Journal of Pharmaceutical Analysis 6 (2016) 64–69

HOSTED BY

Contents lists available at ScienceDirect

Journal of Pharmaceutical Analysis

journal homepage: www.elsevier.com/locate/jpa
www.sciencedirect.com

Short Communication

Optimization of wavelength range and data interval in chemometric analysis of complex pharmaceutical mixtures ☆

Michele De Luca, Giuseppina Ioele, Claudia Spatari, Gaetano Ragno *

Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, 87036 Rende, Italy

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form
28 September 2015
Accepted 9 October 2015
Available online 22 October 2015

Keywords:
Chemometrics
Spectrophotometry
Principal component analysis
Partial least squares
Multivariate curve resolution

ABSTRACT

The performance of different chemometric approaches was evaluated in the spectrophotometric determination of pharmaceutical mixtures characterized by having the amount of components with a very high ratio. Principal component regression (PCR), partial least squares with one dependent variable (PLS1) or multi-dependent variables (PLS2), and multivariate curve resolution (MCR) were applied to the spectral data of a ternary mixture containing paracetamol, sodium ascorbate and chlorpheniramine (150:140:1, m/m/m), and a quaternary mixture containing paracetamol, caffeine, phenylephrine and chlorpheniramine (125:6. 25:1.25:1, m/m/m/m). The UV spectra of the calibration samples in the range of 200–320 nm were pre-treated by removing noise and useless data, and the wavelength regions having the most useful analytical information were selected using the regression coefficients calculated in the multivariate modeling. All the defined chemometric models were validated on external sample sets and then applied to commercial pharmaceutical formulations. Different data intervals, fixed at 0.5, 1.0, and 2.0 point/nm, were tested to optimize the prediction ability of the models. The best results were obtained using the PLS1calibration models and the quantification of the species of a lower amount was significantly improved by adopting 0.5 data interval, which showed accuracy between 94.24% and 107.76%.

1. Introduction

Spectrophotometric analytical techniques are widely used in the pharmaceuticals and food quality controls during the batch production or stability controls. This choice is justified by the simplicity of sample preparation and execution as well as by the short analysis time and relatively lower cost than other analytical techniques [1,2]. However, the techniques based on ordinary spectrophotometry are affected by low resolution and are often unsatisfactory in the analysis of complex mixtures [3–6]. Several pharmaceuticals are multicomponent mixtures and often are difficult to be analyzed because of overlapping signals or the presence of components in much lower concentration than the others. In recent years, the advent of computerized instrumentation coupled to multivariate analysis techniques has allowed to increase the potential of the spectrophotometric analysis with the ability to simultaneously process a large number of spectral data recorded in turn by a high number of samples [7,8]. Analysis of complex pharmaceutical mixtures by applying different chemometric procedures on spectral data has been reported in many papers [9–13]. Multivariate curve resolution-alternating least squares (MCR–ALS) has been applied to the study of complex mixtures to resolve different components in pharmaceutical formulation [14–19].

In building a calibration sample set, an appropriate design of experiments (DOE) can affect the prediction ability of the multivariate models. In the present work, a simple latex design (SLD) distributed on five concentration levels was applied in order to select sets of reference mixtures covering the entire experimental domain corresponding to the content of the commercial pharmaceutical specialties [8]. Moreover, in the chemometric treatment of complex data sets, it is usually preferable to reduce the data in order to select those that carry useful analytical information and at the same time minimize those that carry redundant or useless information. In many cases, the choice of the most useful data influences the predictive ability of the multivariate models and this procedure can be very useful in the determination of the components at very low concentration that are often hidden by the more concentrated components [8].

In the first step of a multivariate regression method, principal component analysis (PCA) identifies orthogonal directions of maximum variance of the original data, and places the data in a space of lower dimensionality made from the components that have the highest variance. PCA combines the original variables into

☆Peer review under responsibility of Xi'an Jiaotong University.
* Corresponding author.
E-mail address: gaetano.ragno@unical.it (G. Ragno).

http://dx.doi.org/10.1016/j.jpha.2015.10.001

a smaller number of orthogonal variables called principal components (PCs). The first PCs are considered in the modeling because of containing the most useful information, whereas the last ones can be discarded [20–23].

The principal component regression (PCR), partial least square 1 (PLS1), partial lest squares 2 (PLS2) and multivariate component analysis (MCR) models were applied to two pharmaceutical formulations, the first one containing three active pharmaceutical ingredients (APIs) and the other four APIs, which are very difficult to be analyzed by means of conventional spectrophotometric methods for the presence of some components in quantities much lower than the others. The increase in the predictive power of the models was studied by varying the instrumental parameter "data interval" between 0.5 and 2 nm in the recording of the spectra used in calibration.

The ternary formulation consisted of paracetamol (PAR), sodium ascorbate (ASC) and chlorpheniramine maleate (CHL), with a ratio of 150:140:1 (m/m/m). The quaternary mixture contained paracetamol (PAR), caffeine (CAF), phenylephrine hydrochloride (PHE) and chlorpheniramine maleate (CHL) with a ratio of 125:6.25:1.25:1 (m/m/m/m). Both mixtures are commonly used as analgesics and antipyretic specialties.

The analytical performance of the applied algorithms was tested on the data matrices from synthetic mixtures and commercial pharmaceutical preparations.

2. Materials and methods

2.1. Chemicals

The active pharmaceutical ingredients ASC (98%), CAF (99%), CHL (99%), PAR (100%), and PHE (98%) were purchased from Sigma-Aldrich (Milan, Italy) and used as received. The pharmaceutical specialties Dequa-Flu® (Aspen Pharmacare SpA) and Zerinolflu® (Boehringer Ingelheim SpA) were obtained commercially. Pure water and ethanol were of instrumental purity grade (J. T. Baker, Holland). All other reagents were of the highest purity commercially available.

2.2. Instruments

Absorption spectra were recorded on a Perkin-Elmer Lambda 40P spectrophotometer under the following conditions: quartz cell 10 mm; wavelength range 200–350 nm; scan rate 1 nm/s; time response 1 s; spectral band 1 nm; data density 0.5, 1.0 and 2.0 point/nm. Spectral acquisition and elaboration were performed by the software package UV Winlab 2.79.01 (Perkin-Elmer). Application of PCR and PLS algorithms was supported by the software package 'The Unscrambler X 10.3' (Camo Process As., Oslo, Norway). MCR-ALS routines were performed under Matlab computer environment and implemented as MATLAB functions. They were used as described in previous works [24,25]. Source files containing these algorithms are available on the website "www.mcrals.info".

2.3. Standard solutions

Stock solutions of the single compounds were prepared in ethanol by dissolving nearly 20.0 mg of each drug in 100 mL calibrated flasks. A calibration set of 18 ternary mixtures was prepared with PAR concentration in the range of 5.05–30.3 mg/L, ASC in the range of 2.04–30.60 mg/L and CHL in the range of 0.20–5.05 mg/L. A second calibration set of 38 quaternary solutions was prepared with the drugs in the following concentration ranges: PAR 5.10–30.60 mg/L, CAF 0.50–5.00 mg/L, CHL 0.20–2.01 mg/L, and PHE 0.21–2.10 mg/L.

The calibration mixtures were prepared by adopting an SLD distributed on five concentration levels. Two further independent validation sets, comprising 15 ternary mixtures and 15 quaternary mixtures, respectively, were then prepared to validate the calibration models. Statistical analysis was performed on data from analysis of three replicates for sample.

2.4. Pharmaceutical samples

Five tablets for each pharmaceutical specialty were reduced to fine powder and suspended in 100 mL of ethanol. The suspension was sonicated for 10 min and then filtered through a 45 μm membrane filter. Samples for analysis were obtained after proper dilution with ethanol.

3. Chemometric elaboration

PCR and PLS are known as factor analysis methods. In the first step of calibration, concentrations and analytical signals from reference samples are used to build a mathematical model. In the following prediction step, this model is used to evaluate the concentration of an unknown sample.

PCR considers all spectral data simultaneously (X variables) and correlates the concentration components (Y variables) with these data in the second phase of multiple regression. On the other hand, PLS modeling processes information from both spectral and concentration data (X and Y) and projects information in the new space of principal components.

In applying multivariate calibration to spectrophotometry, X variables or descriptors are represented by the absorptivity values of the samples at various wavelength values, whereas the Y variables or responses consisted of concentration values. In building multivariate models, PCs have to reach the optimal number because the prediction error decreases with them until they reach an optimal value. The most known validation procedure is full-cross validation, in which one reference is removed from the calibration set at a time and the same sample is predicted by using the calibration built with the remaining references. The number of PCs was chosen by adopting the root mean square error of cross validation (RMSECV), which estimates the error when unknown samples are predicted with the calibration model. The best prediction ability of the models corresponds to the lowest RMSECV.

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{c} - c_i\right)^2}{n}} \quad (1)$$

where $\hat{c}$ is the predicted value for the references; c_i is the real value for the references; and n is the number of references.

External validation was performed by adopting the parameter mean square error of prediction (RMSEP) and relative error (RE) in prediction:

RMSEP = RMSECV calculated for external samples

$$RE(\%) = 100\sqrt{\frac{\sum_{i=1}^{n}(c_i - \hat{c}_i)}{\sum_{i=1}^{n} c_i^2}} \quad (2)$$

MCR–ALS provides the decomposition of the experimental data matrix describing the chemical system into the contributions of the single species as a bilinear relation between the concentrations and the pure spectra, following the generalized law of Lambert–Beer in multi-sample and multivariable version [11,12]. In matrix form, the bilinear model is expressed in the following way:

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \quad (3)$$