Available online at www.sciencedirect.com
ScienceDirect
Journal of Pharmaceutical and Biomedical Analysis 42 (2006) 517–522
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Short communication

Augmentation of near infrared diffuse reflectance and transmittance spectral data for the development of robust PLSBC models for classifying double blind clinical trial tablets

T. Van den Kerkhof*, R. De Maesschalck, K. Vanhoutte, M.C. Coene

Janssen Pharmaceutica NV, Pharmaceutical Research and Development, Turnhoutseweg 30, B-2340 Beerse, Belgium

Received 22 November 2005; accepted 8 May 2006
Available online 23 June 2006

Abstract

The water content of clinical trial tablets can be different between and within different tablet batches, depending on the relative humidity conditions during their production, packaging, storage and analysis. These water variations lead to important spectral variations in the near infrared spectral region which can lead to a wrong identification if the classification model was based on unrepresentative data towards the water content. As model development for clinical trial studies needs to be extremely fast – within one working day – with generally only one batch available, the principle of data augmentation has to be applied to render more robust classification models. Therefore, tablets available for constructing the model are being processed in order to increase or decrease their water content and to make them more representative for tablets to be tested in the future. The inclusion of a deliberate water variation is the most efficient way to develop a model, for which no additional model redevelopment will be required to pass the system suitability tests and to obtain a correct identification.

Keywords: Clinical trial studies; NIR; PLSBC; Data augmentation

1. Introduction

Near infrared spectroscopy (NIR) in transmission or reflectance mode has been successfully applied in the pharmaceutical industry for several applications such as e.g. the end-point determination of a hydrogenation reaction [2] or a drying process [3,4], blend homogeneity [5–7], content uniformity [8,9] and polymorph determination [10,11]. In a previous article [1] we described the PLSBC approach for using NIR spectroscopy as identification test of blister packed double blind clinical trial tablets. The largest benefit of NIR spectroscopy for this application is its speed. The classical way for identifying the tablets is to apply UV spectroscopy or liquid chromatography, which usually requires dissolving and filtering the sample. This procedure is rather time-consuming since, including sample preparation, it can take up to 30 minutes cycle time. NIR measurements need no sample preparation. One measurement can be performed in seconds up to minutes depending on the type of NIR instrument and its settings. To make a fair comparison, also the time invested into developing and validating the method must be taken into account. For the identification with UV spectroscopy or liquid chromatography, no special method has to be developed, since existing methods, e.g. content uniformity, can be applied. For NIR spectroscopy, a new method needs to be developed and validated. Concerning cycle time, NIR becomes more efficient as the number of samples to be measured increases. For double blind clinical trial tablets, the number of samples to be tested for the clinical trial route of one particular drug generally ranges from 800 to 12,000 tablets. This is still a relatively small amount of samples and therefore it is important to limit the time for method development and validation. In that way, the result of using NIR instead of UV spectroscopy or liquid chromatography can be a substantial gain in time or on the other hand an improvement of quality as more tablets can be tested with the same resources.

For NIR spectroscopy, identification methods can be developed by applying the PLSBC approach, which is based on partial least squares (PLS) regression combined with β-error driven class boundaries. The approach was shown to be easy to develop

* Corresponding author. Tel.: + 32 14 60 37 03; fax: +32 14 60 50 34.
E-mail address: tvdkerkh@prdbe.jnj.com (T. Van den Kerkhof).

0731-7085/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2006.05.007

and validate [1]. A typical method can be developed and validated within one working day, which is very fast. Next to the speed of the development, it is even more important that the method is robust, meaning that the risk for misclassification of samples is virtually zero. The PLSBC approach was developed in order to limit the probability for misclassification as much as possible by constructing the class boundaries in function of the β-error. The probability of misclassification however does not only depend on the type of classification model and the class boundaries that are used. Another parameter which has a very large impact on the robustness of the classification, is the representativeness of the data applied for building the classification model (training set).

To obtain a robust classification model for the identification of clinical trial tablets, NIR spectra representing all tablets to be tested in the future must be available. At the time the model has to be developed, generally only one batch is available for constructing the classification model. As a result, the model may not be robust towards additional variation which may be present in newly manufactured tablet batches. In this paper, the effect of relatively small changes between batches of the same formulation towards misclassification is investigated over a large time frame and several batches. It is also shown how one can obtain robust classification models using only a limited number of tablets of one batch.

2. Experimental

2.1. Samples

The samples of two clinical studies, for investigating oral immediate release tablets of galantamine and topiramate, were studied. The galantamine tablets used in this work are coated circular tablets with a total weight of 250 mg and were manufactured to contain 4, 8 and 12 mg of galantamine, or to be a matching placebo. These tablets were measured in transmission mode. In order to introduce additional variation between tablets of one batch, several tablets are being processed by storing them in a different environment before measuring. To increase the water content of the tablets, they were stored at 75% relative humidity for 3 h. Storage at 50 °C for 30 min was used to decrease their water content. These storage conditions were chosen in such a way that the total weight difference of the tablets between the conditions was approximately 5 mg or 2% (w/w).

The topiramate tablets are white layered coated oblong tablets with a total weight of 319 mg and were manufactured to contain 25 mg of topiramate, or to be a matching placebo. These tablets were measured in the diffuse reflection mode. Different water contents were achieved in the same way as for the galantamine tablets. The weight difference between the two conditions was 5 mg or 1.6% (w/w).

2.2. Instrumentation and software

The FT-NIR spectra were recorded on a Bruker Vector 22/N-T spectrometer, using a Tungsten Halogen source in combination with a Quartz beamsplitter and an InGaAs detector. The interferograms were recorded with a resolution of 8 cm^{-1}, averaged over 16 (reflection) or 64 (transmittance) scans, Blackman-Harris 3-Term apodized and Fourier transformed with a zero filling factor of 2.

Calibration models were built using the Quant 2 software package version 4.2, an add-on to the general OPUS spectroscopy software version 4.2. Quant 2 contains the PLS regression technique needed for the PLSBC approach. Setting the class boundaries was performed by using a validated Excel sheet (Microsoft).

3. Discussion and results

The sources of variation within NIR spectra of tablets of the same theoretical composition are physical parameters such as the particle size, particle morphology and density (tablet compression force). These parameters generally can be eliminated from the spectra by applying a suitable spectral pre-treatment method such as e.g. standard normal variate (SNV) [12] or first and second derivative [13]. The excipients in the tablet or the coating can vary slightly between different batches or samples (excipient homogeneity). Although mostly not relevant in vivo, these factors can have an influence on the NIR spectra. In practice, the use of a minimal amount of tablets of one batch is enough to include this variation into the model since the manufacturing of batches is very consistent. To assure that no significant additional variation is included in the spectra of newly produced batches, a system suitability test (SST) is performed. The SST includes the prediction of a few tablets of each new batch with a known identity by using the model. If misclassification should occur, the model must be redeveloped by including a suitable number of NIR spectra of the new batches and be revalidated. As will be shown in this paper, this situation almost never occurs in practice.

A much more important parameter which is often adding a lot of additional variation and which has not been fully covered by the calibration tablets, is the water content of the tablets. For most tablets, water can be taken up very fast depending on the relative humidity of the environment where the NIR measurements are performed. Fig. 1 shows the increase and decrease of water in a topiramate and galantamine tablet as a function of the relative humidity. By storing these tablets in a high relative humidity environment, a substantial increase of the water content can be achieved in a very short period. Even in blisters, the tablets can contain a different amount of water due to the fact that mostly in the packaging facilities often no or only limited humidity control is used. Small amounts of water present in the tablets lead to very strong and broad absorbance bands at 7300–7000 (first overtone of O–H stretch) and 5300–5000 cm^{-1} (combination band of O–H stretch and O–H deformation) in their NIR spectra, which is demonstrated by the topiramate tablets in Fig. 2. First derivative spectra are calculated in order to remove the baseline offsets and to make the spectral regions, which undergo a change, more visible. The spectral variations due to the drug and water content of the topiramate tablets are shown in Fig. 3. Next to the water variations, spectral changes due to the pres-