *Proc. Cont. Qual.*, 1992, **2**, 131.
84 Corti, P., Savini, L., Dreassi, E., Ceramelli, G., Montecchi, L., and Lonardi, S., *Pharm. Acta Helv.*, 1992, **67**, 57.
85 Dreassi, E., Ceramelli, G., Corti, P., Lonardi, S., and Perruccio, P. L., *Analyst*, 1995, **120**, 1005.
86 Dreassi, E., Ceramelli, G., Savini, L., Corti, P., Perruccio, P. L., and Lonardi, S., *Analyst*, 1995, **120**, 319.
87 Dreassi, E., Ceramelli, G., Corti, P., Perruccio, P. L., and Lonardi, S., *Analyst*, 1996, **121**, 219.
88 van der Vlies, C., Kaffka, K. J., and Plugge, W., *Pharm. Technol. Eur.*, 1995, **7**(4), 43.
89 Plugge, W., and van der Vlies, C., *J. Pharm. Biomed. Anal.*, 1996, **14**, 891.
90 Martens, H., and Næs, T., *Multivariate Calibration*, Wiley, New York, 1991.
91 Lo, S., and Brown, C. W., *Appl. Spectrosc.*, 1992, **46**, 790.
92 Shah, N. K., and Gemperline, P. J., *Trends Anal. Chem.*, 1989, **8**, 357.
93 Wold, S., and Sjostrom, M., in *Chemometrics: Theory and Applications*, ed. Kowalski, B. R., American Chemical Society, Washington DC, 1977.
94 Gemperline, P. J., Webber, L. D., and Lox, I. O., *Anal. Chem.*, 1989, **61**, 138.
95 Gemperline, P. J., and Boyer, N. R., *Anal. Chem.*, 1995, **67**, 160.
96 Shah, N. K., and Gemperline, P. J., *Anal. Chem.*, 1990, **62**, 465.
97 Dempster, M. A., Jones, J. A., Last, I. R., MacDonald, B. F., and Prebble, K. A., *J. Pharm. Biomed. Anal.*, 1993, **11**, 1087.
98 Dempster, M. A., MacDonald, B. F., Gemperline, P. J., and Boyer, N. R., *Anal. Chim. Acta*, 1995, **310**, 43.
99 Ciurczak, E. W., and Maldacker, T. A., *Spectroscopy*, 1986, **1**(1), 36.
100 Honigs, D. E., Hieftje, G. M., and Hirschfeld, T., *Appl. Spectrosc.*, 1984, **38**, 317.
101 Lodder, R. A., and Hieftje, G. M., *Appl. Spectrosc.*, 1988, **42**, 1351.
102 Lodder, R. A., and Hieftje, G. M., *Appl. Spectrosc.*, 1988, **42**, 1500.
103 Lodder, R. A., Selby, M., and Hieftje, G. M., *Anal. Chem.*, 1987, **59**, 1921.
104 Ciurczak, E. W., *Pharm. Technol.*, 1991, **15**, 140.
105 Wargo, D. J., and Drennen, J. K., *J. Pharm. Biomed. Anal.*, 1996, **14**, 1415.
106 Sekulic, S. S., Ward, H. W., Brannegan, D. R., Stanley, E. D., Evans, C. L., Sciavolino, S. T., Hailey, P. A., and Aldridge, P. K., *Anal. Chem.*, 1996, **68**, 509.
107 Gimet, R., and Luong, A. T., *J. Pharm. Biomed. Anal.*, 1987, **5**, 205.
108 Aldridge, P. K., Evans, C. L., Ward, H. W., Colgan, S. T., Boyer, N., and Gemperline, P. J., *Anal. Chem.*, 1996, **68**, 997.
109 Norris, T., Aldridge, P. K., and Sekulic, S. S., *Analyst*, 1997, **122**, 549.
110 Buchanan, B. R., Ciurczak, E. W., Grunke, A. Q., and Honigs, D. E., *Spectroscopy*, 1988, **3**, 54.
111 MacDonald, S. A., and Hieftje, G. M., *Appl. Spectrosc.*, 1996, **50**, 1161.
112 Halsey, S. A., Technical Note, NIRSystems, Silver Spring, MD, 1994.
113 Corti, P., Dreassi, E., Murratzu, C., Corbini, G., Ballerini, L., and Gravina, S., *Pharm. Acta Helv.*, 1989, **64**, 140.
114 Jouan-Rimbaud, D., Khots, M. S., Massart, D. L., Last, I. R., and Prebble, K. A., *Anal. Chim. Acta*, 1995, **315**, 257.
115 Jouan-Rimbaud, D., Walczak, B., Massart, D. L., Last, I. R., and Prebble, K. A., *Anal. Chim. Acta*, 1995, **304**, 285.
116 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and de la Pezuela, C., *Anal. Chim. Acta*, 1996, **333**, 147.
117 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and de la Pezuela, C., *Talanta*, 1993, **40**, 1671.
118 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and de la Pezuela, C., *Analyst*, 1997, **122**, 761.
119 Buchanan, B., and Honigs, D., *Trends Anal. Chem.*, 1986, **5**, 154.
120 Næs, T., and Isaksson, T., *NIR News*, 1994, **5**(6), 16.
121 Honigs, D. E., Hieftje, G. M., Mark, H. L., and Hirschfeld, T. B., *Anal. Chem.*, 1985, **57**, 2299.
122 Næs, T., *J. Chemom.*, 1987, **1**, 121.
123 Isaksson, T., and Næs, T., *NIR News*, 1995, **6**(1), 13.
124 Puchwein, G., *Anal. Chem.*, 1988, **60**, 569.
125 Mark, H., *Anal. Chem.*, 1987, **59**, 790.
126 Ferré, J., and Rius, F. X., *Anal. Chem.*, 1996, **68**, 1565.
127 Davies, T., *Spectroscopy*, 1996, **8**(4), 27.
128 Fearn, T., Research report No. 105, University College, London, 1992.
129 *Near Infrared Spectral Analysis Software. User's Guide*, NIRSystems, Silver Spring, MD, 1990.
130 Thomas, E. V., and Haaland, D. M., *Anal. Chem.*, 1990, **62**, 1091.
131 Draper, N., and Smith, H., *Applied Regression Analysis*, Wiley, New York, 2nd edn., 1981.
132 Honigs, D. E., Hieftje, G. M., and Hirschfeld, T. B., *Appl. Spectrosc.*, 1983, **38**, 317.
133 Cowe, I. A., McNicol, J. W., and Cuthbertson, D. C., *Analyst*, 1985, **110**, 1233.
134 Næs, T., and Isaksson, T., *NIR News*, 1993,**4**(3), 14.
135 Miller, C. E., *NIR News*, 1993, **4**(6), 3.
136 Long, R. L., Gregoriov, V. G., and Gemperline, P. J., *Anal. Chem.*, 1990, **62**, 1791.
137 Gemperline, P. J., Long, J. R., and Gregoriov, V. G., *Anal. Chem.*, 1991, **63**, 2313.
138 Næs, T., Kvaal, K., Isaksson, T., and Miller, C., *J. Near Infrared Spectrosc.*, 1993, **1**, 1.
139 Walczak, B., and Wegscheimer, W., *Anal. Chim. Acta*, 1993, **283**, 508.
140 Liu, Y., Upadhyaya, B. R., and Naghedolfeizi, M., *Appl. Spectrosc.*, 1993, **47**, 12.
141 Næs, T., Isaksson, T., and Kowalski, B. R., *Anal. Chem.*, 1990, **62**, 664.
142 Næs, T., and Isaksson, T., *Appl. Spectrosc.*, 1992, **46**, 34.
143 Friedman, J. H., and Stuetzle, W., *J. Am. Stat. Assoc.*, 1981, **76**, 817.
144 Beebe, K. R., and Kowalski, B. R., *Anal. Chem.*, 1988, **60**, 2273.
145 Vogt, N. B., *Chemom. Intell. Lab. Syst.*, 1989, **7**, 119.
146 Wold, S., Wold, N. K., and Skagerberg, B., *Chemom. Int. Lab. Syst.*, 1989, **7**, 53.
147 Ciurczak, E. W., Torlini, R. P., and Demkowicz, M. P., *Spectroscopy*, 1986, **1**(7), 36.
148 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., González, F., and Pous, R., in *Near Infra-red Spectroscopy (Bridging Gap between Data Analysis and NIR Applications*), ed. Hildrum, K. I., Isaksson, T., Næs, T., and Tandberg, A., Ellis Horwood, Chichester, 1992, ch. 64, pp. 401–406.
149 Ilari, J. L., Martens, H. and Isaksson, T., *Appl. Spectrosc.*, 1988, **42**, 722.
150 Kirsch, J. D., and Drennen, J. K., *J. Pharm. Biomed. Anal.*, 1995, **13**, 1273.
151 Langkilde, F. W., and Svantesson, A., *J. Pharm. Biomed. Anal.*, 1995, **13**, 409.
152 Lin, L., and Brown, C. W., *Appl. Spectrosc.*, 1992, **46**, 1809.
153 Delwiche, S. R., Norris, K. H., and Pitt, R. E., *Appl. Spectrosc.*, 1992, **46**, 782.
154 Lin, J., and Brown, C. W., *Appl. Spectrosc.*, 1993, **47**, 62.
155 Sinsheimer, J. E., and Poswalk, N. M., *J. Pharm. Sci.*, 1968, **57**, 2007.
156 Issa, R. M., El-Marsafy, K. M., and Gohar, M. M., *Ann. Quím.*, 1988, **84**, 312.
157 Böhme, W., Liekmeier, W., Horn, K., and Wilhelm, C., *Labor Praxis*, 1989, January/February, 44.
158 Warren, R. J., Zarembo, J. E., Chong, C. W., and Robinson, M. J., *J. Pharm. Sci.*, 1970, **59**, 109.
159 Vomhof, D. W., and Thomas, J. H., *Anal. Chem.*, 1970, **42**, 1230.
160 Ludvik, J., Hilgard, S., and Volke, J., *Analyst*, 1988, **113**, 1729.
161 Garrigues, S., Gallignani, M., and de la Guardia, M., *Anal. Chim. Acta*, 1993, **281**, 259.
162 Last, I. R., and Prebble, K. A., *J. Pharm. Biomed. Anal.*, 1993, **11**, 1071.
163 Lonardi, S., Viviani, R., Mosconi, L., Bernuzzi, M., Corti, P., Dreassi, E., Murratzu, C., and Corbini, G., *J. Pharm. Biomed. Anal.*, 1989, **7**, 303.
164 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and Rovira, E., *J. Pharm. Biomed. Anal.*, 1997, **16**, 255.
165 Corti, P., Dreassi, E., Corbini, G., Lonardi, S., Viviani, R., Mosconi, L., and Bernuzzi, M., *Pharm. Acta Helv.*, 1990, **65**, 28.

166 Dubois, P., Martínez, J. R., and Levillain, P., *Analyst*, 1987, **112**, 1675.
167 Meurens, M., Dupuy, N., Mirouze, F. D., Legrand, P., and Huvenne, P. J., in *Near Infra-red Spectroscopy (Bridging Gap between Data Analysis and NIR Applications*), ed. Hildrum, K. I., Isaksson, T., Næs, T., and Tandberg, A., Ellis Horwood, Chichester, 1992, ch. 7, pp. 47–49.
168 Ciurczak, E. W., and Torlini, R. P., *Spectroscopy*, 1987, **2**(3), 41.
169 Zappala, A. F., and Post, A., *J. Pharm. Sci.*, 1977, **66**, 292.
170 Osborne, B. G., *Analyst*, 1987, **112**, 313.
171 Chasseur, J. C., *Chim. Oggi*, 1987, **6**, 21.
172 Gottfries, J., Depui, H., Fransson, M., Jogeneelen, M., Josefson, M., Langkilde, F. W., and Witte, D. T., *J. Pharm. Biomed. Anal.*, 1996, **14**, 1495.
173 Drennen, J. K., and Lodder, R. A., *J. Pharm. Sci.*, 1990, **79**, 622.
174 Jönsson, U., Eriksson, E. H., and Sellberg, B., presented at the International Conference of Pharmaceutical Applications of NIR Spectroscopy, Stockholm, Sweden, 1996.
175 Jensen, R., Peuchant, E., Castagne, I., Boirac, A. M., and Roux, G., *Spectrosc. Int. J.*, 1988, **6**, 63.
176 Han, S. M., and Faulkner, P. G., *J. Pharm. Biomed. Anal.*, 1996, **14**, 1681.
177 Dean, T., and Isaksson, T., *NIR News*, 1993, **4**(2), 8.
178 Dean, T., and Isaksson, T., *NIR News*, 1993, **4**(4), 14.
179 Shenk, J., and Westerhaus, M. O., *NIR News*, 1993, **4**(5), 13.
180 Bouveresse, E., and Massart, D. L., *Vib. Spectrosc.*, 1996, **11**, 3.
181 Shenk, J., and Westerhaus, M. O., *US Pat.* 4 866 644, 1989.
182 Shenk, J., and Westerhaus, M. O., *Crop Sci.*, 1991, **31**, 1694.
183 Dardenne, P., Biston, R., and Simmaeve, G., in *Near Infra-red Spectroscopy (Bridging Gap between Data Analysis and NIR Applications*), ed. Hildrum, K. I., Isaksson, T., Næs, T., and Tandberg, A., Ellis Horwood, Chichester, 1992, ch. 72, pp. 453–458.
184 Bouveresse, E., Massart, D. L., and Dardenne, P., *Anal. Chim. Acta*, 1994, **297**, 405.
185 Bouveresse, E., and Massart, D. L., *Anal. Chem.*, 1995, **67**, 1381.
186 Mark, H., and Workman, J. J., Jr., *Spectroscopy*, 1988, **3**(11), 28.
187 Wang, Y., Veltkamp, D. J., and Kowalski, B. R., *Anal. Chem.*, 1991, **63**, 2750.
188 Wang, Y., and Kowalski, B. R., *Anal. Chem.*, 1993, **65**, 1301.
189 Wang, Y., and Kowalski, B. R., *Appl. Spectrosc.*, 1992, **46**, 764.
190 Bouveresse, E., Hartmann, C., Massart, D. L., Last, I. R., and Prebble, K. A., *Anal. Chem.*, 1996, **68**, 982.
191 Jones, J. A., Last, I. R., MacDonald, B. F., and Prebble, K. A., *J. Pharm. Biomed. Anal*., 1993, **11**, 1227.
192 Forina, M., Drava, G., Armanino, C., Boggia, R., Lanteri, S., Leardi, R., Corti, P., Conti, P., Giangiacomo, R., Galliena, C., Bigoni, R., Quartari, I., Serra, C., Ferri, D., Leoni, O., and Lazzeri, L., *Chemom. Intell. Lab. Syst*., 1995, **27**, 189.
193 Forina, M., Armanino, C., and Giangiacomo, R., in *Near Infra-red Spectroscopy (Bridging Gap between Data Analysis and NIR Applications*), ed. Hildrum, K. I., Isaksson, T., Næs, T. and Tandberg, A., Ellis Horwood, Chichester, 1992, ch. 14, pp. 91–96.
194 Blank, T. B., Sun, S. T., Brown, S. D., and Monfre, S. L., *Anal. Chem*., 1995, **68**, 2987.
195 Reeve, J. B., III, *Appl. Spectrosc.*, 1995, **49**, 181.
196 Reeve, J. B., III, *Appl. Spectrosc.*, 1995, **49**, 295.
197 Reeve, J. B., III, *J. AOAC Int.*, 1996, **76**,741.
198 Reeve, J. B., III, *Appl. Spectrosc*., 1996, **50**, 154.
199 Lin, J., and Brown, C. W., *Appl. Spectrosc.*, 1993, **47**, 239.
200 Lin, J., and Brown, C. W., *Anal. Chem.*, 1993, **65**, 287.
201 Lin, J., and Brown, C. W., *J. Near Infrared Spectrosc.*, 1993, **1**, 109.
202 Application note A3-987, Guided Wave.
203 Martens, H., Næs, T., and Bjorsvik, H. R., *Wave Guide*, 1988, **1**(1), 4.
204 Mackison, R., Brinkworth, S. J., Belchamber, R. M., Aries, R. E., Cutler, D. J., Deeley, C., and Mould, H. M., *Appl. Spectrosc*., 1992, **46**, 1020.

Paper 8/02531B
Received April 2, 1998
Accepted June 9, 1998