# Evaluation of a near infrared procedure for qualification of a pharmaceutical preparation

**S. Maspoch, M. Blanco, J. Coello, H. Iturriaga and N. Pou**

*Departament de Química, Unitat de Química Analítica, Facultat de Ciències Cn, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona) Spain*

## Introduction

The application of near infrared (NIR) spectroscopy to the identification of raw materials and final products in pharmaceutical industry is, nowadays, a well-established and known analytical method.[1] Identification (defined as the determination of a certain chemical entity) of a pure compound doesn't involve any conceptual problem. Construction of NIR libraries containing the spectra of all the compounds assayed in the identification process, validation and use of the libraries has become routine analysis. However, extension of the identification concept to a blended product (mixture of an active compound and several excipients) is much more complex, since not only the chemical identity has to be proved, but also the relative proportions between components. This "extended identification" is usually referred as "qualification" (assessment of origin, or qualities, or chemical and physical specifications)[2] or also as "pharmaceutical identification".[3]

The aim of this work is to study the influence on the selectivity of a library of small changes in the composition of a pharmaceutical preparation due to variations on the concentration of the active compound or the excipients. To increase the selectivity, a two steps qualification process is proposed using sub-libraries.[4] Besides the general library, a sub-library is built including only all those mutually related substances that result in ambiguous identification with the general library. The general library, containing spectra of the active compound, the excipients and production samples of the pharmaceutical preparation, has been built using correlation coefficient as measure of the similarity. After finding the appropriate threshold value in the validation process of the library, the selectivity of the qualification process to small changes in the concentration of the active compound and/or the excipients has been tested.

To increase the selectivity of the qualification, a sub-library containing only spectra of production samples has been developed and several procedures for the qualification (correlation coefficient,[5] maximum distance in wavelength space[6] and residual variance on principal components space[7]) have been evaluated and its results compared.

Also, the effect on the qualification of the contamination of the active compound (an enantiomerically pure form) with racemic form has been evaluated.

## Experimental

### Instruments and software

NIR spectra were recorded on a NIRSystems 6500 spectrophotometer (Foss) equipped with an AP6641ANO4P fibre-optic module. The system was governed by Vision v.2.22 software, which includes routines for acquisition and processing of spectra.




Near Infrared Spectroscopy: Proceedings of the 11th International Conference.

www.nirpublications.com

The experimental set up also included a Turmix Mill blade grinder (Barcelona, Spain), a Turbula T2C shaker mixer from WAB (Basel, Switzerland) to homogenize laboratory-made samples, and a CISA (Barcelona, Spain) 250 μm sieve.

### Production samples

The pharmaceutical preparation, presented as coated tablets, has a nominal content of 14.1% on the active compound (which is in the enantiomerically pure D form). The excipients have a concentration of (1): 53% (2): 18.5% (3): 10.4% (4): 2%. There is also a 2% of lacquer. 30 samples from different production batches were available.

### Laboratory made samples

The selectivity of the identification method to changes in the concentration of the components was studied by using laboratory made samples obtained by overdosing or underdosing samples from different production batches, by addition of the active compound or excipients, respectively, to an accurately known amount of a previously ground production sample. Nominal concentration of the active compound was varied +5, +10, +15, +20 and +25 %. Nominal concentration of the excipients (1), (2) and (3) was varied +10 and +20 %.

The influence of a contamination of the enantiomerically pure active compound by the racemic form was evaluated by using 19 samples prepared with equal composition to normal production samples, 4 of them with the active compound contaminated with 10, 15, 25 and 50 % of the racemic form.

All laboratory samples were homogenized in a solid shaker mixer for 30 min and then a first NIR spectrum was recorded. NIR spectra were recorded after 10 min of additional shaking. When two consecutive spectra overlapped each other on the computer screen, the sample was considered homogeneous; otherwise, the mixing process was repeated until two identical spectra were obtained.

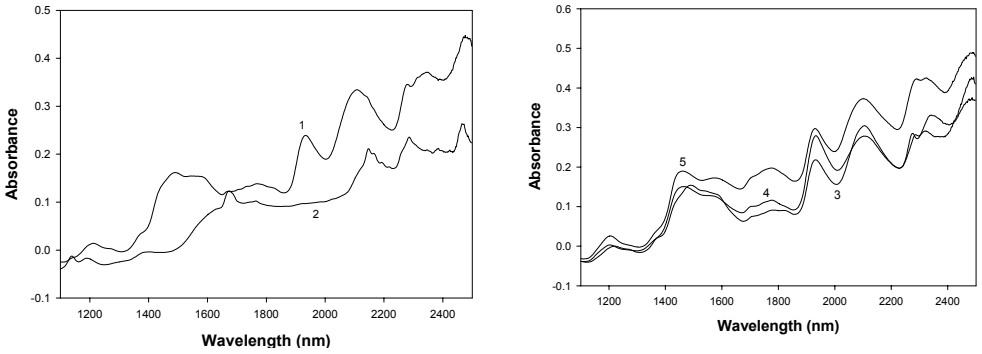


**Figure 1. NIR spectra for (1) Production sample, (2) Active compound; (3) Excipient 1; (4) Excipient 2; (5) Excipient 3.**

### Recording of NIR spectra

All production samples were ground prior to recording of their NIR spectra (particle size < 250 μm). Measurements were made in the diffuse reflectance mode, using a fibre optic probe directly inserted into de powder. The spectrum of each sample was recorded in triplicate over the

wavelength range 1100-2498 nm, using an average of 32 scans. The sample was turned over among the three consecutive measurements. All analyses were performed by using the average spectra.

Figure 1 shows the NIR spectrum for the active compound, a production sample and the main excipients.

### Processing of data

Libraries were constructed from second-derivative spectra in order to facilitate discrimination among different products and decrease spectral variability due to scattering. Vision software was used to calculate the derivatives with a segment size of 10 and a gap size of 0. Wavelength range was 1130-2200 nm.

## Results and discussion

### Identification method for production samples

In order to include the variability in the manufacturing process, a principal component analysis was perform to select, from the 30 production samples available, those 10 that presented maximum variability in the scores plot, which were used to develop the library; the other 20 were used for validation.

A library, containing those 10 spectra for the pharmaceutical preparation, 6 for the active compound and 6 for each of the main excipients, was initially compiled. The library thus obtained exhibited no internal conflicts and correctly identified every spectrum used for self-validation.

**Table 1. Correlation coefficients obtained in the identification of a production sample.**

| | **Product included in the library** | | | | |
|---|---|---|---|---|---|
| **Identified as** | Production | Active Comp. | Excipient (1) | Excipient (2) | Excipient (3) |
| Production | 0.999 | 0.213 | 0.931 | 0.875 | 0.873 |
| Active Comp. | 0.223 | 1.000 | –0.091 | –0.074 | –0.090 |
| Excipient (1) | 0.925 | –0.094 | 1.000 | 0.911 | 0.896 |
| Excipient (2) | 0.866 | –0.072 | 0.913 | 1.000 | 0.943 |
| Excipient (3) | 0.869 | –0.093 | 0.898 | 0.943 | 1.000 |

Table 1 shows the correlation coefficient obtained in the identification of a production sample and their components.

Correlation coefficients are different enough to avoid any possible confusion between the spectrum of each compound included in the library.

The repeatability of the method was tested analysing 12 times a single sample, the same day and by the same operator. Intermediate precision was also tested, with two production samples, three different days and by two different operators. In all cases, the samples were correctly identified, with correlation coefficient between 0.999 and 1.000. Positive identification was assigned to all correlation coefficients above 0.99; in the even of any sample possessing a coefficient above that threshold for two products, it was identified with that exhibiting the highest coefficient.

### Selectivity of the identification to changes in the concentration of the components

Figure 2 shows the difference of 2$^{nd}$ derivative NIR spectra of regular production samples and samples doped with active compound and the major excipient. All the anomalous samples were analysed using the library constructed before (correlation coefficient method) and all of them were incorrectly identified as normal production samples.

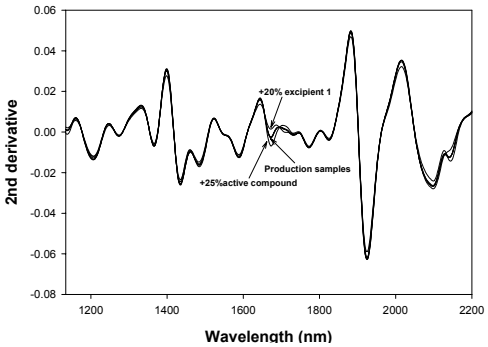


**Figure 2. 2nd derivative spectra of regular production samples and doped with +25% of active compound and +20% of the major excipient.**

This proves that the correlation coefficient is very suitable for constructing libraries as it has the advantage that it has low sensitivity to changes in signal values, so it is scarcely sensitive to slight instrument or composition variations. But this is inconvenient when changes in the concentration make little signal variation. An improvement of the selectivity of the library is needed to achieve a correct qualification of a production sample.

## Increase of the selectivity of the library

To increase the selectivity of a library it is possible to develop sub-libraries to perform a qualification in several steps.[4] In the first step, a general library is constructed were the discrimination between the pharmaceutical compound and their pure components is achieved. In a second step, a sub-library including all those mutually related substances that result in ambiguous identification with the general library is also constructed, not necessary using the same discrimination method that the first one, validated and used to discriminate those samples that couldn't be distinguished before.

So, a two steps qualification process with two libraries was developed. First, the discrimination between the pharmaceutical compound and their pure components was achieved with the first library and, later, a second method was applied to qualify samples with a composition different to the nominal values.

**Table 2. Values obtained in the analysis of samples with different composition from the nominal. Residual variance in principal components space and maximum distance in wavelength space methods were used. The variation of the concentration over the nominal value (in %) is indicated. Bold values indicate samples incorrectly qualified.**

| Variation of the concentration (%) | | | | Residual | Maximum |
|---|---|---|---|---|---|
| Act. Comp. | Exc. (1) | Exc. (2) | Exc. (3) | Variance | Distance |
| +5.0 | –0.8 | –0.8 | –0.8 | 0.918 | 9.46 |
| +10.0 | –1.6 | –1.6 | –1.6 | 0.958 | 10.32 |
| +15.0 | –2.4 | –2.4 | –2.4 | 0.971 | 8.88 |
| +20.0 | –3.2 | –3.2 | –3.2 | 0.978 | 10.60 |
| +25.0 | –4.2 | –4.1 | –4.2 | 0.995 | 14.41 |
| –11.3 | +10.0 | –11.3 | –11.3 | 0.955 | 7.91 |
| –22.6 | +20.0 | –22.6 | –22.6 | 0.995 | 16.88 |
| –2.3 | –2.3 | +10.0 | –2.3 | **0.820** | **4.17** |
| –4.5 | –4.5 | +20.0 | –4.5 | 0.912 | 5.68 |
| –1.2 | –1.2 | –1.2 | +10.0 | **0.818** | **3.79** |
| –2.3 | –2.3 | –2.3 | +20.0 | **0.835** | 4.46 |

The sub-library contained the spectra of the ten production samples previously chosen. Maximum distance in wavelength space and residual variance in principal components space were

assayed for the development of the second qualification step. Threshold values were obtained from validation samples and repeatability and intermediate precision studies.

Using the maximum distance in the wavelength space and the spectral ranges of 1130–1450 and 1600–2200 nm (to avoid Wood's anomaly) a threshold of 4.2 was obtained. With a model of 2 principal components, threshold for the residual variance was found to be 0.90.

Table 2 presents the composition variation of the doped samples and the results obtained in their qualification process by both methods. Maximum distance in the wavelength space produced slightly better results with only two samples incorrectly qualified, one of them just in the threshold. It has to be remarked that in these samples, the concentration of the active compound and the main excipient changed less than 3% of the nominal values, showing the high selectivity achieved.

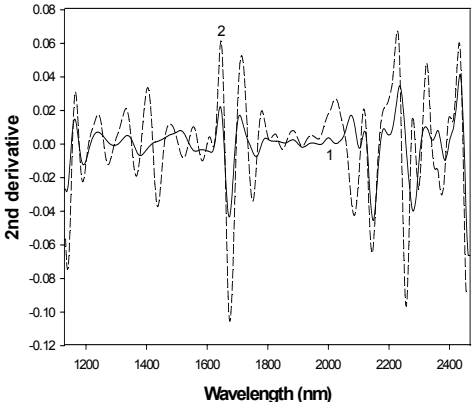


**Figure 3. 2nd derivative spectra of (1) active compound (D form) and (2) active compound (racemic form).**

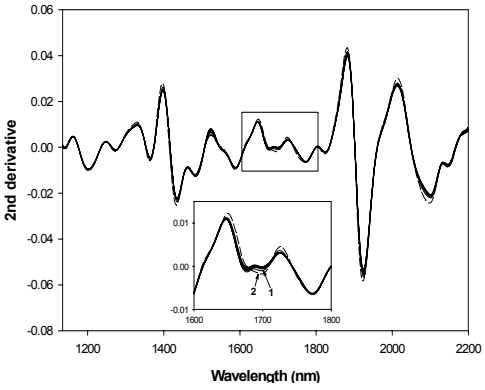


**Figure 4. 2nd derivative spectra of laboratory-made samples. Active compound impurified with (1) 25% of racemic form; (2) 50% of racemic form.**

## Effect of a contamination with racemic form

The active compound is an enantiomerically pure form, with a different NIR spectrum than that of the racemic form. Figure 3 shows both spectra. So, it was attempted to distinguish samples with the active compound contaminated with the racemic form. The presence of the racemic form has a slight effect in the NIR spectrum (Figure 4).

**Table 3. Qualification of the nine samples assayed.**

| | | Positive | Negative |
|---|---|---|---|
| **n** | **Samples** | **Qualification** | **Qualification** |
| 5 | Normal | 5 | — |
| 1 | 10% racemic | 1 | — |
| 1 | 15% racemic | — | 1 |
| 1 | 25% racemic | — | 1 |
| 1 | 50% racemic | — | 1 |

A library containing 10 of the 15 laboratory-made normal samples was built. The other five were used for validation and to establish the threshold values. Best results were obtained with the method of maximum distance in wavelength space, with a threshold of 4.3. This method allows to distinguish samples contaminated with more than 10 % of the racemic form, as it is shown in table 3.

## Conclusion

Identification libraries constructed by using correlation coefficient are very easy to build since they need a relative few number of samples and are very efficient in the identification of a chemical composition. However, they are not able to distinguish changes in the concentrations of the components of a preparation.

The qualification in two steps can increase the selectivity of the identification process. In the first step, clearly different products are identified using correlation coefficient method and, in the second one, the more similar compounds may be differentiated in a specially built library that contains all the expected variation in a normal production process, so a careful selection of all available samples is needed. A more powerful discrimination method is also needed.

In the case studied in this work, maximum distance in the wavelength space and residual variance in principal components space produce similar results, slightly better for the former. Variations in the active compound concentration higher than 3% and also a contamination by the racemic form above 10% can be detected.

## Acknowledgement

The authors are grateful to the DGIGyT. Spain, for funding this work as a part of the projects BQU2000-0234 and BQU2001-2019. N. Pou acknowledges additional support from the DGU of the Generalitat de Catalunya in the form of an FPI scholarship.

## References

1. European Pharmacopoeia, Council of Europe, Strasbourg, France, 3rd Edn, (1997).
2. J.M. Chalmers and R. Griffiths (Eds), *Handbook of Vibrational Spectroscopy*. John Wiley & Sons, Chichester, UK (2002).
3. The European Agency for the Evaluation of Medicinal Products. CPMP/ QWP/ 3309/ 01 draft, November (2001).
4. M. Blanco and M.A. Romero, *Analyst* **126,** 2212 (2001).
5. J.J. Workman, Jr, P.R. Mobley, B.R.Kowalski and R. Bro, *Appl. Spectrosc. Rev,* **31,** 73 (1996).
6. J. Gemperline and N.R. Boyer, *Anal. Chem,* **67**, 160 (1995).
7. *VISION User Manual*, Foss NIRSystems, Silver Spring, MD, (1998).