International Journal of Pharmaceutics, 19 (1984) 27–34
Elsevier

IJP 00639

Granule and tablet formulae study by principal component analysis

L. Benkerrour [1], D. Duchêne [1], F. Puisieux [1] and J. Maccario [2]

[1] *Laboratoire de Pharmacie Galénique, et* [2] *Laboratoire de Mathématiques, Faculté de Pharmacie de Paris-Sud, 92290 Chatenay Malabry (France)*

(Received November 22nd, 1982)
(Modified version received October 6th, 1983)
(Accepted October 13th, 1983)

Summary

The authors wanted to demonstrate the interest of the use of Principal Component Analysis (PCA) for the interpretation of galenical results. To do this, they applied this method to the study of the effect on granules and tablets of diluent type, and concentration and viscosity of the wetting liquid.

Following this work, it seems that PCA can be recommended when the number of formulae and parameters studied culminate in a large number of results. In fact, PCA enables better definition and limitation of the number of parameters involved. Where formula optimization is desired, a method such as Simplex can be used in a better clarified framework.

Introduction

The purpose of this work is to study the modifications of granule and tablet characteristics under the effect of diluents, and the concentration and viscosity of the wetting liquid.

When numerous assays are performed on granules and tablets, the galenic interpretation of the results may be difficult. Our opinion is that it is fruitful to

Correspondence: D. Duchêne, Laboratoire de Pharmacie Galénique, Faculté de Pharmacie de Paris-Sud, rue Jean Baptiste Clément, 92290 Chatenay Malabry, France.

perform a comprehensive description of the results by using a multidimensional statistical tool such as Principal Component Analysis (PCA).

This method is preferred to the classical Analysis of Variance (ANOVA) as a first approach, for at least 2 reasons: (a) contrary to the PCA method, ANOVA requires rigorous planning of the experiments which more or less assumes some prior knowledge of the effects of formulae variations; and (b) the investigator usually determines more than one characteristic for one formula and is thus led either to a multiplication of ANOVAs, or the use of a Multivariate ANOVA whose results may be more difficult to interpret, while those of PCA can easily cope with such multidimensional data.

Principles of PCA

Although the method described by Caillez and Pages (1976) is classic, we found it useful to give a brief description of the method in order to arrive at a better understanding of our comments.

Later on, a given formula will be considered as an 'individual' for which n characteristics or parameters are measured, e.g. granule strength, friability, porosity, microporosity, etc.

To picture the formulae in terms of their n parameters would necessitate representation in an n-dimensional Euclidian space. This kind of representation is, of course, impossible. However, to obtain readable pictures, one may orthogonally project this n-dimensional space onto a plane. The problem then is to choose the best plane so that the less the original information is altered, the better the plane. Geometrically, this may be stated: to choose the plane where the relative positions of the projections of formulae are closest to those in the original n-dimensional space.

Algebraically, the best plane is defined by the two eigenvectors corresponding to the highest eigenvalues of the parameter correlation matrix. In other words, we might say that an n-dimensional ellipsoid fits around a 'cloud' of formulae. This ellipsoid will have approximately the same shape as the 'cloud'. The two major stretch directions of the ellipsoid, usually called principal components or factorial axes or factors, define the desired plane. The computations involved may be numerous, but are not beyond the capacities of microcomputers. We performed ours in less than 10 s on a main frame computer (Univac 1110, Paris-Sud Informatique).

Positive correlations between the parameters, or similarities between the formulae, will give adjoining points, and negative correlations, or dissimilarities between formulae, will give distant points. Conversely, if two formulae are represented close to each other, they give almost identical results, and if they are far from each other, they give very different results. The same holds for the parameters measured on granules and tablets corresponding to these formulae.

This governs the interpretation of the figure which is the major result of PCA and should be read in simple terms of vicinity or distance. Interpretation is easiest when the original cloud decomposes into disjointed subsets.

In this work, we applied this technique for a better understanding of the effects of wetting liquid and binder.

Experimental Results

We used lactose as the water-soluble diluent and tricalcium phosphate as the water-insoluble diluent. These products, alone or mixed (50 : 50) were added to 1% of erythrosine as tracer. These were then granulated with an aqueous solution of guar gum at concentrations of 0.5, 1 and 1.5% (Table 1). The two brands (A and B) of guar gum used differ in their rate of polymerization and hence in their viscosity (Table 2). The quantities of solution employed vary according to the solubility or insolubility of the powders (Table 1), but are constant for one type of powder regardless of the gum used and its concentration.

Overall 18 formulae were realized with each type of gum following a well-defined and rigorously constant protocol described by Benkerrour (1980). Classical assays were performed: (a) on granules whose granulometry ranged from 400 to 600 $\times 10^{-6}$ m; and (b) on tablets made at a constant pressure of 200 MPa (single punch machine Frogerais, Type OA, flat 12 mm diameter punches) with the lubricated granules (0.3% magnesium stearate + 2% talc).

The results of a classical galenic study have already been published by Benkerrour et al. (1982). In the present analysis, each formula is described by 12 characters: 7 are related to the granules and 5 to the tablets (Table 3).

The matrix correlation of the results is given in Table 4, and the graphic representation of characteristics and formulae in Fig. 1. Axes 1 and 2 account for nearly 90% of the total inertia of the original formula 'cloud' (77.5% for the first axis, 12% for the second), which is a good hint as to the amount of information accounted for by projection on this plane.

TABLE 1

COMPOSITION OF THE DIFFERENT FORMULAE INVESTIGATED

Formula	1	2	3
tricalcium phosphate	100	50	0
lactose	0	50	100
gum solution	370	190	80

TABLE 2

APPARENT VISCOSITIES OF GUAR GUM SOLUTIONS

GUM	A			B		
Concentration (%)	0.5	1.0	1.5	0.5	1.0	1.5
Apparent viscosity (mPas)	132.5	643.0	1494.0	7.5	19.5	90.0

TABLE 3

PRINCIPAL COMPONENT ANALYSIS (DATA)

	G_1	G_2	G_3	G_4	G_5	G_6	G_7	C_1	C_2	C_3	C_4	C_5
$A_1 1$	16.0	184	12.0	9.5	79.0	476	554	87	8.0	2.2	289	38
$A_2 1$	14.9	186	11.8	9.8	84.0	485	539	88	8.0	1.8	284	37
$A_3 1$	18.6	188	11.7	9.4	85.5	481	538	88	8.0	1.6	286	36
$B_1 1$	17.4	184	12.0	9.5	82.0	478	533	86	9.0	2.4	289	39
$B_2 1$	16.4	182	11.0	9.9	87.0	474	529	87	8.0	2.0	292	37
$B_3 1$	15.4	186	11.8	9.6	86.5	475	546	88	8.0	1.7	285	36
$A_1 2$	20.6	182	13.2	9.1	77.0	292	373	81	6.0	1.7	221	19
$A_2 2$	21.9	176	13.6	9.3	91.5	259	327	82	6.0	1.6	208	16
$A_3 2$	18.6	181	13.8	9.0	91.0	268	336	83	6.0	1.5	228	18
$B_1 2$	19.9	172	13.4	9.4	88.0	258	313	81	6.0	1.9	221	19
$B_2 2$	19.3	174	13.8	9.7	86.5	258	325	82	5.5	1.8	245	20
$B_3 2$	16.0	178	13.5	9.2	87.0	255	330	81	6.0	1.6	211	16
$A_1 3$	29.8	217	14.1	8.7	85.0	148	268	76	3.5	4.0	145	8
$A_2 3$	22.7	200	15.0	8.6	92.0	141	272	75	4.0	3.2	154	8
$A_3 3$	24.0	202	14.9	8.2	92.0	156	279	78	5.0	2.6	151	9
$B_1 3$	26.8	212	13.2	7.7	86.0	139	320	79	4.0	3.3	151	8
$B_2 3$	22.7	200	14.5	8.5	91.0	143	278	79	4.5	3.2	147	8
$B_3 3$	25.7	203	14.1	8.4	91.0	132	295	79	5.0	2.8	150	9

RESULTS OF FORMULAE IN TERMS OF FORMULAE

Measurements on granules			
G1	percentage of fines	(%)	measured on the whole of the batch manufactured
G2	bulk volume before tapping	(cm^3)	mean of 3 tests
G3	settling rate	(%)	mean of 3 tests
G4	flow rate	(g/s)	mean of 5 tests
G5	granule strength	(%)	measured on 5 g of grains
G6	micropore volume	(mm^3/g)	mean of 2 tests
G7	total pore volume	(mm^3/g)	mean of 2 tests
Measurements on tablets			
C1	ratio of axially transmitted pressure to applied pressure	(%)	mean of 10 tests
C2	hardness for constant pressure	(N)	mean of 10 tests
C3	friability	(%)	mean of 10 tablets
C4	tablet porous volume	(mm^3/g)	mean of 2 tests
C5	massic area	(m^2/g)	mean of 2 tests

Results of PCA

In Fig. 1, the positions of the formulae depend mainly on the type of diluent, which divides the formulae into 3 homogeneous and well-defined groups. This shows that the effect of wetting liquid viscosity, or brand and concentration of guar gum it contains, are far less important than the water-solubility of the diluent.

TABLE 4

CORRELATION MATRIX OF THE TWELVE MEASURED PARAMETERS (ALL THE FIGURES ARE MULTIPLIED BY 1000)

	G1	G2	G3	G4	G5	G6	G7	C1	C2	C3	C4	C5
G1	1000											
G2	781	1000										
G3	679	394	1000									
G4	−813	−786	−692	1000								
G5	328	193	578	−388	1000							
G6	−826	−538	−926	818	−551	1000						
G7	−748	−372	−942	693	−567	974	1000					
C1	−825	−566	−915	769	−446	958	939	1000				
C2	−850	−609	−874	788	−469	966	933	948	1000			
C3	789	894	473	−692	176	−598	−468	−676	−661	1000		
C4	−872	−683	−886	884	−529	972	914	948	948	−705	1000	
C5	−858	−639	−908	867	−556	989	941	955	967	−662	994	1000

The relative heterogeneity of group 3 of the formulae (diluent = 100% lactose), due mainly to low guar gum concentrations ($B_1$3 and $A_1$3, 0.5%), shows that the effect of wetting liquid may be increased by a totally water-soluble diluent.

The abnormal position of formula $A_1$2 (guar gum A at 0.5%, diluent 50% lacrose, 50% tricalcium phosphate) is due to the particularly low value, for this formula, of granule strength (G5) as may be seen in Table 3. The question that remains unanswered is whether this is a specific behaviour of the formula or an experimental error. Various hints exist in favour of the second interpretation, because of similar positions of formulae with low guar gum concentrations found elsewhere ($A_1$1 and $B_1$1, $A_1$3 and $B_1$3).

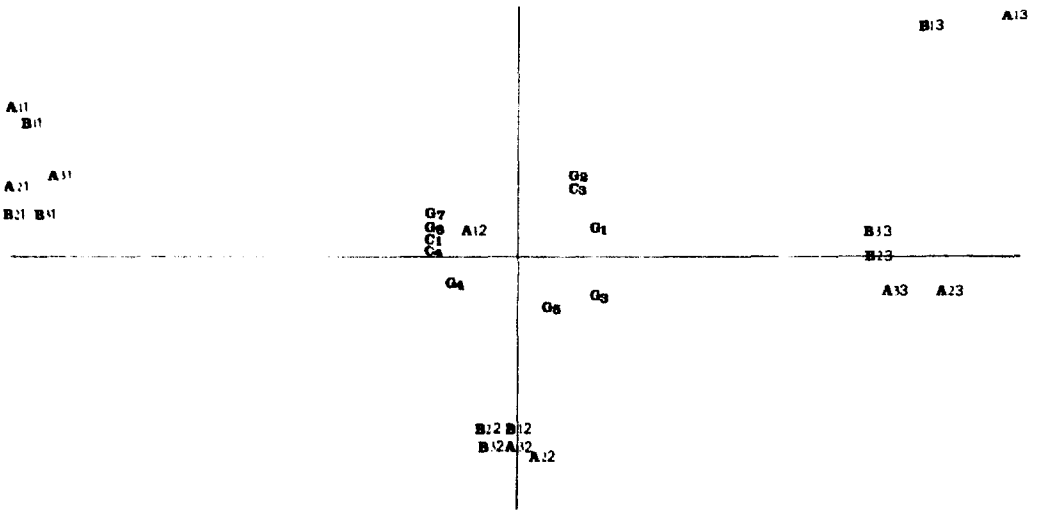

Fig. 1. Graphic representation of characteristics and formulae according to the PCA method.

TABLE 5

GROUPING OF PARAMETERS ACCORDING TO THEIR CORRELATIONS WITH AXES (RANKED BY DECREASING CORRELATION)

Axis no. 1		
C4	tablet porous volume	0.984
C5	massic area	0.982
C6	micropore volume	0.949
C2	tablet hardness	0.930
C1	ratio of pressure	0.920
Axis no. 2		
G2	bulk volume before tapping	0.430
C3	tablet friability	0.339
Axis no. 3		
G5	granule strength	0.456

Amongst the characteristics, two couples are very positively correlated. First, hardness for constant pressure (C2) and ratio of pressure (C1), and secondly the porous volume of tablets (C4) and massic area (C5).

Their correlation coefficients are respectively $r(C1/C2) = 0.948$ and $r(C4/C5) = 0.994$.

The first couple shows that the better the axial transmission pressure of the formulae, the harder the tablets and vice versa.

The value of r for the second couple is just the manifestation of the well-known interdependency of these characteristics.

The set of parameters is roughly divided into two groups, one on the left side of axis 1, and the second on its right side, showing opposite behaviour, or negative correlations between these groups.

Additional computations also allow the evaluation of correlation coefficients of the parameter with the axes themselves. Results are given in Table 5, where a third axis is mentioned which is orthogonal to the projection plane of Fig. 1. It will be seen below that this axis is of little importance.

Discussion of the results of PCA

The third axis mentioned accounts for about 5% of the data shown in Table 3. Such a low value may be considered as an experimental error if one recollects that nearly 90% are already accounted for by the other two axes. The correlation of this third axis with only the G5 parameter (granule strength) backs up this interpretation. We have already seen that the $A_1 2$ formula gives a discrepancy in the results for this parameter. In other terms, it may be said that the variations of formulae positions along the third axis are negligible compared with their variations along the other two axes, or else that granule strength is not significantly modified by the variations in diluent or wetting liquid.

The first axis, according to the parameters correlated with the formulae and also their positions, may be called a 'porosity' axis, with high porosity values on the left side and low values on the right. By relating this to the positions of the formulae, one is led to the conclusion that an insoluble diluent will increase porosity, while a soluble diluent will decrease porosity.

Remarks. As the major cause of dispersion along this first axis is the type of diluent, it may be stated that the type and concentration of the guar gum in the wetting liquid is of secondary importance for 'porosity'. The term porosity is to be understood in a generic sense. In fact the tablet porous volume (C4) is highly positively correlated with its massic area (C5), granule micropore volume (G6), pressure ratio (C1), and tablet hardness under constant pressure (C2), and all these parameters are increased with insoluble diluent and virtually insensitive to variations in wetting liquid.

The second axis is less easily commented on. In fact the parameters correlated to it (granule bulk volume before tapping (G2) and tablet friability (C3)) are also correlated to the first axis. This means that variations in these two parameters are explained partly by the first axis, i.e. solubility variations of the diluent. But there remains to be explained a number of variations accounted for by the second axis, by construction statistically uncorrelated with the first.

Along this second axis, it may be seen that the formulae are distributed in terms of the type of wetting liquid. This is true for formulae 1 and 3 whose concentrations 1 ($A_1$1, $B_1$1, and $A_1$3, $B_1$3) are higher than the others. The second axis may then be viewed as describing the effect of guar gum concentration, particularly sensitive when the diluent is hydrosoluble. It can be verified in Table 3 that hardness increases with concentration.

Formulae 2 (50% lactose, 50% tricalcium phosphate) exhibit the lowest and most homogeneous friability. This is due to the binding effect of the soluted and recrystallized lactose added to the low natural friability of tricalcium phosphate. As pointed out by Opakunle and Spring (1976), there exists a composition of a mixture of soluble and insoluble products that gives minimal friability.

Conclusion

PCA cannot answer all the questions raised by the problems of defining a new tablet formulation. But PCA enables us to deal with a large number of parameters and to group them according to their variations in terms of formulae modifications. From these groupings, more or less quantitative causality relationships can often be deduced which will orient further trials.

Moreover, one can limit the number of measured parameters by making a choice based on correlations between them. And in the case where an optimization is sought, another method, such as the Simplex method, may be used in a narrower and better defined frame.

In the case of a systematic study, PCA may reveal unknown interactions or effects.

Finally, it appears to us that PCA is a tool to be recommended when the lack of prior information may motivate the investigator to multiply tested formulae and measured parameters. If such a multiplication is of an intellectual comfort, it is well known that it leads to serious difficulties in the interpretation of the results. PCA, as a synthetic descriptive method, mitigates the consequences of such an attitude.

Acknowledgement

Computations and graphical displays were obtained with the software ADDAD, Institut de Statistiques des Universités de Paris, Université Paris VI.

References

Benkerrour, L., Etude rhéologique de deux gommes guar. Influence de la solubilité du diluant et de la viscosité du liant sur les caractéristiques des grains et des comprimés. Thèse de 3ème Cycle (Pharmacie), No. 25, Faculté de Pharmacie de Paris-Sud, 1980.

Benkerrour, L., Puisieux, F. and Duchêne, D., Influence de la nature du diluant, de la viscosité et de la concentration en liant du liquide de mouillage, sur les caractéristiques des grains et des comprimés. Pharm. Acta Helv., 57 (1982) 301–308.

Caillez, F. and Pages, J.P., Introduction à l'Analyse des Données, Smash, Paris, 1976.

Opakunle, W.O. and Spring, M.S., The granulation of binary mixtures: the effects of the properties of the component powders on granules. J. Pharm. Pharmacol., 28 (1976) 915–918.