Chapter 18

Optimization Techniques in Pharmaceutical Formulation and Processing

Joseph B. Schwartz

Philadelphia College of Pharmacy, Philadelphia, Pennsylvania

Robert E. O'Connor

Pharmaceutical Sourcing Group Americas, a division of Ortho-McNeil Pharmaceutical, Bridgewater, New Jersey

Roger L. Schnaare

Philadelphia College of Pharmacy, Philadelphia, Pennsylvania

I. INTRODUCTION

A significant portion of this book is devoted to the concepts involved in formulating drug products in their various forms. Physical, chemical, and biological properties all must be given due consideration in the selection of components and processing steps for that dosage form. The final product must be one that meets not only the requirements placed on it from a bioavailability standpoint, but also the practical mass production criteria of process and product reproducibility. In the current regulatory climate, formulation and process justification is a requirement for preapproval inspections for all new drug applications. In fact, development reports for both formulation and process are reviewed during these inspections. It is in the best interest of the pharmaceutical scientist to understand the theoretical formulation and target processing parameters, as well as the ranges for each excipient and processing parameter. Optimization techniques provide both a depth of understanding and an ability to explore and defend ranges for formulation and processing factors. With a rational approach to the selection of the several excipients and manufacturing steps for a given product, one qualitatively selects a formulation. It is at this point that optimization can become a useful tool to quantitate a formulation that has been qualitatively determined. Optimization is not a screening technique.

The word "optimize" is defined as follows: to make as perfect, effective, or functional as possible [1]. The last phrase, "as possible," leads one immediately into the area of decisions making, since one might ask: (a) perfect by whose definition; (b) for what characteristics; and (c) under what conditions? The term "optimization" is often used in pharmacy relative to formulation and to processing, and one will find it in the literature referring to any study of the formula. In developmental projects, one generally experiments by a series of logical steps, carefully controlling the variables and changing one at a time until a satisfactory system is produced. If the experimenter had sufficient help or sufficient time, he or she would eventually perfect the formulation, but under the circumstances the "best" one is often simply the last one prepared. It is satisfactory, but how close is it to the optimum, and how does the experimenter know?

No matter how rationally designed, the trial-and-error method can be improved upon. It is the purpose

of this chapter to discuss the general principles behind the techniques of optimization and to review the specific techniques that have been applied to pharmaceutical systems.

II. OPTIMIZATION PARAMETERS

A. Problem Types

There are two general types of optimization problem: constrained and unconstrained. Constraints are restrictions placed on the system by physical limitations or perhaps by simple practicality (e.g., economic considerations). In unconstrained optimization problems there are no restrictions. For a given pharmaceutical system one might wish to make the hardest tablet possible. The constrained problem, on the other hand, would be stated: make the hardest tablet possible, but it must disintegrate in less than 15 minutes.

Within the realm of physical reality, and most important in pharmaceutical systems, the unconstrained optimization problem is almost nonexistent. There are always restrictions that the formulator wishes to place or must place on a system, and in pharmaceuticals, many of these restrictions are in competition. For example, it is unreasonable to assume, as just described, that the hardest tablet possible would also have the lowest compression and ejection forces and the fastest disintegration time and dissolution profile. It is sometimes necessary to trade off properties, that is, to sacrifice one characteristic for another. Thus, the primary objective may not be to optimize absolutely (i.e., a maxima or minima), but to realize an overall pre selected or desired result for each characteristic or parameter. Drug products are often developed by teaching an effective compromise between competing characteristics to achieve the best formulation and process within a given set of restrictions.

An additional complication in pharmacy is that formulations are not usually simple systems. They often contain many ingredients and variables, which may interact with one another to produce unexpected, if not unexplainable, results.

B. Variables

The development of a pharmaceutical formulation and the associated process usually involves several variables. Mathematically, they can be divided into two groups. The independent variables are the formulation and process variables directly under the control of the formulator. These might include the level of a given ingredient or the mixing time for a given process step. The dependent variables are the responses or the characteristics of the in-progress material or the resulting drug delivery system. These are a direct result of any change in the formulation or process.

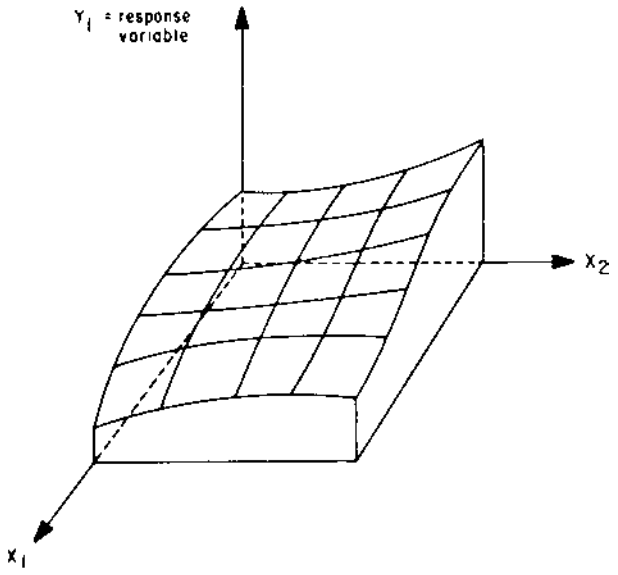


Fig. 1 Response surface representing the relationship between the independent variables X_1 and X_2 and the dependent variable Y_1.

The more variables one has in a given system, the more complicated becomes the job of optimization. But regardless of the number of variables, there will be a relationship between a given response and the independent variables. Once this relationship is known for a given response, it defines a response surface, such as that represented in Fig. 1. It is this surface that must be evaluated to find the values of the independent variables, X_1 and X_2, which give the most desirable level of the response, Y. Any number of independent variables can be considered; representing more than two becomes graphically impossible, but mathematically only more complicated.

III. CLASSIC OPTIMIZATION

Classic optimization techniques result from application of calculus to the basic problem of finding the maximum or minimum of a function. The techniques themselves have limited application, but they might be useful for problems that are not too complex and do not involve more than a few variables. The concept, however is important.

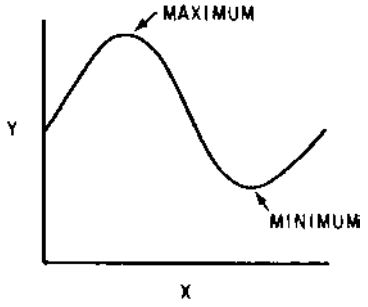


Fig. 2 Graphic location of optimum (maximum or minimum).

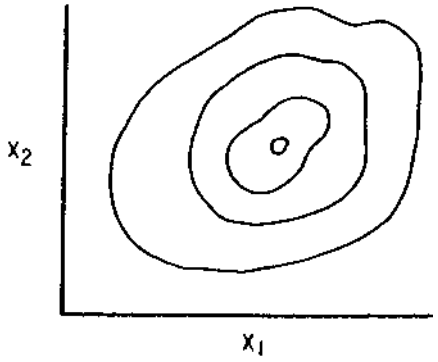


Fig. 3 Contour plot. Contours represent values of the dependent variable Y.

The curve in Fig. 2 might represent the relationship between a response Y and a single independent variable X in a hypothetical system, and since we can see the whole curve, we can pick out the highest point or lowest, the maximum or minimum. Use of calculus, however, makes the task of plotting the data or equation unnecessary. If the relationship, that is, the equation for Y as a function of X, is available [Eq. (1)]:

$$Y = f(X) \tag{1}$$

we can take the first derivative, set it equal to zero, and solve for X to obtain the maximum or minimum. For many functions of X, there will be more than one solution when the first derivative is set equal to zero. The various solutions may all be maxima or minima, or a mixture of both.

There are also techniques to determine whether we are dealing with a maximum or a minimum, that is, by use of the second derivative. And there are techniques to determine whether we simply have a maximum (one of several local peaks) or the maximum. Such approaches are covered in elementary calculus texts and are well presented relative to optimization in a review by Cooper and Steinberg [2].

When the relationship for the response Y is given as a function of two independent variables, X_1 and X_2,

$$Y = f(X_1, X_2) \tag{2}$$

the problem is slightly more involved. Graphically, there are contour plots (Fig. 3) on which the axes represent the two independent variables, X_1 and X_2, and the contours (analogous to elevations, as on a contour map) represent a specific level of Y. Again, we can select an optimum graphically. Mathematically appropriate manipulations with partial derivatives of the function can locate the necessary pair of X values for the optimum.

The situation with multiple variables (any more than two) becomes graphically impossible. It is still possible by mathematics, but very involved, making use of partial derivatives, matrices, determinants, and so on. The reader is referred to optimization texts for further details. Because of the complications involved and because the classic calculus methods apply basically to unconstrained problems, more practical methods are generally used.

IV. STATISTICAL DESIGN

The techniques most widely used for optimization may be divided into two general categories: one in which experimentation continues as the optimization study proceeds, and another in which the experimentation is completed before the optimization takes place. The first type is represented by evolutionary operations and the simplex method, and the second by the more classic mathematical and search methods. (Each of these is discussed in Sec. V.)

For the techniques of the second type, it is necessary that the relation between any dependent variable and the one or more independent variables be known. To obtain the necessary relationships, there are two possible approaches: the theoretical and the empirical.

If the formulator knows a priori the theoretical equation for the formulation properties of interest, no experimentation is necessary. However, much of the work in pharmaceutics has been in the pursuit of such relationships, and to our knowledge most have not been determined. Therefore, it remains the task of the formulator to generate the relationships between the variables for the particular formulation and process.

In a text on experimental design, Davis states [3]:

> Theoretically, the behavior of chemical reactions, or for that matter the behavior of any system, is governed

by ascertainable laws, and it should be possible to determine optimum conditions by applying such laws. In practice, however, the underlying mechanisms of the system are frequently so complicated that an empirical approach is necessary.

To apply the empirical or experimental approach for a system with a single independent variable, the formulator experiments at several levels, measures the property of interest, and obtains a relationship, usually by simple regression analysis or by the least-squares method. In general, however, there is more than one important variable, so the experimenter must enter into the realm of "statistical design of experiments and multiple linear regression analysis." Statistical design and multiple linear regression analysis are separate and rather large fields, and, again, the reader is referred to appropriate texts [3–5,40]. The concept of interest to the pharmacist planning to utilize optimization techniques is that there are methods available for selecting one's experimental points so that (a) the entire area of interest is covered or considered, and (b) analysis of the results will allow separation of variables (i.e., statistical analysis can be performed, which allows the experimenter to know which variable caused a specific result).

One of the most widely used experimental plans is that of the factorial design, or some variation of it (two of the techniques in the following section utilize it). By multiple regression techniques, the relationships between variables, then, are generated from experimental data, and the resulting equations are the basis of the optimization. These equations define the response surface for the system under investigation.

V. APPLIED OPTIMIZATION METHODS

There are many methods that can be, and have been, used for optimization, classic and otherwise. These techniques are well documented in the literature of several fields. Deming and King [6] presented a general flowchart (Fig. 4) that can be used to describe general optimization techniques. The effect on a real system of changing some input (some factor or variable) is observed directly at the output (one measures some property), and that set of real data is used to develop mathematical models. The responses from the predictive models are then used for optimization. The first two methods discussed here, however, omit the mathematical-modeling step; optimization is based on output from the real system.

A. Evolutionary Operations

One of the most widely used methods of experimental optimization in fields other than pharmaceutical technology is the evolutionary operation (EVOP). This technique is especially well suited to a production situation. The basic philosophy is that the production

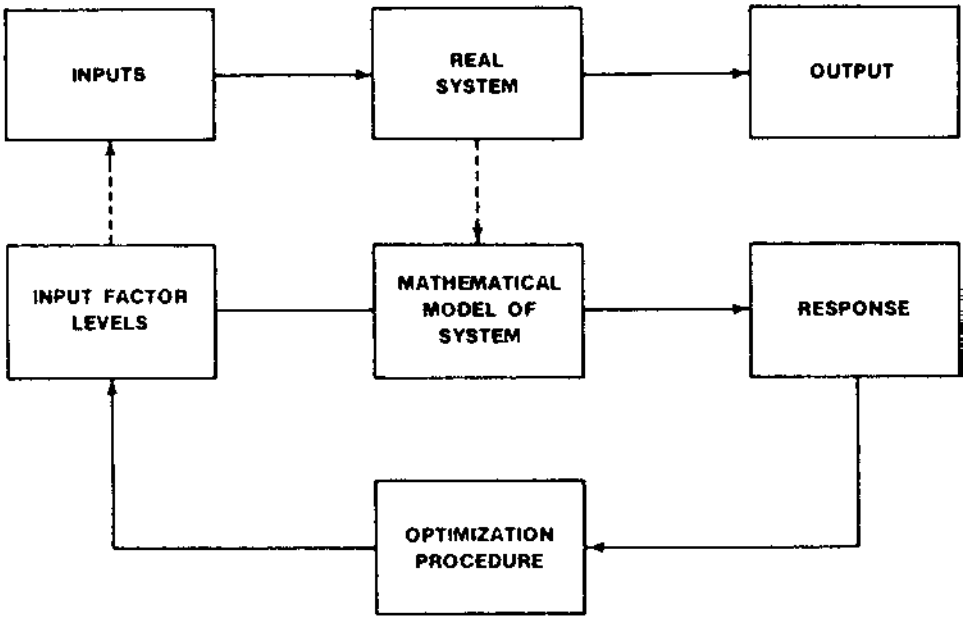


Fig. 4 Flowchart for optimization.

procedure (formulation and process) is allowed to evolve to the optimum by careful planning and constant repetition. The process is run in a way such that it both produces a product that meets all specifications and (at the same time) generates information on product improvement.

By this method the experimenter makes a very small change in the formulation or process but makes it so many times (i.e., repeats the experiment so many times) that he or she can determine statistically whether the product has improved. If it has, the experimenter makes another change in the same direction many times and notes the results. This continues until further changes do not improve the product or perhaps become detrimental. The experimenter then has the optimum—the peak.

In an industrial process, this large number of experiments is usually not a problem, since the process will be run over and over again. The application of this technique to tablets has been advocated by Rubinstein [7]. It has also been applied to an inspection system for parenteral products [8].

In most pharmaceutical situations, however, there is often insufficient latitude in the formula or process to allow the necessary experimentation. The pharmaceutical industry is subject to regulatory constraints that make EVOP impossible to employ in validated production processes and, therefore, impractical and expensive to use. Moreover, EVOP is not a substitute for good laboratory-scale investigation and, because of the necessarily small changes utilized, is not particularly suitable to the laboratory. In pharmaceutical development, more efficient methods are desired.

B. The Simplex Method

The simplex approach to the optimum is also an experimental method and has been applied more widely to pharmaceutical systems. Originally proposed by Spendley et al. [9], the technique has even wider appeal in areas other than formulation and processing. A particularly good example to illustrate the principle is the application to the development of an analytical method (a continuous flow analyzer) by Deming and King [6].

A simplex is a geometric figure that has one more point than the number of factors. So, for two factors or independent variables, the simplex is represented by a triangle. Once the shape of a simplex has been determined, the method can employ a simplex of fixed size or of variable sizes that are determined by comparing the magnitudes of the responses after each successive calculation. Figure 5 represents the set of simplex movements to the optimum conditions using a variable size technique.

The two independent variables (the axes) show the pump speeds for the two reagents required in the analysis reaction. The initial simplex is represented by the lowest triangle; the vertices represent the spectrophotometrie response. The strategy is to move toward a better response by moving away from the worst response. Since the worst response is 0.25, conditions are selected at the vortex, 0.6, and, indeed, improvement is obtained. One can follow the experimental path to the optimum, 0.721.

For pharmaceutical formulations, the simplex method was used by Shek et al. [10] to search for an optimum capsule formula. This report also describes the necessary techniques of reflection, expansion, and contraction for the appropriate geometric figures. The same laboratories applied this method to study a solubility problem involving butoconazole nitrate in a multicomponent system [11].

Bindschaedler and Gurny [12] published an adaptation of the simplex technique to a TI-59 calculator and applied it successfully to a direct compression tablet of acetaminophen (paracetamol). Janeczek [13] applied the approach to a liquid system (a pharmaceutical solution) and was able to optimize physical stability. In a later article, again related to analytical techniques, Deming points out that when complete knowledge of the response is not initially available, the simplex method is probably the most appropriate type [14]. Although not presented here, there are sets of rules for the selection of the sequential vertices in the procedure, and the reader planning to carry out this type of procedure should consult appropriate references.

C. The Lagrangian Method

This optimization method, which represents the mathematical techniques, is an extension of the classic method and was the first, to our knowledge, to be applied to a pharmaceutical formulation and processing problem. Fonner et al. [15] chose to apply this method to a tablet formulation and to consider two independent variables. The active ingredient, phenylpropanolamine HCl, was kept at a constant level, and the levels of disintegrant (corn starch) and lubricant (stearic acid) were selected as the independent variables, X_1 and X_2. The dependent variables include tablet hardness, friability, volume, in vitro release rate, and urinary excretion rate in human subjects.

This technique requires that the experimentation be completed before optimization so that mathematical

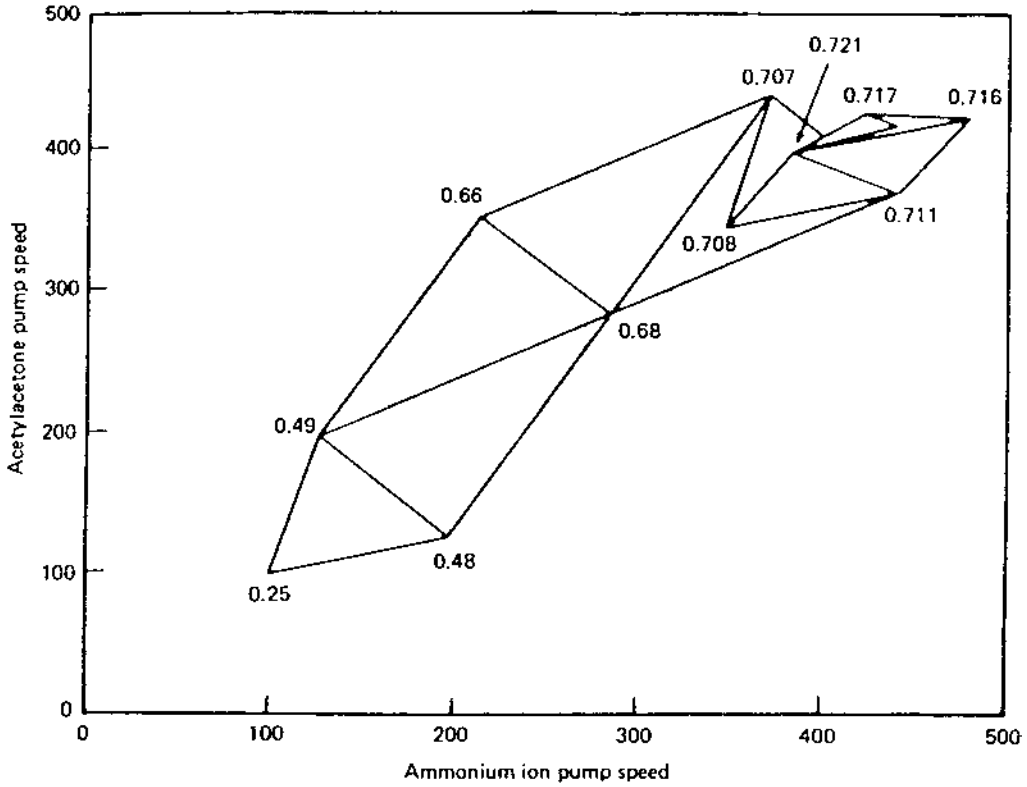


Fig. 5 The simplex approach to optimization. Response is spectrophotometric reading at a given wavelength. (From Ref. 6.)

models can be generated. The experimental design here was a full 3^2 factorial, and, as shown in Table 1, nine formulations were prepared. Polynomial models relating the response variables to the independent variables were generated by a backward stepwise regression analysis program. The analyses were performed on a polynomial of the form

$$y = B_0 + B_1X_1 + B_2X_2 + B_3X_1^2 + B_4X_2^2 + B_5X_1X_2 + B_6X_1X_2^2 + B_7X_1^2X_2 + B_8X_1^2X_2^2 \quad (3)$$

and the terms were retained or eliminated according to standard stepwise regression techniques. In Eq. (3), y represents any given response and B_i represents the regression coefficient for the various terms containing

Table 1 Tablet Formulations

	Ingredient per tablet (mg)			
Formulation no.	Phenylpropanolamine HCl	Dicalcium phosphate · $2H_2O$	Starch	Stearic acid
1	50	326	4 (1%)	20 (5%)
2	50	246	84 (21%)	20
3	50	166	164 (41%)	20
4	50	246	4	100 (25%)
5	50	166	84	100
6	50	86	164	100
7	50	166	4	180 (45%)
8	50	86	84	180
9	50	6	164	180

Source: Ref. 15.

levels of the independent variables. One equation is generated for each response or dependent variable.

A graphic technique may be obtained from the polynomial equations, as represented in Fig. 6. Figure 6a shows the contours for tablet hardness as the levels of the independent variables are changed. Figure 6b shows similar contours for the dissolution response, $t_{50\%}$. If the requirements on the final tablet are that hardness be 8–10 kg and $t_{50\%}$ be 20–33 min, the feasible solution space is indicated in Fig. 6c. This has been obtained by superimposing Fig. 6a and b, and several different combinations of X_1 and X_2 will suffice.

Slightly different constraints are used to illustrate the mathematical technique. In this example, the constrained optimization problem is to locate levels of stearic acid (X_1) and starch (X_2) that minimize the time of in vitro release (y_2) such that the average tablet volume (y_4) did not exceed 9.422 cm^2 and the average friability (y_3) did not exceed 2.72%.

To apply the Lagrangian method, this problem must be expressed mathematically as follows:

$$\text{Minimize } y_2 = F_2(X_1, X_2) \tag{4}$$

such that

$$y_3 = f_3(X_1, X_2) \leq 2.72 \tag{5}$$

$$y_4 = F_4(X_1, X_2) \leq 0.422 \tag{6}$$

and

$$5 \leq X_1 \leq 45 \tag{7}$$

$$1 \leq X_2 \leq 41 \tag{8}$$

Equations (7) and (8) serve to keep the solution within the experimental range.

The foregoing inequality constraints must be converted to equality constraints before the operation begins, and this is done by introducing a slack variable q, for each. The several equations are then combined into a Lagrange function F, and this necessitates the introduction of a Lagrange multiplier, λ, for each constraint.

Then, following the appropriate steps (i.e., partial differentiation of the Lagrange function) and solving the resulting set of six simultaneous equations, values are obtained for the appropriate levels of X_1 and X_2, to yield an optimum in vitro time of 17.9 mm ($t_{50\%}$). The solution to a constrained optimization program may depend heavily on the constraints applied to the secondary objectives.

oA technique called *sensitivity analysis* can provide information so that the formulator can further trade off one property for another. For sensitivity analysis the formulator solves the constrained optimization problem for systematic changes in the secondary objectives. For example, the foregoing problem restricted tablet friability, y_3, to a maximum of 2.72%. Figure 7 illustrates the in vitro release profile as this constraint is tightened or relaxed and demonstrates that substantial improvement in the $t_{50\%}$ can be obtained up to about 1–2%. Subsequently, the plots of the independent variables, X_1 and X_2, can be obtained as shown in Fig. 8. Thus the formulator is provided with the solution (the formulation) as he changes the friability restriction.

The several steps in the Lagrangian method can be summarized as follows:

1. Determine objective function.
2. Determine constraints.
3. Change inequality constraints to equality constraints.
4. Form the Lagrange function, F:
   a. One Lagrange multiplier λ for each constraint
   b. One slack variable q for each inequality constraint
5. Partially differentiate the Lagrange function for each variable and Set derivatives equal to zero.
6. Solve the set of simultaneous equations.
7. Substitute the resulting values into the objective functions.

Although many steps in the procedure may be carried out by computer, the application requires significant mathematical input from the person involved.

Buck et al. [16] expanded on the previous work and proposed that the statistical design technique can be incorporated into an overall management philosophy for proposed product design. The authors discussed four phases in this philosophy, which are defined as (a) a preliminary planning phase, (b) an experimental phase, (c) an analytical phase, and (d) a verification phase. They include case studies of a tablet design and a suspension design to illustrate the efficient and effective procedures that might be applied. Representation of such analysis and the available solution space is shown for the suspension in Figs. 9 and 10.

D. Search Methods

In contrast with the mathematical optimization methods, search methods do not require continuity or differentiability of the function—only that it be

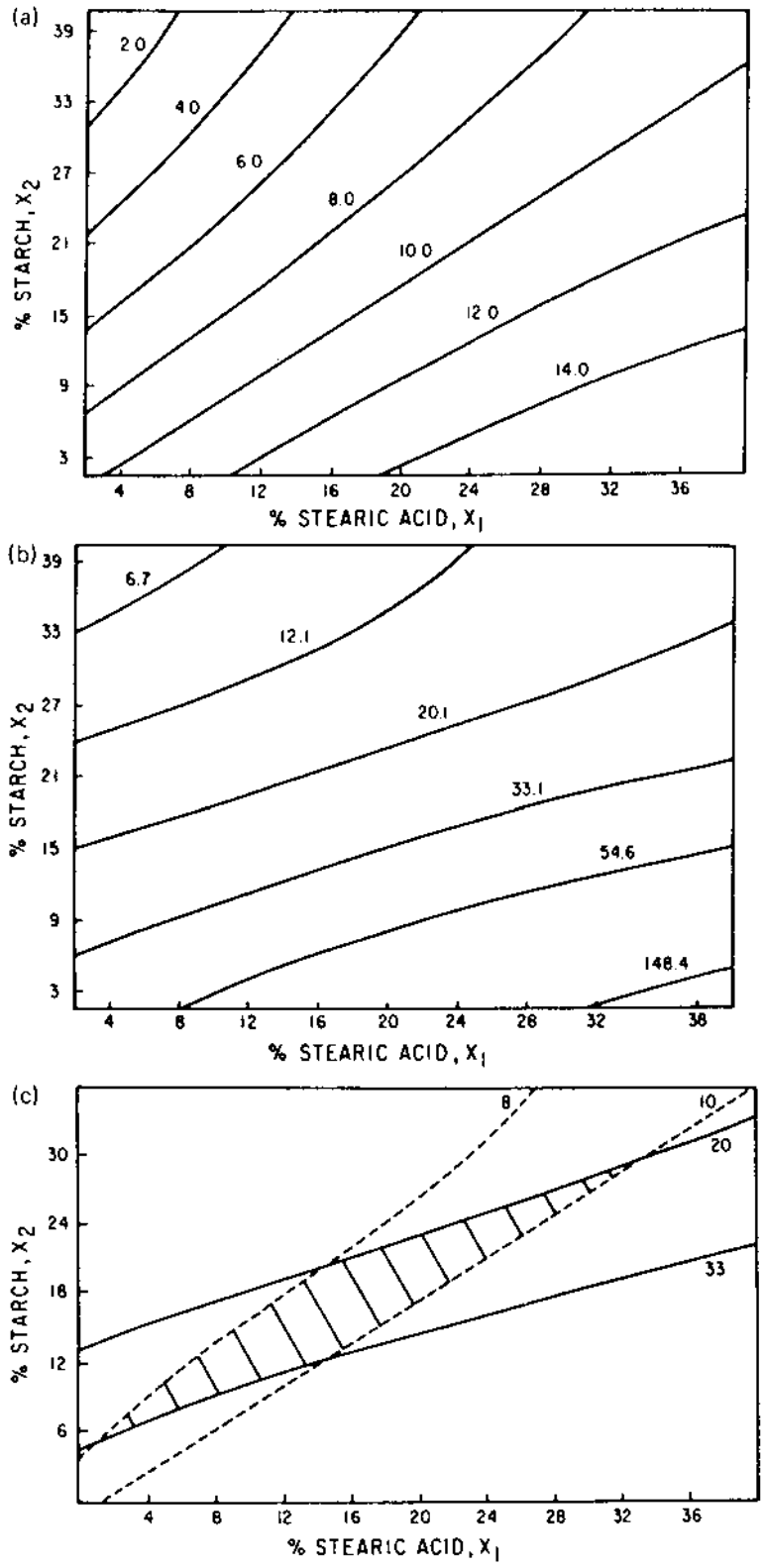


Fig. 6 Contour plots for the Lagrangian method: (a) tablet hardness; (b) dissolution ($t_{50\%}$); (c) feasible solution space indicated by crosshatched area. (From Ref. 15.)

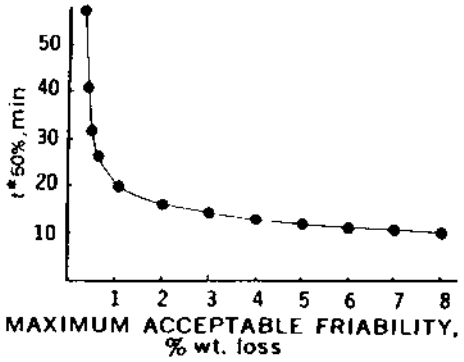


Fig. 7 Optimum in vitro $t_{50\%}$ release rate as a function of restrictions on tablet friability. (From Ref. 15.)

computable. In these methods the response surfaces, as defined by the appropriate equations, are searched by various methods to find the combination of independent variables yielding the optimum.

Although the Lagrangian method was able to handle several responses or dependent variables, it was generally limited to two independent variables. A search method of optimization was also applied to a pharmaceutical system and was reported by Schwartz et al. [17]. It takes five independent variables into account and is computer-assisted. It was proposed that the procedure described could be set up such that persons unfamiliar with the mathematics of optimization and with no previous computer experience could carry out an optimization study.

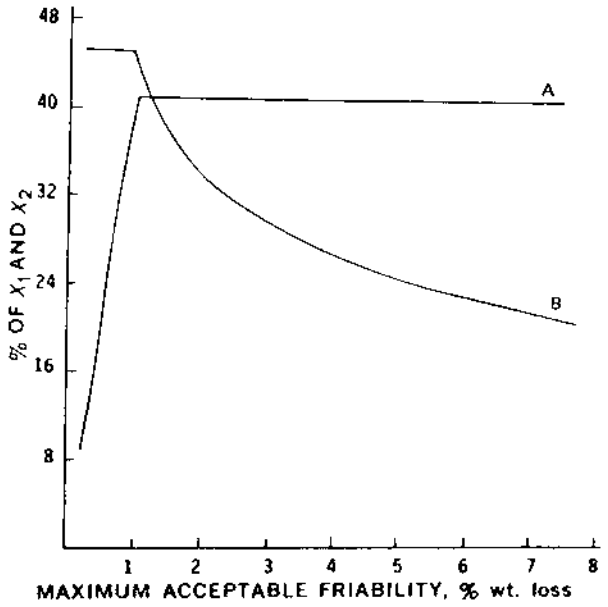


Fig. 8 Optimizing values of stearic acid and starch as a function of restrictions on tablet friability: (A) percent starch; (B) percent stearic acid. (From Ref. 15.)

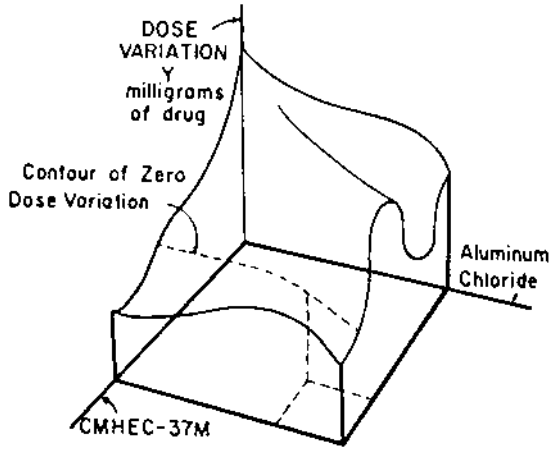


Fig. 9 Response surface concept and results of the second case study. (From Ref. 16.)

The system selected here was also a tablet formulation. The five independent variables or formulation factors selected for this study are shown in Table 2. The dependent variables are listed in Table 3. Since each dependent variable is considered separately, any number could have been included.

The experimental design used was a modified factorial and is shown in Table 4. The fact that there are five independent variables dictates that a total of 27 experiments or formulations be prepared. This design is known as a five-factor, orthogonal, central, composite, second-order design [3]. The first 16 formulations represent a half-factorial design for five factors at two levels, resulting in $\frac{1}{2} \times 2^5 = 16$ trials. The two levels are represented by $+1$ and -1, analogous to the high and low values in any two-level factorial design. For the remaining trials, three additional levels were selected: zero represents a base level midway between the aforementioned levels, and the levels noted as 1.547 represent extreme (or axial) values.

The translation of the statistical design into physical units is shown in Table 5. Again the formulations were prepared and the responses measured. The data were subjected to statistical analysis, followed by multiple regression analysis. This is an important step. One is not looking for the best of the 27 formulations, but the

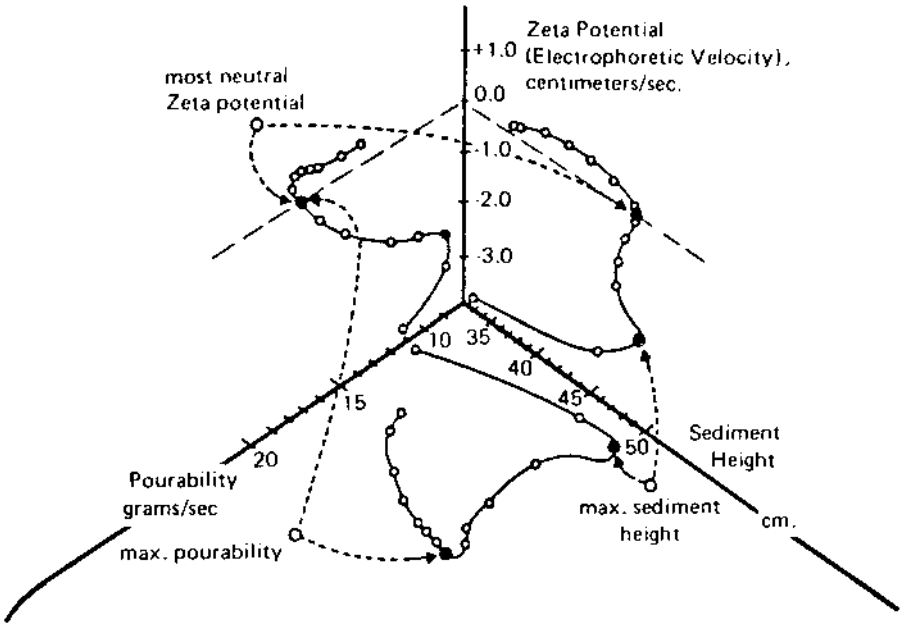


Fig. 10 Secondary properties of various suspensions yielding zero dose variation. (From Ref. 16.)

"global best." The type of predictor equation used with this type of design is a second-order polynomial of the following form:

$$Y = a_0 + a_1X_1 + \cdots + a_5X_5 + a_{11}X_1^2 + \cdots + a_{55}X_5^2 + a_{12}X_1X_2 + a_{13}X_1X_3 + \cdots + a_{45}X_4X_5 \quad (9)$$

Table 2 Formulation Variables (Independent)

X_1	Diluent ratio
X_2	Compressional force
X_3	Disintegrant level
X_4	Binder level
X_5	Lubricant level

Table 3 Response Variables (Dependent)

Y_1	Disintegration time
Y_2	Hardness
Y_3	Dissolution
Y_4	Friability
Y_5	Weight uniformity
Y_6	Thickness
Y_7	Porosity
Y_8	Mean pore diameter

where Y is the level of a given response, a_{ij} the regression coefficients for second-order polynomial, and X_i the level of the independent variable. The full equation has 21 terms, and one such equation is generated for each response variable. The usefulness of the equation is evaluated by the R^2 value, or the index of determination, which is an indication of the fit. In most cases the fit was satisfactory, and the equations were used. One possible disadvantage of the procedure as it is set up is that not all pharmaceutical responses will fit a second-order regression model. In fact, further analysis was attempted, and the results indicated that one of the responses was adequately described by a modified third-order model (inter-action terms were eliminated.) However, a significant advantage of the digital system utilized is that it can be modified to accept other mathematical models—another order polynomial, any other empirical relationship, or a mathematical model based on first principles.

For the optimization itself, two major steps were used: the feasibility search and the grid search. The feasibility program is used to locate a set of response constraints that are just at the limit of possibility. One selects the several values for the responses of interest (i.e., the responses one wishes to constrain), and a search of the response surface is made to determine whether a solution is feasible. For example, the constraints in Table 6 were fed into the computer and were relaxed one

Table 4 Experimental Design

Trial	Factor level in experimental units				
	X_1	X_2	X_3	X_4	X_5
1	−1	−1	−1	−1	1
2	1	−1	−1	−1	−1
3	−1	1	−1	−1	−1
4	1	1	−1	−1	1
5	−1	−1	1	−1	−1
6	1	−1	1	−1	1
7	−1	1	1	−1	1
8	1	1	1	−1	−1
9	−1	−1	−1	1	−1
10	1	−1	−1	1	1
11	−1	1	−1	1	1
12	1	1	−1	1	−1
13	−1	−1	1	1	1
14	1	−1	1	1	−1
15	−1	1	1	1	−1
16	1	1	1	1	1
17	−1.547	0	0	0	0
18	1.547	0	0	0	0
19	0	−1.547	0	0	0
20	0	1.547	0	0	0
21	0	0	−1.547	0	0
22	0	0	1.547	0	0
23	0	0	0	−1.547	0
24	0	0	0	1.547	0
25	0	0	0	0	−1.547
26	0	0	0	0	1.547
27	0	0	0	0	0

Source: Adapted from Ref. 17.

at a time until a solution was found. The first feasible solution was found at disintegration time = 5 min, hardness = 10 kg, and dissolution = 100% at 50 min: This program is designed so that it stops after the first possibility; it is not a full search. The formulation obtained may be one of many possibilities satisfying the constraints.

The next step, the grid search, is essentially a brute-force method in which the experimental range is divided into a grid of specific size and methodically searched. The method is called an exhaustive grid search. From an input of the desired criteria, the program prints out all points (formulations) that satisfy the constraints.

The purpose of the preliminary step of the feasibility program is simply to limit the number of solutions in the grid search. In addition to providing a printout of each formulation, the grid search program also gives the corresponding values for the responses. At this point, the experimenter can trade off one response for another, and the fewer possibilities there are, the easier the job. Thus, the best or most acceptable formulation is selected from the grid search printout to complete the optimization.

The two steps just discussed require that one or more responses be constrained, and a question may arise as to which ones to select. The formulator may have certain basic constraints, such as a minimum hardness value, but it is nevertheless important to know which property or properties can be used to distinguish between the available choices. Generally, this is done by an educated guess, based on experience with the system and with pharmaceutical systems in general.

Table 5 Experimental Conditions

Factor	−1.547 eu	−1 eu	Base 0	+1 eu	+1.547 eu
X_1 = Calcium phosphate/ lactose ratio (1 eu = 10 mg)	24.5/55.5	30/50	40/40	50/30	55.5/24.5
X_2 = compression pressure (1 eu = 0.5 ton)	0.25	0.5	1	1.5	1.75
X_3 = Corn starch disintegrant (1 eu = 1 mg)	2.5	3	4	5	5.5
X_4 = Granulaitng gelatin (1 eu = 0.5 mg)	0.2	0.5	1	1.5	1.8
X_5 = Magnesium stearate (1 eu = 0.5 mg)	0.2	0.5	1	1.5	1.8

Source: Ref. 17.

Table 6 Specifications for Feasibility Search

Variable	Constraint	Experimental range[a]
Disintegration time (min)	1(1)[b]	1.33–30.87
	3(2)	
	5(3)	
Hardness (kg)	12(1)[b]	3.82–11.60
	10(2)	
	8(3)	
Dissolution (% at 50 min)	100(1)[b]	13.30–89.10
	90(2)	
	80(3)	

[a]It is possible to request values for a response that are more desirable than any data obtained in the set of 27 experiments.
[b](1) = first choice.

However, there is a mathematical method for selecting those variables that best distinguish between formulations—those variables that change most drastically from one formulation to another and that should be the criteria on which one selects constraints. A multivariate statistical technique called *principal component analysis* (PCA) can effectively be used to answer these questions. PCA utilizes a variance-covariance matrix for the responses involved to determine their interrelationships. It has been applied successfully to this same tablet system by Bohidar et al. [18].

In addition to the programs to select the optimum discussed previously, graphic approaches are also available and graphic output is provided by a plotter from computer tapes. The output includes plots of a given response as a function of a single variable (Fig. 11) or as a function of all five variables (Fig. 12). The abscissa for both types is produced in experimental units, rather than physical units, so that it extends from -1.547 to $+1.547$ (see Table 5). Use of the experimental units allows the superpositioning of the single plots (see Fig. 11) to obtain the composite plots (see Fig.12).

An infinite number of these plots is possible, since for each curve represented, four of the five variables must remain constant at some level. This is analogous to a partial derivative situation, and the slope of any one graph does indeed represent a partial derivative of the response for one of the independent variables. It will change, depending on the level of the other four variables.

Contour plots (Fig. 13) are also generated in the same mariner. The specific response is noted on the graph, and, again, the three fixed variables must be held at some desired level. For the contour plots shown, both axes are in experimental units (eu). This technique is automated so that a formulator with no previous computer experience and no familiarity with the mathematics of optimization can follow the steps necessary to complete such a study. Those steps may be summarized as follows:

1. Select a system.
2. Select variables:
   a. Independent
   b. Dependent
3. Perform experiments and test product.
4. Submit data for statistical and regression analysis.
5. Set specifications for feasibility program.
6. Select constraints for grid search.
7. Evaluate grid search printout.
8. Request and evaluate:.
   a. "Partial derivative" plots, single or composite
   b. Contour plots

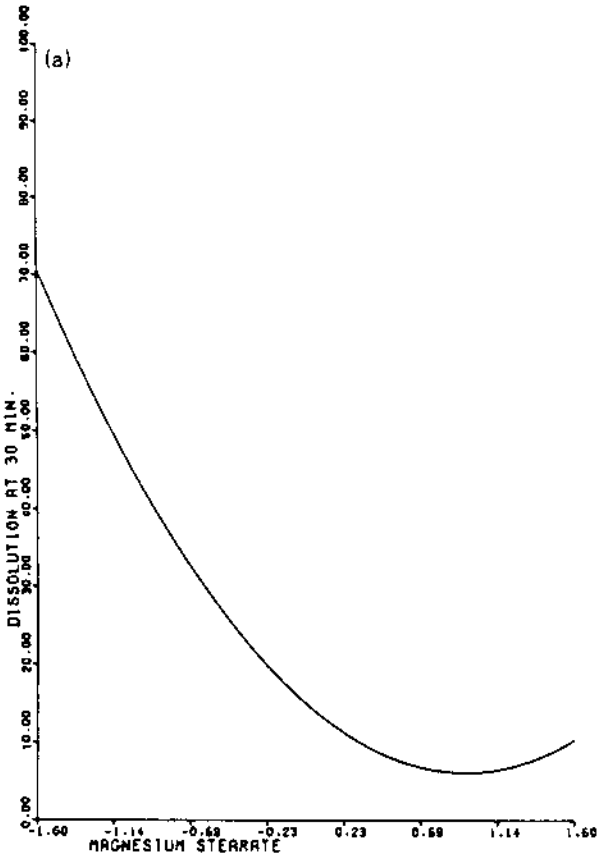


Fig. 11 Computer-generated plots for a single variable.

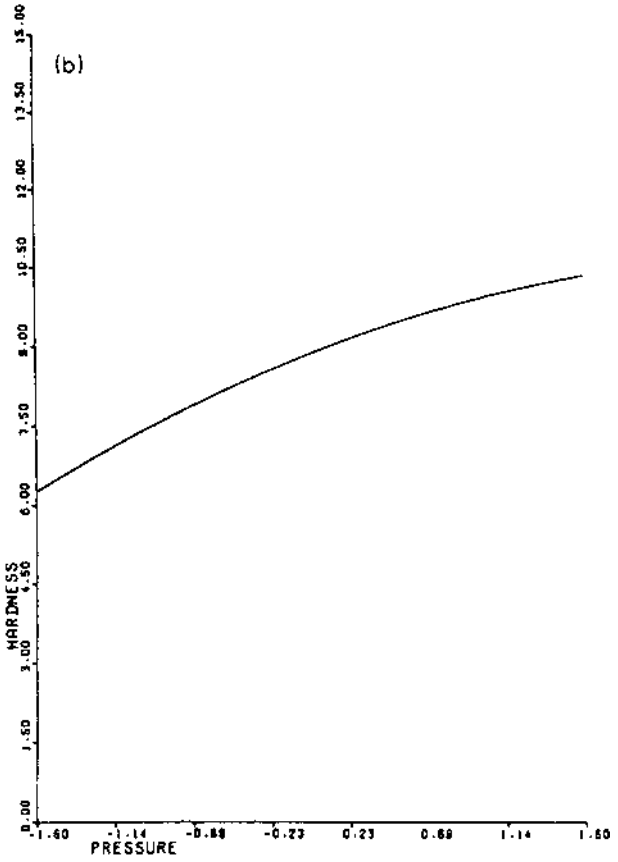


Fig. 11 (continued).

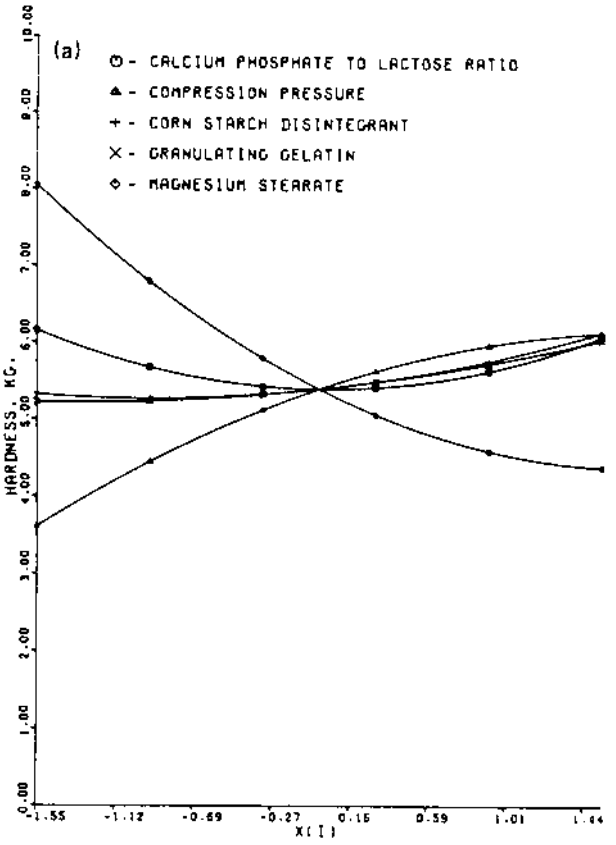


Fig. 12 Computer-generated composite plots.

The last step, which concerns the graphic techniques, may be requested at any time after the regression analysis has been performed and will probably be appropriate at several different stages of a project.

The key to successful application of the experimental optimization techniques is based on adequate experimental design. A system based on this experimental design (see Table 4), but utilizing a special analog computer for analysis, was presented by Claxton [19] as the Firestone Computer/Optimizer.

This approach demonstrates that use of only a part of this procedure will represent a step forward over the trial-and-error method of formula and process modification. It is not always necessary to carry these studies to completion. For example, once the designed experimentation has been completed, one might be able to accomplish the task simply by analyzing the graphs; therefore, further mathematical treatment or search programs will not be necessary. Some of the examples in the following section illustrate this fact.

E. Canonical Analysis

Canonical analysis, or canonical reduction, is a technique used to reduce a second-order regression equation, such as Eq. (9), to an equation consisting of a constant and squared terms, as follows:

$$Y = Y_0 + \lambda_1 W_1^2 + \lambda_2 W_2^2 + \lambda_3 W_3^2 + \cdots \tag{10}$$

The technique allows immediate interpretation of the regression equation by including the linear and interaction (cross-product) terms in the constant term (Y_0 or stationary point), thus simplifying the subsequent evaluation of the canonical form of the regression equation. The first report of canonical analysis in the statistical literature was by Box and Wilson [37] for determining optimal conditions in chemical reactions. Canonical analysis, or canonical reduction, was described as an efficient method to explore an empirical response surface to suggest areas for further experimentation. In canonical analysis or canonical reduction, second-order regression equations

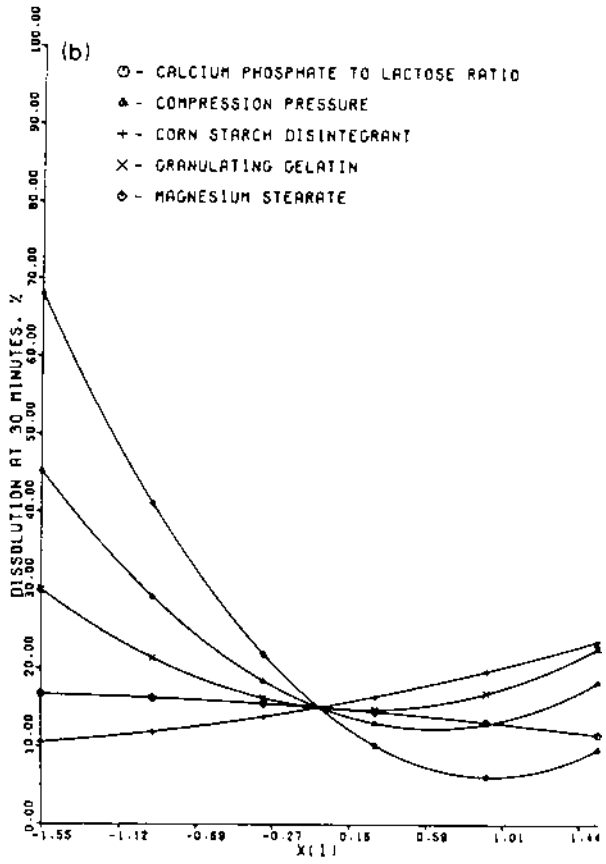


Fig. 12 (continued).

are reduced to a simpler form by a rigid rotation and translation of the response surface axes in multidimensional space, as shown in Fig. 14 for a two-dimension system. This mathematical technique, which makes use of eigenvalues and eigenvectors, is based in matrix algebra and is described in textbooks on response surface methodology [38,39].

A reported application of canonical analysis involved a novel combination of the canonical form of the regression equation with a computer-aided grid search technique to optimize controlled drug release from a pellet system prepared by extrusion and spheronization [28,29]. Formulation factors were used as independent variables, and in vitro dissolution was the main response, or dependent variable. Both a minimum and a maximum drug release rate was predicted and verified by preparation and testing of the predicted formulations. Excellent agreement between the predicted values and the actual values was evident for the four-component pellet system in this study.

VI. OTHER APPLICATIONS

In the last few years, optimization techniques have become more widely used in the pharmaceutical industry. Some of these have appeared in the literature, but a far greater number remain as "in-house" information, using the same techniques indicated in this chapter, but with modifications and computer programs specific to the particular company. An excellent review of the application of optimization techniques in the pharmaceutical sciences was published in 1981 [20]. This covers not only formulation and processing, but also analysis, clinical chemistry, and medicinal chemistry.

Designed experimentation, involving mostly some type or modification of factorial design, has been used to study many different types of formulation problems. These include a pharmaceutical suspension [21], a controlled-release tablet formulation [22], and a tablet-coating operation [23]. In the latter case, Dincer and Ozdurmus studied an enteric film coating and utilized the steepest descent graphic method to select the optimum.

Adaptation of the modified factorial techniques to desktop computers has also been accomplished [24, 25]. Down et al. [25] presented this concept and applied the programs to a tablet problem. The statistics involved were presented in some detail. A similar design was also used to study a high-performance liquid chromatography (HPLC) analysis [26]. In an unusual application, optimization techniques were even used to study the formulation of a culture medium in the field of virology [27].

Other applications of the previously described optimization techniques are beginning to appear regularly in the pharmaceutical literature. A literature search in *Chemical Abstracts* on process optimization in pharmaceuticals yielded 17 articles in the 1990–1993 time-frame. An additional 18 articles were found between 1985 and 1990 for the same narrow subject. This simple literature search indicates a resurgence in the use of optimization techniques in the pharmaceutical industry. In addition, these same techniques have been applied not only to the physical properties of a tablet formulation, but also to the biological properties and the in-vivo performance of the product [30,31]. In addition to the usual tablet properties the authors studied the following pharmacokinetic parameters: (a) time of the peak plasma concentration, (b) lag time, (c) absorption rate constant, and (d) elimination rate constant. The graphs in Fig. 15 show that for the drug hydrochlorothiazide, the time of the plasma peak and the absorption rate constant could, indeed, be

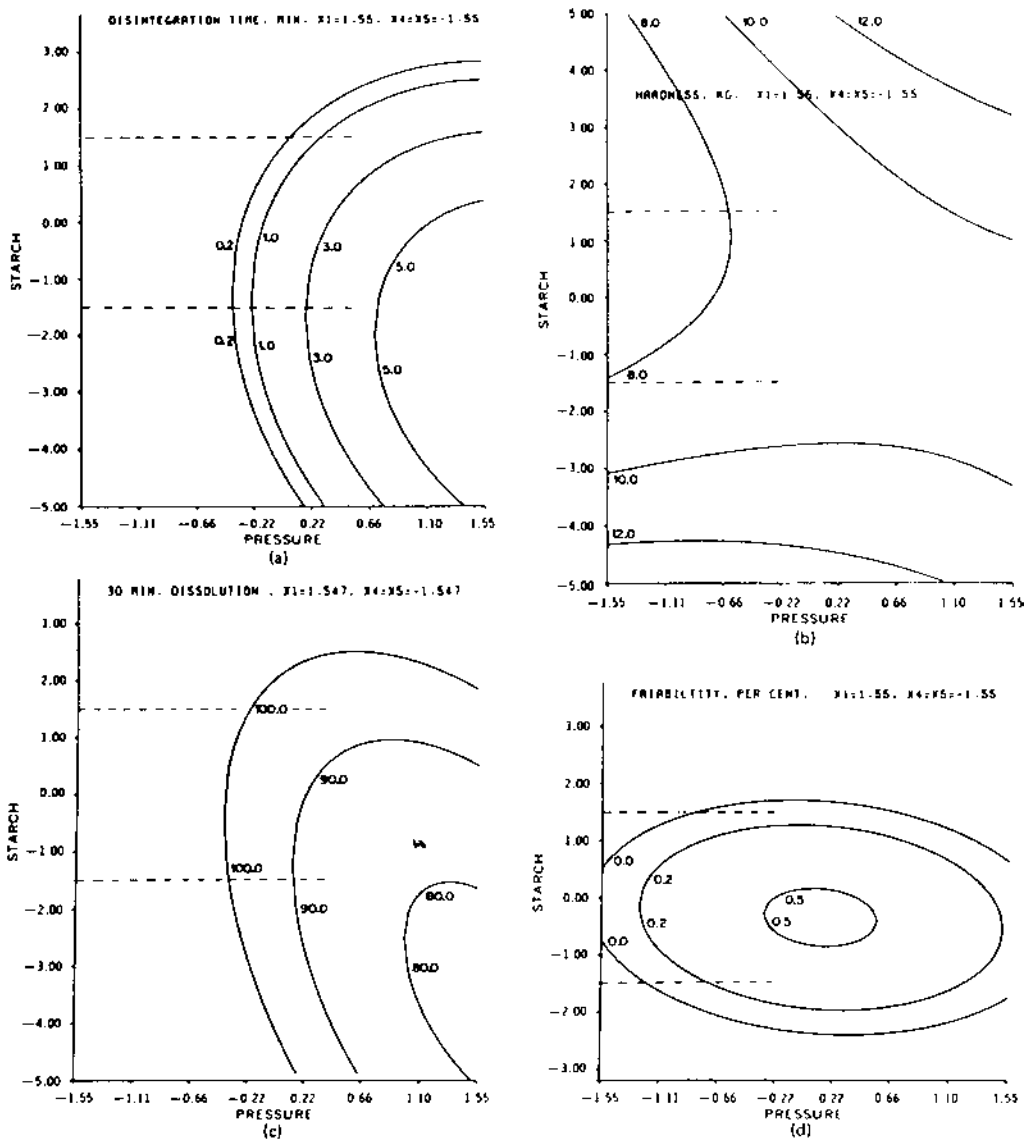


Fig. 13 Contour plots for (a) disintegration time; (b) tablet hardness; (c) dissolution response (%); (d) tablet friability as a function of disintegrant level and compressional force. Dashed lines on ordinate denote limits of experimental range (−1.547 to +1.547 eu; see text for details).

controlled by the formulation and processing variables involved.

VII. COMPUTERS AND SYSTEMS

It is obvious that the use of computers will facilitate the data analysis steps in the procedures discussed and will be needed for any mathematical analysis or search methods. In fact, a textbook has appeared describing the practical application of computer-aided optimization and provides direction for the implementation of these techniques to formulation [41]. Most of the examples presented have made use of computers in some way, and a few were completely performed by computer.

Several companies have adapted these experimental analysis techniques to computer software, but have kept the programs in-house. Representatives of a few

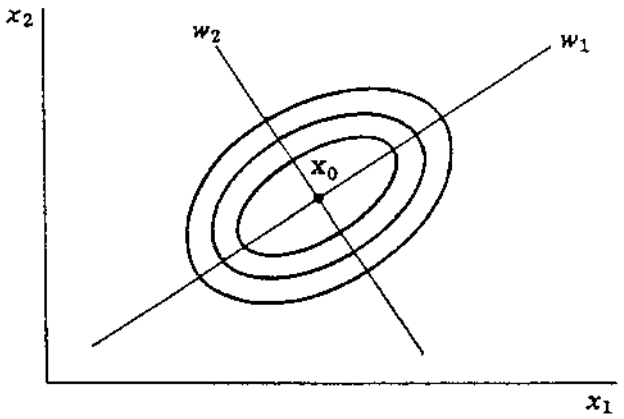


Fig. 14 Two-dimensional representation of the rigid rotation and translation involved in canonical analysis.

have, in fact, presented data at various conferences [24,32–34]. However, there are several commercially available programs that may be bought or licensed and several courses in experimentation address this subject.

The interested reader might be alerted to courses offered by the American Chemical Society, Dupont Corporation, or the Foremost Corporation. Some of these programs offer the use of statistical or response surface software. Specific computer packages are also available through Statistical Analysis Systems (SAS), IBM, and RS/Discover (RS1), which are designed for mainframe computers. The number of software packages available for standard desk-top office computers is large and is expected to increase. Several software packages—eCHIP, XStat, JMP, and Design Expert—are commonly used in the pharmaceutical industry, but these titles do not provide a complete list of available programs.

VIII. CONCLUDING REMARKS

As the list of applications illustrates, the techniques of optimization are not limited to tablets or even to solids. Any dosage form and any process should be amenable to this type of experimentation and analysis. From the most simple formulation to the most complicated one, there are ingredient levels and processing steps that can be varied, and any information on the result of such variation should be useful to the formulator.

Properly designed experimentation and subsequent analysis can not only lead to the optimum or most desirable product and process, but, if carried far enough, can shed light on the mechanism by which the independent variables affect the product properties. There are appropriate statistical techniques involving the use of selective regression analysis by which such analyses can be carried out [35]. Because this technique answers the question, "What independent variables most affect each response studied?" the application to selection of critical formulation and processing variables is obvious. This could provide supporting statistical evidence for the identification of critical variables in today's regulatory environment.

By appropriate analysis and generation of model (regression) equations (which are continuous), the formulator is able to select not the best of the formulations experimentally prepared, but the best within an experimental range; the optimum may be a combination of ingredients that the formulator has never prepared (and might never think to prepare). In the 30 years since the techniques of optimization were introduced to the pharmaceutical literature, the number of published studies on delivery systems has grown exponentially. There are numerous examples of the use of design of experiments, related statistical analysis, response surface methodology, and other methods for optimization in the recent literature. In many cases the techniques are used to study the variables in a system, rather than make any major changes.

Franz et al. [42] reviewed these techniques completely, along with statistical screening techniques and other experimental methods, with an excellent list of publications. A few selected publications from the recent literature demonstrate the wide variety of formulation and processing problems to which these techniques can be applied and the varying methods selected for optimization.

Porter et al. [43] applied the method to study the process variables in the tablet-coating operation. Remon et al. [44] studied high shear granulation and microwave drying to minimize dust production along with other responses using design of experiments, specifically a central composite design. Pujari and Chandra [45] reported on riboflavin production, optimizing the culture growth media via Plackett-Burmann screening methodology followed by factorial design.

Wu et al. [46] used the approach of an artificial neural network and applied it to drug release from osmotic pump tablets based on several coating parameters. Gabrielsson et al. [47] applied several different multivariate methods for both screening and optimization applied to the general topic of tablet formulation: they included principal component analysis and

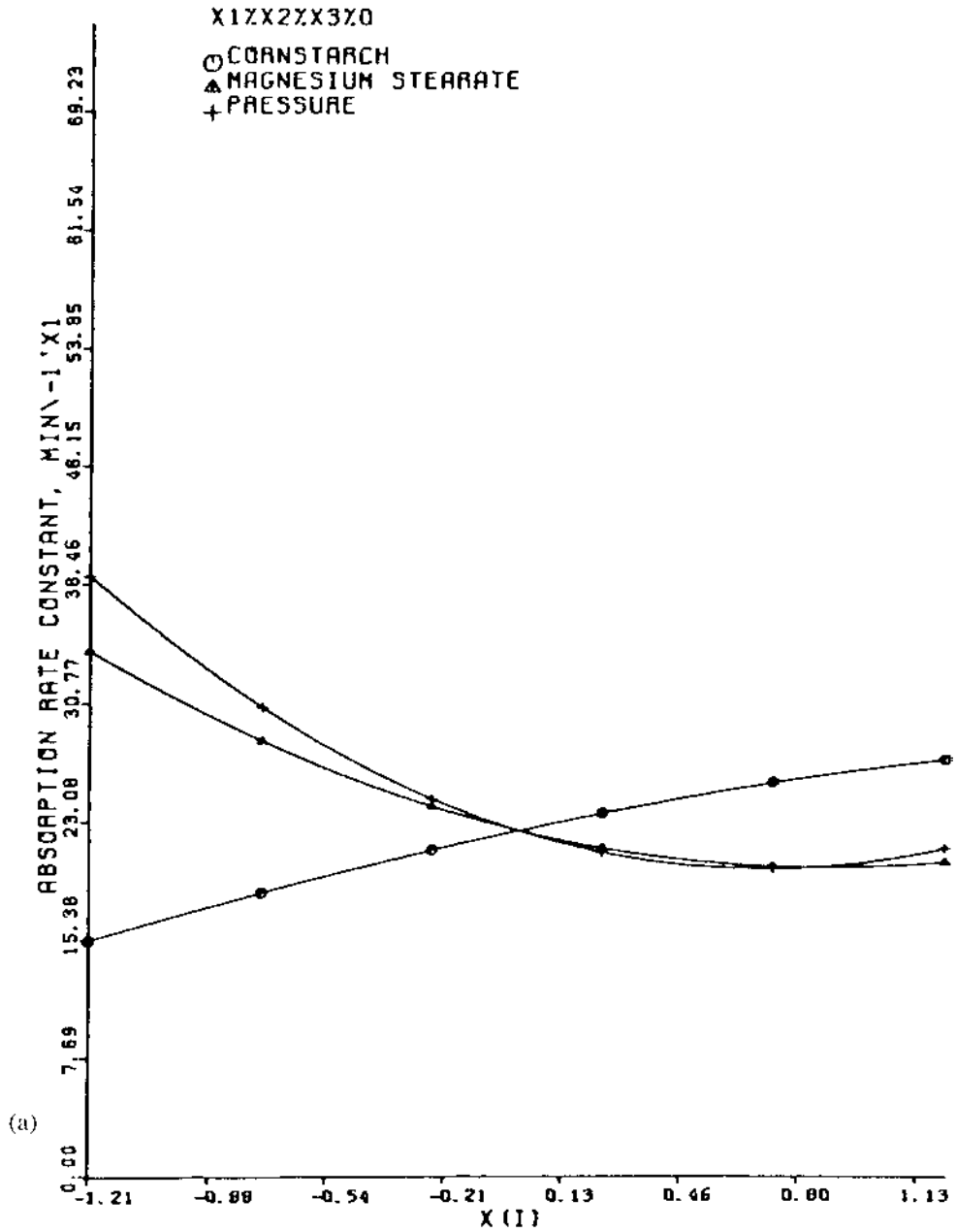
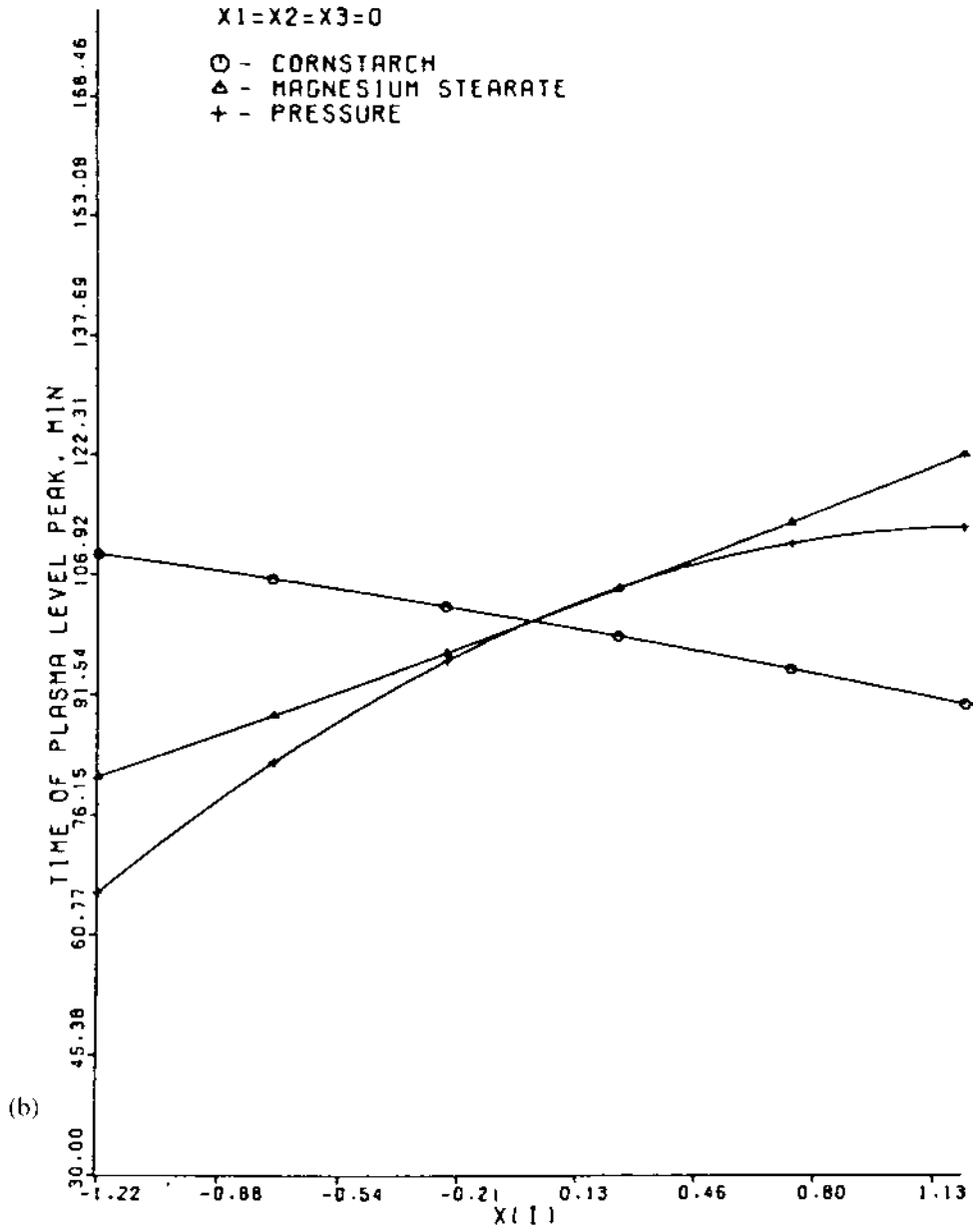


Fig. 15 Computer plots for (a) absorption rate constant and (b) time of plasma peak. (From Ref. 31.)

factorial design. Marti-Mestres et al. [48] studied submicron emulsions with sunscreens using simplex centroid design. Shiromani and Clair [49] performed a statistical comparison of high shear versus low shear granulation using a common formulation and a central composite design.

These techniques of optimization can be useful, even if selecting the optimum is not the primary objective. The formulator may have no intention of drastically changing a given formulation. Many times a very small change in processing or ingredient level can dramatically improve a particular property. The use of such information in "troubleshooting" situations has been demonstrated [36].

The independent variables have been, or should have been, selected by the formulator, and there is no substitute for experience. Experience with the system or with pharmaceutical systems in general can guide



(b)

Fig. 15 (continued).

the formulator to select those variables most likely to have an effect on those levels which are most practical. The results of an optimization study, especially the graphic output, can give direction for product improvement—no matter why the improvement is necessary or desirable.

Once experimental data have been collected and relationships generated by regression analysis (or even derived from first principles), the formulator has many options available for subsequent analysis, These need not be restricted to mathematical techniques or to elaborate computerized systems.

A side benefit of this designed type of experimentation is its potential usefulness in product and process validation. The subject of validation is of great interest to those in the operations area, but if approached rationally, validation must begin in the product development phase. The designs usually

selected lend themselves to the concept of processing limits and "challenge." The resulting data can be applied to scale-up, can aid in the transfer of information to the operations area, and should be the basis of the protocol design for validation.

The emphasis, once again, is that appropriate statistical design is an important consideration. For a formulator planning such a study, it should be noted that the independent variables can be anything that he or she can quantitate and control; and the dependent variables can be anything that he or she can quantitate. From the data resulting from the required number of experiments, one is able to generate a mathematical model to which the appropriate optimization technique is applied (e.g., graphic, mathematical, or the search method).

The final conclusion is the ultimate benefit: The more the formulator knows about a system, and the better that he or she can define it, the more closely it can be controlled.

REFERENCES

1. *Webster's New Collegiate Dictionary*, G. & C. Merriam, Springfield, MA, 1974.
2. L. Cooper and D. Steinberg, *Introduction to Methods of Optimization*, W. B. Saunders, Philadelphia, 1970.
3. O. L. Davis (ed.), *The Design and Analysis of Industrial Experiments*, Macmillan (Hafner Press), New York, 1967. [2nd ed., Longman, Harlow, Essex, England, 1978.]
4. W. G. Cochran and G. M. Cox, *Experimental Designs*, John Wiley & Sons, New York, 1957, Chap. 8A.
5. G. E. P. Box, W. G. Hunter, and I. S. Hunter, *Statistics for Experimenters: An Introduction of Data Analysis and Model Building*, John Wiley & Sons, New York, 1978.
6. S. N. Deming and P. G. King, Res./Dev., p. 22, May (1974).
7. M. H. Rubinstein, Manuf. Chem. Aerosol News, p. 30, Aug. (1974).
8. T. N. DiGaetano, Bull. Parenter. Drug Assoc., 29, 183 (1975).
9. W. Spendley, G. R. Hext, and F. R. Himsworth, Technometrics, 4, 441 (1962).
10. E. Shek, M. Ghani, and R. Jones, J. Pharm. Sci., 69, 1135 (1980).
11. S. T. Anik and L. Sukumar, J. Pharm. Sci., 70, 897 (1981).
12. C. Bindschaedler and R. Gurny, Pharm. Acta HeIv., 57(9), 251 (1982).
13. D. Janeczek, M. S. thesis, Philadelphia College of Pharmacy and Science (1979).
14. S. L. Morgan and S. N. Deming, Anal. Chem., 46, 1170 (1974).
15. D. E. Fonner, Jr., J. R. Buck, and G. S. Banker, J. Pharm. Sci., 59, 1587 (1970).
16. J. R. Buck, G. E. Peck, and G. S. Banker, Drug Dev. Commun., 1, 89 (1974–75).
17. J. B. Schwartz, J. R. Flamholz, and R. H. Press, J. Pharm. Sci., 62, 1165 (1973).
18. N. R. Bohidar, F. A. Restaino, and J. B. Schwartz, J. Pharm. Sci, 64, 966 (1975).
19. W. E. Claxton, paper presented to the Industrial Pharmaceutical Technology Section, American Pharmaceutical Association, Academy of Pharmaceutical Sciences, San Francisco meeting, Mar. 1971.
20. D. A. Doornbos, Pharm. Weekbl. Sd. Ed., 3, 33 (1981).
21. J. L. O'Neill and R. E. Dempski, through J. Soc. Cosmet. Chem., 32, 287 (Sept./Oct. 1981).
22. M. Harris, J. B. Schwartz, and J. McGinity, Drug Dev. Intl. Pharm., (in press).
23. S. Dincer and S. Ozdurmus, S. Pharm. Sci., 66, 1070 (1977).
24. D. R. Savello. K. J. Koenig, K. R. Nelson, J. B. Schwartz, and C. G. Thiel, paper presented to the American Pharmaceutical Association, Academy of Pharmaceutical Sciences. Atlanta meeting. Nov. 1975.
25. G. R. B. Down. R. A. Miller, S. K. Chopra, and J. F. Millar, Drug Dev. Iid. Pharm. 6, 311 (1980).
26. M. L. Cotton and G. R. B. Down, J. Chromatogr., 259, 17 (1983).
27. E. M. Scattergood, J. B. Schwartz, M. O. Villarejos, W. J. McAleer, and M. R. Hilleman. Drug Dev. Ind. Pharm., 9, 745 (1983).
28. R. E. O'Connor, N. R. Bohidar, and J. B. Schwartz, Optimization by Canonical Analysis Applied to controlled Release Pellets, Abstracts of the 1st National AAPA Meeting. Washington, DC. Nov. 2–6, 1986.
29. R. E. O'Connor, The Drug Release Mechanism and Optimization of a Microcrystalline Cellulose Pellet System, Ph.D. dissertation, Philadelphia College of Pharmacy & Science, 1987.
30. A. N. Karabelas, R. W. Mendes, and J. B. Schwartz, paper presented to the Industrial Pharmaceutical Technology Section, American Pharmaceutical Association, Academy of Pharmaceutical Sciences, San Antonio meeting, Nov. 1980.
31. J. B. Schwartz, J. Soc. Cosmet. Chem., 32, 287, Sept/Oct. (1981).
32. A. R. Lewis and R. Poska, paper presented to the Midwest Regional IPT meeting, American Pharmaceutical Association, Academy of Pharmaceutical Sciences, May 1981.
33. D. R. Savello, Workshop presentation, American Pharmaceutical Association, Academy of Pharmaceutical Sciences, Las Vegas meeting, April 1982.
34. S. E. Brown, paper presented at RxPo Conference, New York, June 1984.
35. N. R. Bohidar, F. A. Restaino, and J. B. Schwartz, Drug Dev. Ind. Pharm., 5, 1975 (1979).
36. J. B. Schwartz, J. R. Flamholz, and R. H. Press, J. Pharm. Sci., 62, 1518 (1973).

37. G. E. P. Box and K. B. Wilson, S. R. Statist. Soc. Ser. B, (Methodological), 13, 1 (1954).
38. G. E. P. Box and N. R. Draper, *Empirical Model Building and Response Surfaces*, John Wiley & Sons, New York, 1987.
39. R. H. Myers, *Response Surface Methodology*, Allyn & Bacon, Boston, 1971.
40. D. G. Kleinbaum, L. L. Kupper, and K. E. Muller, *Applied Regression Analysis and Other Multi-variate Methods*, PWS-Kent, Boston, 1988.
41. A. H. Bohl (ed.), *Computer-Aided Formulation*, VCH, New York, 1990.
42. R. M. Franz, D. M. Cooper, J. E. Brown, and A. R. Lewis, in: *Pharmaceutical Dosage Forms: Disperse Systems*, Vol. 1, 2nd ed., Marcel Dekker, New York, 1996.
43. S. C. Porter, R. P. Verseput, and C. R. Cunningham, Pharm. Tech., 21(10), 60 (1997).
44. F. Kiekens, M. Cordoba-Diaz, and J. P. Remon, DDIP, 25(12), 1289 (1999).
45. V. Pujari and T. S. Chandra, Proc. Biochem., 36, 31 (2000).
46. T. Wu, W. Pan, J. Chen and R. Zhang, DDIP, 26(2), 211 (2000).
47. J. Gabrielsson, A. Nystrom, and T. Lundstedt, DDIP, 26(3): 275 (2000).
48. G. Marti-Mestres, F. Niellood, R. Fortune, C. Fernandez, and H. Maillols, Drug Dev. Ind. Pharm., 26(3), 349 (2000).
49. P. K. Shiromani and J. Clair, Drug Dev. Ind. Pharm., 26(3), 357 (2000).