ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
18 (1998) 57–65

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

Experimental design for a pharmaceutical formulation: optimisation and robustness[1]

B. Campisi [a,*], D. Chicco [b], D. Vojnovic [b], R. Phan-Tan-Luu [c]

[a] *Department of Economics and Commodity Science, University of Trieste, via Valerio 6, I-34127 Trieste, Italy*
[b] *Department of Pharmaceutical Sciences, University of Trieste, P.le Europa 1, I-34127 Trieste, Italy*
[c] *LMRE, Centre de St. Jérôme, Avenue Escadrille Normandie-Niemen, F-13397 Marseille, Cedex 20, France*

Received 22 October 1997; received in revised form 29 December 1997; accepted 5 January 1998

Abstract

In pharmaceutical industries, the formulator is usually faced with the optimisation of the excipient mixture composition aimed to prepare a product with the required characteristics. Experimental research methodology represents an efficient approach for solving such optimisation problems. Planning mixture experiments using specific designs allows to analyse the blending properties of each mixture component and estimate an empirical model approximating the response of interest as a function of excipient proportions. In this study the evolution of theophylline solubility in a four-component system with constraints was analysed using two mixture design approaches: a classical mixture component proportion approach and a mathematically independent variable approach. An optimal region characterised by high solubility values was found and further explored in order to verify the insensitivity of theophylline solubility to slight variations of the excipient mixture composition.

Keywords: Experimental mixture design; Component proportion approach; Mathematically independent variable approach; Cosolvent mixture optimisation; Theophylline solubility

1. Introduction

A common problem in the preformulation of pharmaceutical dosage forms is the optimisation of the excipient mixture composition aimed to obtain a product with the required characteristics. When the target is to evaluate one or more blend properties, the experimenter can make use of a set of tools and techniques known under the name of experimental research methodology. In general, adopting such an experimental approach means defining the problem which one is going to cope with by determining the objectives, the possible constraints on the component proportions, and the response variables under study. In this way, the experimental region of interest and the strategy to follow can be defined.

A classical mixture problem in pharmaceutical technology is represented by cosolvent systems

* Corresponding author. Tel.: +39 40 6767031; fax: +39 40 6763215; e-mail: campisi@uts.univ.trieste.it

[1] Presented at the 7th Meeting on Recent Developments in Pharmaceutical Analysis, Island of Elba, Italy, September 16–20, 1997.

0731-7085/98/$ - see front matter
PII S0731-7085(98)00175-7

used to enhance the solubility of poorly water-soluble drugs in liquid dosage forms. The optimisation of such cosolvent systems using mixture experimental design has been studied for different drugs by some authors. For instance, the extreme vertices design was applied to the estimation of a model for the butoconazole nitrate solubility in a multicomponent system with upper and lower-bound restrictions [1], whereas a Scheffé statistical design was used for solubility prediction of diazepam and phenobarbital in three-component mixtures [2].

More recently, the theophylline solubility in a cosolvent system consisting of polyethylene glycol 400, water, propylene glycol, and ethanol has been investigated [3]. A Scheffé polynomial was obtained for the description and prediction of theophylline solubility as a function of mixture composition. On the basis of this experimental study it was possible to point out a subregion within the factor space where the observed values of theophylline solubility were relatively high.

In the present paper this subregion was further investigated in order to study in detail the constrained experimental region. For this purpose, two strategies were used, namely a classical mixture component proportion approach and a mathematically independent variable approach. A maximal region was found and further explored in order to verify the insensitivity of the theophylline solubility to slight variations of the formulation.

2. Experimental

2.1. Materials

Theophylline (anhydrous) was supplied by Prodotti Gianni (Italy). Ethanol, polyethylene glycol 400 (PEG-400) and propylene glycol (A.C.E.F. S.p.a., Italy) and bidistilled water were used as cosolvents.

2.2. Solubility determinations

The theophylline solubility in each cosolvent mixture, prepared according to the experimental design described further on, was determined following the method already reported in a previous paper [3]. Saturated theophylline solutions were prepared by adding an excess of the drug to 30 ml of each cosolvent blend. The mixtures were shaken for 24 h at 25°C and subsequently filtered through a 0.8 μm membrane filter (Millipore). The theophylline concentration in each mixture was determined, after proper dilution, by spectrophotometric assay (Perkin-Elmer spectrophotometer—Model 552). The absorbance was read at 271 nm versus the corresponding solvent mixture blank and the drug concentration was estimated from a calibration curve obtained from solutions of theophylline in the same blend.

2.3. Mixture experimental design

In this study, mixture experimental design was applied to the analysis of the evolution of theophylline solubility in a four-component system given by polyethylene glycol (x_1), water (x_2), propylene glycol (x_3), and ethanol (x_4). In the specific case, the experimental region of interest is represented by a subregion inside the tetrahedron (Fig. 1) according to the constraints on the component proportions (Eq. (1)) determined on the basis of the experimental results of a previous study [3]

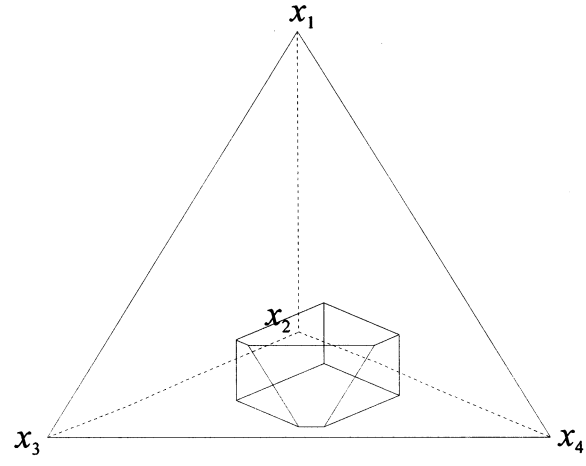

Fig. 1. The constrained experimental region in the four-component system given by the restrictions on the component proportions $0 \leq x_1 \leq 0.2$, $0.1 \leq x_2 \leq 0.7$, $0.1 \leq x_3 \leq 0.45$, and $0.20 \leq x_4 \leq 0.5$.

Table 1
Design point coordinates in the constrained region inside the factor space given by the four blend components

Design points		x_1	x_2	x_3	x_4	Type of boundary
◆	1	0	0.700	0.100	0.200	Vertex
◆	2	0.200	0.500	0.100	0.200	Vertex
◆	3	0	0.400	0.100	0.500	Vertex
◆	4	0.200	0.200	0.100	0.500	Vertex
◆	5	0	0.350	0.450	0.200	Vertex
◆	6	0.200	0.150	0.450	0.200	Vertex
◆	7	0	0.100	0.450	0.450	Vertex
	8	0	0.100	0.400	0.500	Vertex
	9	0.200	0.100	0.450	0.250	Vertex
◆	10	0.200	0.100	0.200	0.500	Vertex
◆	11	0.100	0.600	0.100	0.200	Edge centroid
◆	12	0	0.550	0.100	0.350	Edge centroid
◆	13	0	0.525	0.275	0.200	Edge centroid
◆	14	0.200	0.350	0.100	0.350	Edge centroid
◆	15	0.200	0.325	0.275	0.200	Edge centroid
	16	0.100	0.300	0.100	0.500	Edge centroid
◆	17	0	0.250	0.250	0.500	Edge centroid
	18	0.200	0.150	0.150	0.500	Edge centroid
	19	0.100	0.250	0.450	0.200	Edge centroid
	20	0	0.225	0.450	0.325	Edge centroid
	21	0.200	0.125	0.450	0.225	Edge centroid
	22	0	0.100	0.425	0.475	Edge centroid
	23	0.100	0.100	0.450	0.350	Edge centroid
◆	24	0.100	0.100	0.300	0.500	Edge centroid
◆	25	0.200	0.100	0.325	0.375	Edge centroid
◆	26	0	0.330	0.300	0.370	Face centroid
	27	0.100	0.100	0.375	0.425	Face centroid
	28	0.100	0.450	0.100	0.350	Face centroid
◆	29	0.100	0.425	0.275	0.200	Face centroid
	30	0.200	0.210	0.260	0.330	Face centroid
◆	31	0.100	0.175	0.450	0.275	Face centroid
	32	0.100	0.200	0.200	0.500	Face centroid
	33	0.100	0.270	0.280	0.350	Overall centroid

Symbols (◆) mark the design points selected by the exchange algorithm.

$$\begin{aligned} &0 \le x_1 \le 0.20 \\ &0.10 \le x_2 \le 0.70 \\ &0.10 \le x_3 0.45 \\ &0.20 \le x_4 \le 0.50. \end{aligned} \quad (1)$$

As shown in Fig. 1, the constrained experimental region is represented by an irregular polyhedron with ten vertices, fifteen edges and seven faces. The vertices of the polyhedron were computed using the McLean and Anderson's extreme vertices algorithm [4].

In order to explore the restricted region, two different strategies for mixture design were followed and compared: the well-known component proportion approach and the mathematically independent variable approach.

2.3.1. Component proportion approach

In general, once the experimental region of interest is determined, in order to analyse the evolution of the response variable, an experimental design for mixtures can be chosen, so that a polynomial model can be fitted to the experimental data. In the mixture experiments, the empirical model most commonly used to describe and analyse the evolution of the variable response is

the Scheffé canonical polynomial [4,5]. According to the model order, the experimenter must adopt a particular design. The number of trials to be done depends in fact on the number of coefficients to be estimated. Each Scheffé mixture model can be normally associated to a Scheffé Simplex lattice consisting of enough design points for the model to be estimated and validated. Eq. (2) reports the Scheffé quadratic polynomial for four component {4, 2} chosen to analyse the evolution of the solubility of the theophylline according to the formulation

$$\eta = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 \quad (2)$$

2.3.2. Mathematically independent variable approach

When dealing with mixture experiments, the researcher can opt to work with a system consisting of q-1 orthogonal variables w_1, w_2, ..., w_{q-1}, instead of the q mixture components x_1, x_2, ..., x_q dependent, for definition, one on each other. In general, one of the reasons that may lead one to consider a transformation of the variable system under study is the familiarity with classical designs for independent variables and, in particular, with design optimality criteria [4].

In the (q-1)-dimensional system, the response variation can be approximated by a classical polynomial. For the four-component system under study, the following second-order equation expressed in terms of orthogonal variables was used

$$\eta = \beta_0 + \beta_1 w_1 + \beta_2 w_2 + \beta_3 w_3 + \beta_{11} w_1^2 + \beta_{22} w_2^2 + \beta_{33} w_3^2 + \beta_{12} w_1 w_2 + \beta_{13} w_1 w_3 + \beta_{23} w_2 w_3; \quad (3)$$

In this study, with the intent of testing the equivalence of the two strategies for data treatment, the models in Eqs. (2) and (3) were fitted to the measured response values, and the corresponding contour plots were compared.

Furthermore, the same variable transformation allowed us to apply a classical Doehlert design to the exploration of the maximal region found within the polyhedron defined by the constraints on mixture component proportions.

3. Results and discussion

3.1. Optimisation of the formulation

The mixture experimental design used to estimate the coefficients in Eqs. (2) and (3) was made up of 33 points listed in Table 1. Besides the vertices computed using McLean and Anderson's algorithm, the centroids of the edges joining the vertices, the bounding faces centroids, and the overall centroid were included as additional design points. However, in order to reduce the number of trials, the Fedorov's exchange algorithm was run to generate a new design from the 33 lattice points. This algorithm is one of the iterative methods more commonly used for the construction of D-optimal experimental designs meeting the $\mathbf{X}'\mathbf{X}$ matrix determinant maximisation criterion [4].

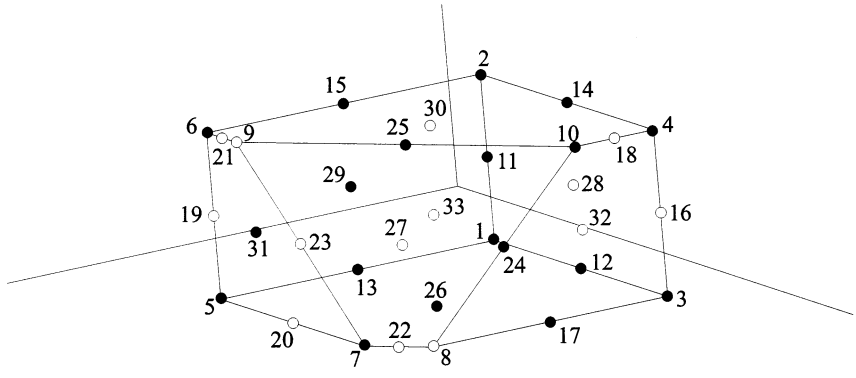

Fig. 2. The vertices and centroids of the constrained factor space chosen as candidate points for selecting a D-optimal design. The black dots mark the 19-point design selected by the exchange algorithm.

Table 2
Mixture variable coordinates in pseudocomponent x'_2 and orthogonal variable w_1 systems, along with the measured response values y_1

Design points	x'_1	x'_2	x'_3	x'_4	w_1	w_2	w_3	y_1 (mg ml^{-1})
1	0	1	0	0	−1	1	−0.077	13.51
2	0.333	0.667	0	0	1	0.556	−0.077	15.39
3	0	0.500	0	0.500	−1	0	−1	26.60
4	0.333	0.167	0	0.500	1	−0.444	−1	21.18
5	0	0.417	0.583	0	−1	−0.167	1	23.26
6	0.333	0.083	0.583	0	1	−0.611	1	17.97
7	0	0	0.583	0.417	−1	−1	0.231	21.22
10	0.333	0	0.167	0.500	1	−0.778	−0.692	14.98
11	0.167	0.833	0	0	0	0.778	−0.077	15.76
12	0	0.750	0	0.250	−1	0.5	−0.538	22.02
13	0	0.708	0.292	0	−1	0.417	0.462	19.06
14	0.333	0.417	0	0.250	1	0.056	−0.538	22.59
15	0.333	0.375	0.292	0	1	−0.028	0.462	23.03
17	0	0.250	0.250	0.500	−1	−0.5	−0.538	26.75
24	0.167	0	0.333	0.500	0	−0.889	−0.385	16.92
25	0.333	0	0.375	0.292	1	−0.778	0.077	18.88
26	0	0.383	0.333	0.283	−1	−0.233	0.015	25.79
29	0.167	0.542	0.292	0	0	0.194	0.462	20.03
31	0.167	0.125	0.583	0.125	0	−0.639	0.769	19.70
33	0.167	0.283	0.300	0.250	0	−0.322	0.015	26.52

Among the D-optimal designs generated using this algorithm, a 19-point design was selected according to the minimum–maximum variance (max var $(\hat{y})/\sigma^2$) criterion which is one of the most commonly used design statistics for this purpose [6]. In Table 1, the 19 design points (points 1–7, 10–15, 17, 24–26, 29, and 31) chosen from the candidate list are marked by the symbol (◆) and by black dots in Fig. 2. The overall centroid of the polyhedron (design point 33) was added to the experimental design in order to test the model.

The polynomials fitted to the observed values y_i, shown in the last column of Table 2, are as in Eqs. (4) and (5)

$$\hat{y} = 23.28x'_1 + 11.99x'_2 + 15.20x'_3 + 9.06x'_4 + 7.74x'_1x'_2 - 6.69x'_1x'_3 + 36.05x'_2x'_3 - 3.17x'_1x'_4 + 65.07x'_2x'_4 + 32.14x'_3x'_4 \quad (4)$$

$$\hat{y} = 23.97 - 2.12w_1 - 3.85w_2 - 2.35w_3 + 0.43w_1^2 - 10.63w_2^2 - 2.36w_3^2 + 0.54w_1w_2 + 0.28w_1w_3 - 3.93w_2w_3 \quad (5)$$

It must be noticed that Eq. (4) is expressed in terms of pseudocomponents x'_i, and not in the original variables x_i. A linear transformation of the variables, $x'_i = (x_i - a_i)/R_a$, where a_i is the lower bound of the component i ($i = 1, 2, ..., q$) and $R_a = 1 - \Sigma_{i=1}^{q} a_i$, was performed in order to obtain a more accurate estimation of the mixture model parameters. In fact, when restraints on a blend composition are considered, the ill-conditioning of the matrix $\mathbf{X'X}$ or collinearity among the original variables often lead to an incorrect least squares solution. One of the possible remedies to this drawback is the transformation of the original components to pseudocomponents [7,8]. In Table 2 the mixture variables are expressed in terms of pseudocomponents x'_i and orthogonal variables w_i.

For the comparison of the two response surfaces (Eqs. (4) and (5)), the corresponding contour plots are shown in Figs. 3 and 4. In order to represent the response evolution in a bidimensional system, one of the variables was to be kept constant. In this case, polyethylene glycol was fixed at two levels: $x_1 = 0$ (Fig. 3(a) and Fig. 4(a)) and $x_1 = 0.20$ (Fig.

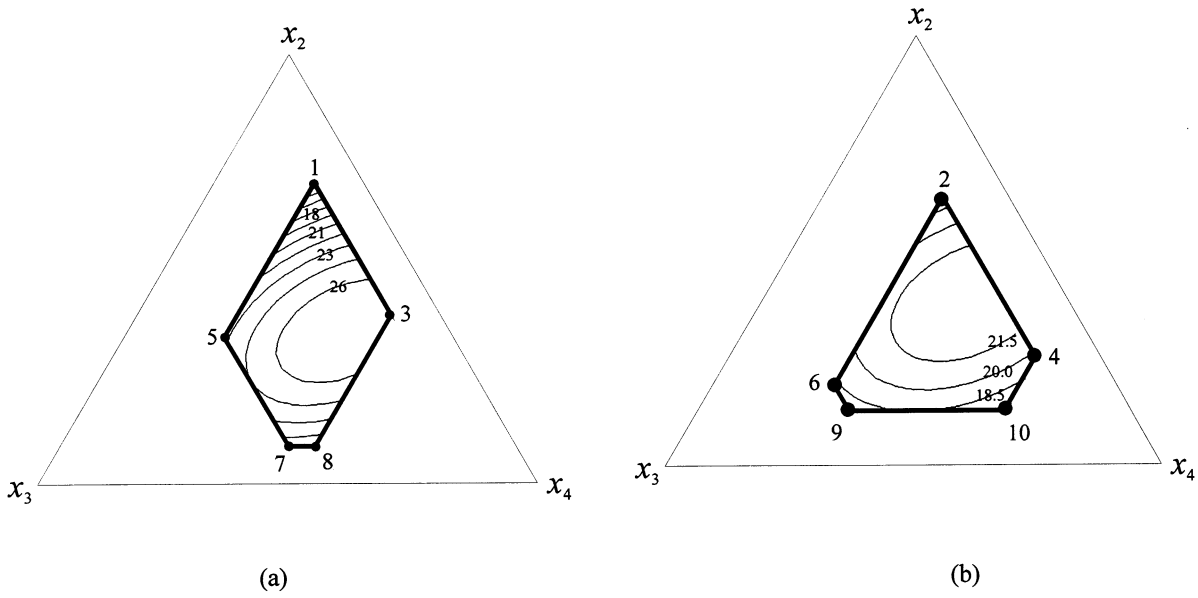

Fig. 3. (a) Contour plot in the restrained region for the mixture component system when $x_1 = 0$; (b) contour plot in the restrained region for the mixture component system when $x_1 = 0.20$.

3(b) and Fig. 4(b)), respectively. It must be pointed out that in the representation of the surface contours in the orthogonal variable system, the w_2 axis corresponds to blend component x_2, whereas on the w_3 axis both components x_3 and x_4 are located.

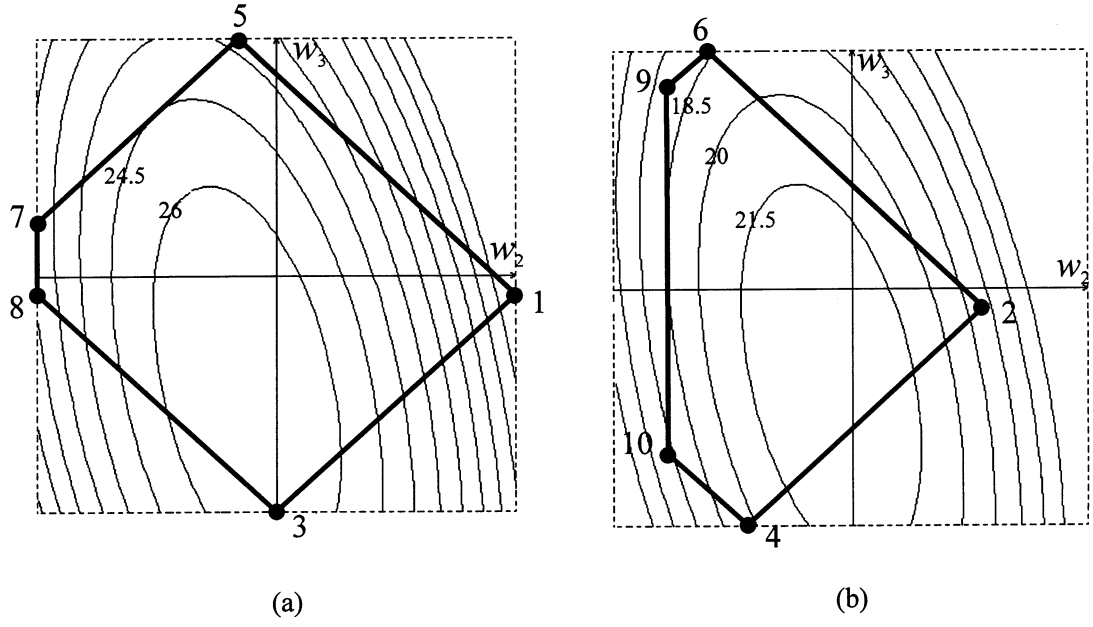

Fig. 4. (a) Contour plot in the restrained region for the orthogonal variable system when $w_1 = -1$; (b) contour plot in the restrained region for the orthogonal variable system when $w_1 = +1$.

Table 3
Boundaries of the optimal region within the polyhedron, along with the coordinates of point $\mathbf{x}_0$

	Component	a_i	x_{0i}	b_i
x_1	Polyethylene glycol	0	0.50	0.10
x_2	Water	0.20	0.30	0.40
x_3	Propylen glycol	0.10	0.25	0.40
x_4	Ethanol	0.30	0.40	0.50

In Figs. 3 and 4, the lines joining the experimental points mark the constrained region wherein a similar depiction of the expected solubility is obtained by both approaches. Formulations along each contour, though representing different excipient mixtures, should yield the same solubility value predicted by the fitted model. On the basis of these surface contours it was possible to identify a maximal region where the blend optimisation could be achieved.

3.2. Robustness of the formulation

The boundaries of the maximal region to be explored for the robustness of the formulation were defined according to the highest observed values of solubility (design points 3, 17, 26 and 33 in Table 2). The definition of these boundaries implied stricter constraints on mixture components as reported in Table 3. This maximal region, whose centre is point $\mathbf{x}_0$ (Table 3), can be represented in the mixture space by an ellipsoid (Fig. 5) of the following form

$$\left(\frac{x_1-0.05}{0.05}\right)^2+\left(\frac{x_2-0.20}{0.10}\right)^2+\left(\frac{x_3-0.30}{0.15}\right)^2+\left(\frac{x_4-0.20}{0.10}\right)^2\le 1 \quad (6)$$

since, in the general q-component case, an ellipsoidal region can be expressed as

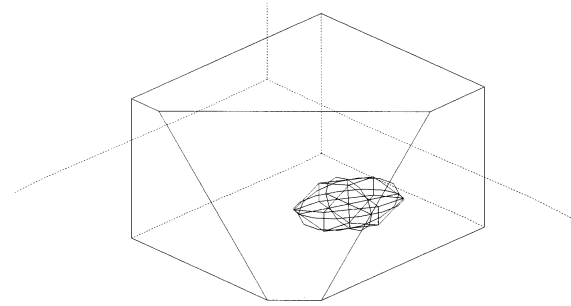

Fig. 5. Ellipsoid in the mixture space representing the maximal region within the polyhedron.

Table 4
Doehlert design point coordinates z_i with the corresponding mixture component proportions x_i, along with the measured response values y_i

Design point	z_1	z_2	z_3	x_1	x_2	x_3	x_4	y_1 (mg ml^{-1})
1	1	0	0	0.023	0.327	0.250	0.400	28.26
2	−1	0	0	0.077	0.273	0.250	0.400	26.18
3	0.5	0.866	0	0.027	0.276	0.296	0.400	28.66
4	−0.5	−0.866	0	0.073	0.324	0.204	0.400	28.35
5	0.5	−0.866	0	0.046	0.350	0.204	0.400	25.62
6	−0.5	0.866	0	0.054	0.250	0.296	0.400	29.24
7	0.5	0.289	0.816	0.031	0.289	0.238	0.443	29.20
8	−0.5	−0.289	−0.816	0.069	0.311	0.262	0.357	28.15
9	0.5	−0.289	−0.816	0.043	0.338	0.262	0.357	27.15
10	0	0.577	−0.816	0.047	0.288	0.308	0.357	32.16
11	−0.5	0.289	0.816	0.057	0.262	0.238	0.443	31.69
12	0	−0.577	0.816	0.053	0.312	0.192	0.443	27.64
13	0	0	0	0.050	0.300	0.250	0.400	30.56
14	0	0	0	0.050	0.300	0.250	0.400	29.15
15	0	0	0	0.050	0.300	0.250	0.400	27.99
16	0	0	0	0.050	0.300	0.250	0.400	29.06

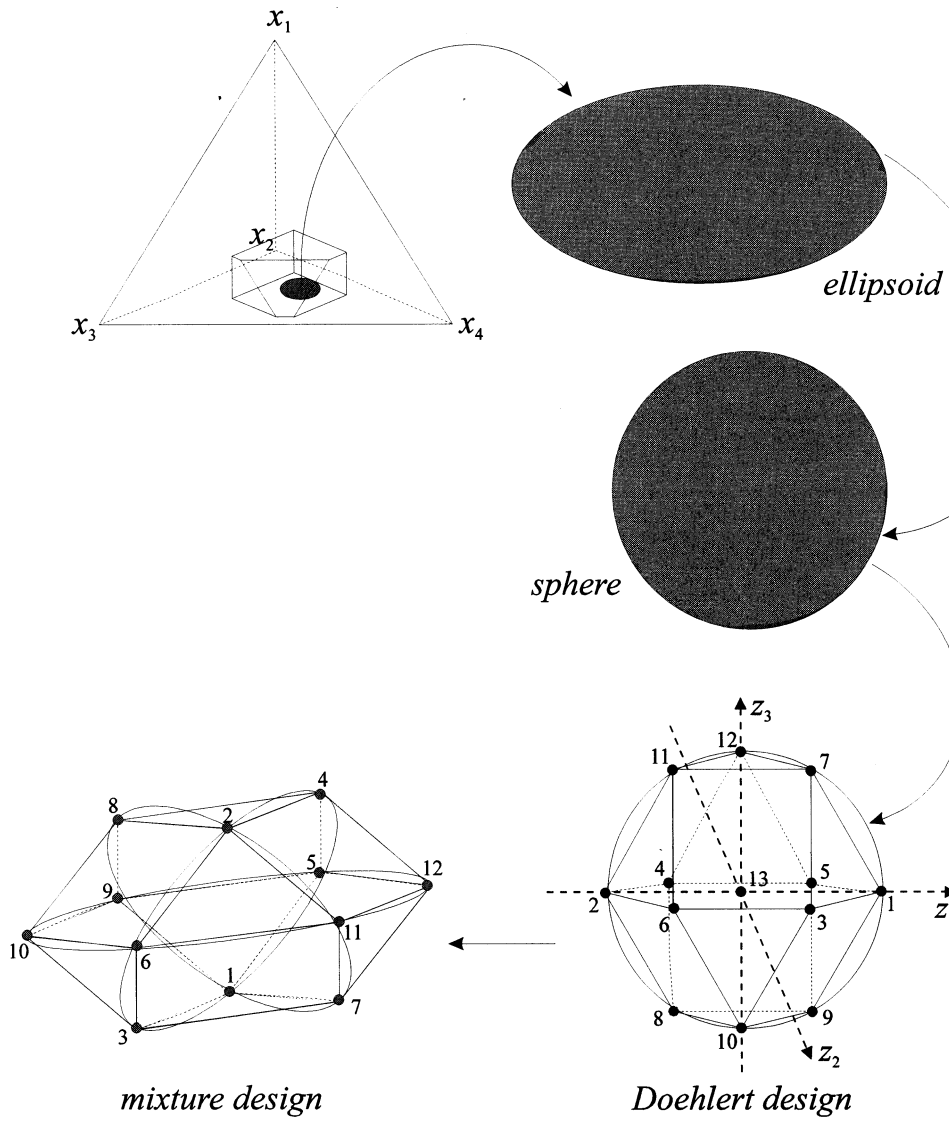

Fig. 6. The transformation steps of ellipsoidal region to the unit spherical region, along with the corresponding experimental design.

$$\sum_{i=1}^{q}\left(\frac{x_i - x_{0i}}{h_i}\right)^2 \leq 1 \qquad (7)$$

where x_{0i} represents the proportion of each mixture component for point $\mathbf{x}_0$ and h_i corresponds to one half of the range of interest for component i.

In order to simplify the system, the ellipsoid was transformed to a more familiar unit spherical region so that a classical design could be adopted. Among the most commonly used experimental design, a Doehlert design was chosen. In Table 4, the coordinates of the experimental design points for both regions are shown, along with the observed values of theophylline solubility. In Fig. 6, the transformation steps of the ellipsoidal region to the unit spherical region, and the corresponding experimental designs are presented.

A classical second-degree model and a Scheffé mixture polynomial were both fitted to the experimental data collected at the Doehlert design points. Both response surfaces turned out not to

be statistically significant and thus rejected. This outcome might lead to the conclusion that the solubility fluctuations observed within the maximal region were not due to the variation of mixture composition and therefore that the formulation could be considered 'robust' from this point of view. However, further studies are needed in order to ascertain the possible presence of noise factors affecting the theophylline solubility in the considered four-component system.

4. Conclusions

In this study the application of a transformation to orthogonal variables for a mixture system was introduced for the optimisation of the theophylline solubility in a four-cosolvent blend.

The equivalence between this approach and the one based on the classical mixture component proportions was pointed out by comparing the two response surfaces obtained using the two strategies. The feasibility of employing a mathematically independent variable system while dealing with a mixture problem allowed a well-known classical design for independent variables to be used. A Doehlert design was run in order to investigate in further detail the theophylline solubility and the formulation robustness within the maximal region found in the constrained factor space.

Acknowledgements

This study has been made possible by a fellowship from the Italian National Research Council (CNR) to the first author for the Research Project 'Quality optimisation of products and processes using experimental design methodology'.

References

[1] S.T. Anik, L. Sukumar, J. Pharm. Sci. 8 (70) (1981) 897–900.
[2] R.J. Belloto, A.M. Dean, M.A. Moustafa, A.M. Molokhia, M.W. Gouda, T.D. Sokoloski, Int. J. Pharm. 23 (1985) 195–207.
[3] D. Vojnovic, D. Chicco, Drug Dev. Ind. Pharm. 7 (23) (1997) 639–645.
[4] J.A. Cornell, Experiments with Mixtures. Designs, Models, and the Analysis of Mixture Data, Wiley, New York, 1990.
[5] H. Scheffé, J.R. Stat. Soc., B 20 (2) (1955) 344–360.
[6] D. Cook, C. J. Nachtsheim, Technometrics 22 (3) (1980) 315–324.
[7] R.D. Snee, Technometrics 17 (2) (1975) 149–159.
[8] R.D. Snee, A.A. Rayner, J. Qual. Technol. 14 (2) (1982) 67–79.