Experimental design in HPLC separation of pharmaceuticals

Jevrem Stojanović, Jovana Krmar, Ana Protić, Bojana Svrkota, Nevena Đajić, Biljana Otašević*

University of Belgrade – Faculty of Pharmacy, Department of Drug Analysis, Vojvode Stepe 450, 11221 Belgrade, Serbia

*Corresponding author: Biljana Otašević, e-mail: biljana.otasevic@pharmacy.bg.ac.rs

Abstract

Design of Experiments (DoE) is an indispensable tool in contemporary drug analysis as it simultaneously balances a number of chromatographic parameters to ensure optimal separation in High Pressure Liquid Chromatography (HPLC). This manuscript briefly outlines the theoretical background of the DoE and provides step-by-step instruction for its implementation in HPLC pharmaceutical practice. It particularly discusses the classification of various design types and their possibilities to rationalize the different stages of HPLC method development workflow, such as the selection of the most influential factors, factors optimization and assessment of the method robustness. Additionally, the application of the DoE-based Analytical Quality by Design (AQbD) concept in the LC method development has been summarized. Recent achievements in the use of DoE in the development of stability-indicating LC and hyphenated LC-MS methods have also been briefly reported. Performing of Quantitative structure retention relationship (QSRR) study enhanced with DoE-based data collection was recommended as a future perspective in description of retention in HPLC system.

Keywords: design of experiments, drug analysis, high pressure liquid chromatography, analytical method development and validation

doi.org/10.5937/arhfarm71-32480

Introduction

The main goal of pharmaceutical analysis is to ensure the quality of active pharmaceutical ingredients (APIs), excipients and/or drug products (DP) (1). Thus, drug analysis comprises many tasks, such as the assessment of pharmaceutical substance purity, identification and quantitation of APIs in dosage forms, testing the stability of APIs and DP, bioanalysis, etc. (2) As strict requirements for the quality of pharmaceuticals are prescribed, the method of analysis must provide reliable results. This is achieved through effective method development. Method validation is a procedure that shows the suitability of the developed analytical method for the intended purpose. It includes the estimation of method robustness, selectivity, accuracy and other validation parameters (3). Nowadays, a robust and reliable pharmaceutical method can be obtained by embedding the quality into method development stage. This concept is known as Analytical Quality by Design (AQbD) and its basic principles, as well as the benefits, are described in more detail in (4).

High pressure liquid chromatography (HPLC) is one of the irreplaceable techniques in modern pharmaceutical analysis. However, despite its widespread use in various fields of pharmaceutical research, optimization of reliable HPLC methods remains a rather challenging task. This claim finds its support in the observation that successful separation depends on the simultaneous control of numerous factors that affect the retention of analytes (5). For decades, this problem has been addressed by trial-and-error approach. However, this methodology is time-consuming, labor intensive and involves excessive consumption of reagents and toxic organic solvents. Most importantly, the optimal set of working conditions can never really be ensured (6, 7). In order to rationalize the examination of the experimental space during the HPLC condition optimization, chemometric approaches are increasingly being applied (8).

Chemometrics is a contemporary chemical sub-discipline that uses mathematics, statistics and formal logic to design the optimal experimental procedure. It provides global knowledge of the system under a study with a minimum number of carefully chosen experiments. Experimental design (also referred to as Design of Experiments; DoE) is the most popular chemometric discipline in the optimization and validation of chromatographic methods as well as in defining design spaces, an essential component of the AQbD paradigm (9, 10). To provide a state-of-the-art understanding of such a broad topic and thus encourage further use of chemometric techniques in drug analysis, this manuscript aims to provide a comprehensive overview of DoE applications in HPLC pharmaceutical research.

Design of experiments –theoretical background

DoE is a multivariate statistical technique intended for effective planning, performing and analyzing data from conducted experiments (11). In DoE studies, one or several input variables (factors) are deliberately changed to reveal their impact on the observed response (experimental outcome of interest). Compared to the traditional

approach that includes tuning one variable at a time (while keeping other parameters unchanged), the DoE methodology has a number of advantages. DoE methodology assesses the interaction between factors, provides global knowledge, delivers high-quality information at every point of the experimental domain and requires fewer experiments than classical approach. Leardi stated that the traditional, One-Factor-At-a-Time (OFAT) methodology should not be used in the future and gave illustrative examples from everyday life about its outdated nature (12).

In summary, the application of DoE in HPLC analysis is fueled by the ability to obtain objective and information-rich results with minimal experimental work. DoE can be used to facilitate chromatographic method development, method optimization and robustness testing.

Step-by-step protocol to ensure a successful application of DoE in HPLC

1) Define the aim of the experiment

According to the classical way of thinking, once you have obtained the results you can draw any information of interest. This approach also includes the belief that the number of experiments should be as large as possible in order to obtain reliable information. On the contrary, in DoE-supported studies the goal (examination of chromatographic behaviour of analyte, separation of closely eluted analytes) should be set before experiments are conducted. This rationalizes HPLC research and improves its quality (12).

2) Conduct preliminary studies

Preliminary experiments should be in line with the conventional protocol for the HPLC method development and validation. Hence, the aim of this step is to select the initial HPLC conditions, such as column type and mobile phase composition. Nevertheless, there is a growing number of studies (13, 14) in which the issue related to the selection of mobile phase is addressed using the DoE.

3) Select the experimental factors and their levels

The quality of DoE-supported HPLC research depends on the proper identification of all the important variables. If an important variable is not detected as a factor (input), irregular behavior of the analytical system can be observed. On the other hand, if an insignificant variable is identified as a factor, sequential stages of research may become unnecessarily complicated (10). A comprehensive overview of the factors that can affect HPLC separation is given in Figure 1. If the analytical process includes the derivatization or extraction, the factors associated with these steps should be investigated separately (7).



Figure 1.Factors affecting HPLC separation – adapted according to reference (9)Slika 1.Faktori bitni za razdvajanje u HPLC sistemu – prilagođeno prema referenci (9)

Within the DoE framework, HPLC experiments are performed at different levels of investigated variables. Factor level corresponds to its magnitude for numeric variables. Lower and higher levels of variables are coded as -1 and +1, respectively. Nominal level of a factor is coded as 0. Using the same scale for different factors makes it possible to compare their effects. Apart from this, coded values allow the representation of qualitative variables within DoE (12).

The number of factors and selection of corresponding levels depend on intended use of a particular design. In general, when DoE is used in method development or robustness testing, a large number of factors should be examined at two levels. In method optimization stage, however, only the most significant factors are investigated at least at three levels. The general procedure is subject to change, depending on the available time, relevant resources and the analyst's experience. Usually, the factor levels are placed symmetrically around the nominal level. The interval between them can be defined based on experience or precision, i.e., the uncertainty with which the value of a factor can be adjusted (15). According to Hibbert (11), the proper choice of factor level is more pivotal than the choice of design type.

4) Select the response

Response is the observed or measured quality or quantity value which we use to estimate factor's significance or which we want to optimize. It is very common to define multiple responses that need to be monitored during the experimentation.

In the screening phase of method development, responses such as retention factors and selectivity are most often measured. In the optimization phase, retention and resolution factors, desirability function and chromatographic response function can be monitored. In robustness testing, quantitative responses (content of APIs or by-products, peak height, peak area) or responses describing quality of separation (resolution factor, relative retention time, retention factor, number of theoretical plates, plate height, tailing factor, peak to valley ratio) can be chosen (7, 10, 16).

5) Plan the experiments

In this step, a convenient experimental design has to be chosen. As stated in (17), it is recommended that the utilized design has good statistical features, such as orthogonality and/or rotability. In orthogonal designs, factors are altered independently of each other (11). If there is no correlation between prediction variance and distance of the estimates from the design center, design is rotatable (18).

Based on the ultimate goal of HPLC experiments, the designs fall into two distinguished categories: screening designs and response surface designs. Screening designs are used to test the factors' impact on the response of interest. Since there are numerous factors that can potentially impact the HPLC retention, screening designs detect factors (and their interactions) with the biggest effects. Factors that do not have a significant impact are discarded, and only the remaining factors are thoroughly examined using a response surface design. Response surface designs are also known as optimization designs, as they are used to establish experimental condition at which the most desirable response and/or criteria (maximum, minimum or range) is obtained (7).

6) Perform the experiments

Experimental work is considered as less important in the DoE methodology, but necessary to conduct and generate results. To obtain reliable results, the experiments must be performed in a random order to minimize the effects of uncontrollable variables (12). It is also advisable to perform experiments in replicates, (replications for several times under the same experimental conditions) which are used to estimate the experimental error (4-6, 11).

7) Analyze the data obtained by the analytical procedure

In this step, experimental data should be converted into information required to draw valid conclusions about the HPLC system.

Statistical data analysis comprises regression analysis and analysis of variance (ANOVA). Regression analysis is used to assess the relationship between factors and response (8). Considering that many variables need to be fitted into the mathematical

model, the most commonly applied regression technique is the multiple-linear regression (MLR). In ANOVA, the significance of one or more factors is assessed by the F-test (for detailed instruction, see (19)). ANOVA is a useful tool in the validation of the derived model.

Graphical data analysis is a complementary tool for detecting significant factors and revealing the relationship between variables. In this regard, mechanistic interpretation of HPLC data is performed via 3D and 2D plots such as (half-)normal probability plots, Pareto charts, contour plots, response surface plots, perturbation plots and residual plots. Discrimination between significant and insignificant HPLC factors is readily performed by the first three listed chart types (20). Contour plots and response surface plots are graphical representations of the mathematical model that best fit the data. These graphs show how the response changes depending on the factors values and are used to find the optimal conditions for a HPLC separation (21-23). However, it is often the case that more than two factors are statistically significant. Thus, the decision on which two factors should be included into contour plot or response surface can be based on the perturbation plot (20,22). Residual plots may suggest the presence of outliers, i.e. the introduction of a systematic error (21).

Many steps in the DoE methodology can be easily done by an experienced analyst who is familiar with the basic principles of DoE. Nevertheless, it is easier to analyze the data using specialized computer programs or general statistical software. Popular software for the interpretation of chromatographic data are Design Expert (Stat-Ease Inc.), MINITAB (Minitab Inc.), MATLAB (The Mathworks Inc.), Statgraphics Centurion (Statpoint Technologies) and JMP (SAS Institute Inc.). Despite being promising free alternatives to the commercial programs, Chemoface and Develve have not been used in the field of HPLC research so far.

Mathematical modeling in DoE

Once the responses have been measured under varying conditions, multivariate data are fitted into linear or quadratic model in dependence on the predefined goal. Mathematical models, as stated before, are derived to explain the relationship between experimental factors and observed responses. They are usually described by the polynomial order (first-order model or linear model; second order-model or quadratic model). Linear and quadratic models for three factors (x_1 , x_2 and x_3) can be written using equations 1 and 2, respectively:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{23} x_2 x_3 + b_{13} x_1 x_3 + b_{123} x_1 x_2 x_3 + \varepsilon$$
(1)

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{11} x_1^2 + b_{22} x_2^2 + b_{33} x_3^2 + b_{12} x_1 x_2 + b_{23} x_2 x_3 + b_{13} x_1 x_3 + b_{123} x_1 x_2 x_3 + \varepsilon$$
(2)

In the given equations, y is a measured response, b_0 is intercept, b_1 , b_2 and b_3 are linear terms (first-order parameters), b_{12} , b_{23} and b_{13} are interaction parameters, b_{11} , b_{22} and b_{33} are quadratic terms (second-order parameters) and ε is a residual. Coefficients in

a model are calculated from the model matrix as explained in reference (24). The absolute value of a coefficient gives an insight into the magnitude of the factor's effect. A negative or positive coefficient indicates whether the response is decreased or increased by increase in factor's value. Residuals represent the differences between the observed and the predicted values of a given response. Ideally, all residuals are 0 and the model predicts perfectly in the investigated experimental domain.

If factors are investigated at two levels (screening designs), only a linear model can be fitted to data, no curvature is obtained, and hence neither the minimum nor the maximum can be found in 3D-graphs. The linear relationship is not suitable for choosing optimal experimental conditions, but it is very useful for assessing the significance of factor effects. If factors are investigated at three levels, the quadratic relationship can be fitted to data and significant curvature can be obtained. Significant curvature, arising from quadratic terms, allows the analyst to understand the behavior of the system and choose the optimal experimental conditions by finding the response minimum or maximum. Multicriteria decision making should be performed to find optimal conditions if there are more responses of interest, given that one model is used for just one response (25).

Classification of experimental designs applied in HPLC method development

DoE methodology encompasses a plethora of various experimental designs. To better understand their applicability in HPLC method development pipeline, these designs are classified as screening designs, optimization designs and mixture designs. The classification scheme is given in Figure 2. It is important to emphasize that this classification is not strict so that designs that are primarily intended for screening can be used for optimization and vice versa. It is left to the analyst to decide which design is appropriate in a particular case, taking into account the characteristics of experimental designs, the available resources and the aim of the experiments.

Screening designs

Two-level factorial designs are invaluable screening tools in HPLC pharmaceutical studies. Factorial-based design can be further classified as: Full Factorial Design (FFD), Fractional Factorial Design (FrFD) and Plackett-Burman Design (PBD).



Figure 2. Classification of experimental designs – adapted from reference (9)
Slika 2. Klasifikacija eksperimentalnih dizajna – prilagođeno prema referenci (9)

Full Factorial Design

In FFD, it is mandatory to perform n^k experiments, where *n* is the number of factor levels and *k* is the number of factors to be examined. FFD is composed of experiments at each combination of factor levels, which provides sufficient degrees of freedom to calculate all linear and interaction terms in a mathematical model. Experimental matrix and model matrix for 2^3 FFD are shown in Table I and II, respectively. Figures 3a and 3b show a graphical representation of the FFD with two and three variables, respectively (26).

ors)
C

Experiment number	A	В	C
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

Tabela I Matrica eksperimenata za 2³ puni faktorski dizajn (A, B i C označavaju faktore)

b ₀	b 1	b 2	b 3	b 12	b 13	b 23	b 123
 1	-1	-1	-1	1	1	1	-1
1	1	-1	-1	-1	-1	1	1
1	-1	1	-1	-1	1	-1	1
1	1	1	-1	1	-1	-1	-1
1	-1	-1	1	1	-1	-1	1
1	1	-1	1	-1	1	-1	-1
1	-1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1	1

Table IIModel matrix for 2³ full factorial designTabela IIModel matrica za 2³ puni faktorski dizajn



Figure 3.Two-level FFDwith 2 (a) and 3 factors (b)Slika 3.FFD na 2 nivoa za ispitivanje 2 (a) i 3 (b) faktora

Two-level full factorial design is typically applied to screen factors that might significantly affect the HPLC system. However, the utilization of FFD is accompanied by a considerable number of experiments (e.g. examination of six factors would demand 64 runs to be performed). To save on costly experiments, FFD matrix is frequently applied in the examination of a limited pool of independent variables (27). In cases where multiple factors (more than 4) need to be assessed, it is possible to safely condensate (reduce) the set of experiments and keep most of the information. In this regard, FrFD and PBD are the most commonly used reduced factorial designs in HPLC.

Fractional factorial design

The number of experiments in FrFD is defined by n^{k-p} , where *n* is the number of levels (usually 2), *k* is the number of factors and *p* is the size of the fraction. In FrFD the number of experiments should always exceed the number of factors. The FrFD cannot provide a sufficient degree of freedom to calculate all interaction terms. Instead, the effects of interaction are confounded (aliased) with the main effects. This is considered

acceptable during the screening phase aimed primarily at determining main effects (10). Experimental matrix and graphical representation of 2^{3-1} FrFD are shown in Table III and Figure 4, respectively.

Run	Α	В	C (A ^x B or - A ^x B)
1	-1	-1	1 or -1
2	1	-1	-1 or 1
3	-1	1	-1 or 1
4	1	1	1 or -1

Table IIIExperimental matrix for 2³⁻¹ fractional factorial design**Tabela III**Matrica eksperimenata za 2³⁻¹ frakcioni faktorski dizajn



Figure 4.Two-level FrFD with 3 factors – two possible subsets of FFDSlika 4.FrFD na 2 nivoa za ispitivanje 3 faktora – dva moguća podseta FFD

If the influence of the interactions of high order is negligible, FrFD makes a great tool for screening a large set of factors. Recently, reported studies used FrFD as screening tool in robust method development. The number of examined factors was typically between 4 and 7. The common degree of fractionation was between 1 and 3 (28-32).

Plackett-Burman design

PBD is a two-level reduced factorial design with favorable features, such as balance and orthogonality. In PBD, the total number of experiments (N) is always multiple of 4 runs (4, 8, 12, 16, 20 and 24), and with N experiments N-1 factors are examined. Thus, the appealing advantage of PBD is the ability to assess a large number of factors with a small number of runs (e.g. the examination of 11 factors requires only 12 experiments). Experimental matrix of PBD with 11 factors is displayed in Table IV. It can be built starting from the first row, usualy designated as generator (33).

Run	x_1	x_2	<i>X</i> 3	<i>X</i> 4	<i>X</i> 5	<i>X</i> 6	X 7	<i>X</i> 8	<i>X</i> 9	X 10	<i>x</i> 11
1	+	+	-	+	+	+	-	-	-	+	-
2	-	+	+	-	+	+	+	-	-	-	+
3	+	-	+	+	-	+	+	+	-	-	-
4	-	+	-	+	+	-	+	+	+	-	-
5	-	-	+	-	+	+	-	+	+	+	-
6	-	-	-	+	-	+	+	-	+	+	+
7	+	-	-	-	+	-	+	+	-	+	+
8	+	+	-	-	-	+	-	+	+	-	+
9	+	+	+	-	-	-	+	-	+	+	-
10	-	+	+	+	-	-	-	+	-	+	+
11	+	-	+	+	+	-	-	-	+	-	+
12	-	-	-	-	-	-	-	-	-	-	-

Table IVExperimental matrix for Plackett-Burman design for 11 factors**Tabela IV**Matrica eksperimenata za Plaket-Burmanov dizajn za 11 faktora

Given that PBD requires a specific number of factors (N-1), there are often not enough potentially significant variables to fill the experimental matrix. For this purpose, *dummy factors* are introduced. *Dummy factors* are imaginary factors, with no real chemical meaning, that cannot possibly affect the observed system. A change in response that is the consequence of *dummy factor* alternation (is it rainy outside or not, did you put your watch on your right or left hand) indicates that some important factor is not included into the examination or the existance of significant interactions between factors. The PBD cannot be used to estimate anything else but main factors. This is a sufficient level of information for screening design. On the other hand, the two-factor and higher-level factor interactions are confounded with main factors and could result in the significance of *dummy factors*. Therefore, *dummy factors* can be used in the estimation of experimental error on one side and in demasking the statistically significant interactions on the other. When using *dummy factors* in standard error estimation, at least three *dummy factors* should be included in the experimental plan (16).

Since PBD is only able to assess the main effects, it is favorably used in robustness testing, where the effects of the interaction are negligible due to the extremely narrow factor ranges. Commonly, 6-8 factors are assessed using PBD (34-38). In latest publications, PBD has been employed in the robustness testing of AQbD-supported LC methods (39) or as a screening tool (40-42).

Optimization designs

Response surface designs include: a) symmetrical designs and b) asymmetrical designs. The former category of designs explores a symmetrical experimental domain and contains a central point. Symmetrical designs are: Three-level FFD, Central Composite

design (CCD), Box-Behnken design (BBD), Taguchi design (TD), and Doehlert design. The latter category of designs examines factors at different numbers of levels causing the investigated domain to take asymmetrical shape. For instance, D-optimal design is an asymmetrical design. Designs that are used to study mixture variables exclusively, i.e. to optimize the composition of HPLC mobile phase, are denoted as Mixture designs (8).

Three-level Full Factorial Design

The three-level full factorial design can be used to optimize the HPLC methods (8,18). However, due to the great number of required experiments (3^k) , such an application is practical if only two or three factors (k = 2, 3) are of interest for the investigation (43).

Central composite design

Central composite design (CCD) is the most complex design used in the optimization of HPLC methods. It requires $2^{k-p} + 2k + C_p$ experiments, where k is the number of factors, p is the size of the fraction and C_p is the central point of the experimental domain. CCD combines full or fractional factorial design (2^{k-p}) , star design (2k axial points) and central point. Unlike screening designs, response surface designs have central points as their obligatory part. Replicates at central points can be added to estimate experimental error and perform ANOVA.

CCD can be performed as three- or five-level design, depending on the distance of axial points from the center (α). If $|\alpha|$ has a value 1, axial points are placed on the faces of the experimental domain. Accordingly, the three-level design is denoted as face-centered CCD. If $|\alpha|$ has value $2^{(k-p)/4}$, axial points are placed on the circle outside the FFD or FrFD experimental domain. This five-level design is referred to as circumscribed CCD. Circumscribed CCD is a rotatable design. Due to the possibility of missing some important data, circumscribed CCD is considered inappropriate when: axial points are very distant from the central point or when such extreme values of examined factors are not feasible or practical (26). Figure 5 shows a graphical representation of face-centered CCD (left) and circumscribed CCD (right) with 3 factors.



Figure 5. Central composite design with 3 factors (a) face-centered (b) circumscribed Slika 5. Centalni kompozicioni dizajn sa 3 faktora (a) ka centru orjentisani (b) kružni

CCD is particularly applied to facilitate the optimization of critical HPLC factors and their interactions (44-47). In most cases, 3 significant factors were optimized using CCD and 5 to 6 runs were performed at the central point of the experimental domain.

Studies that used CCD to assess the robustness of HPLC methods are nicely summarized in (15). When CCD is used for this purpose, the conclusion about method robustness can be drawn using the results from the optimization stage or a new design can be carried out separately.

Box-Behnken design

Box-Behnken design (BBD) is a response surface design at three levels. The total number of experiments in BBD is defined by $2k(k - 1) + C_p$. Compared to CCD, BBD requires fewer experiments (e.g. when 3 factors are studied, the required number of experiments in BBD is 13, while in CCD is 15). BBD is nearly rotatable design based on incomplete factorial design. The BBD experimental matrix for 3 factors is shown in (Table V). This design does not include experiments with all factors being examined at their highest or lowest levels, as can be seen from Figure 6a. BBD is applied when it is known that the optimum lies in the central area of the experimental domain or when experimental points with all factors at extreme levels are not useful or feasible (11, 25).

BBD is mainly used to optimize a HPLC system with 3 important experimental factors (21, 48-51). The experimental matrix is usually composed of 15-17 experiments due to 3-5 replications at the central point. The recently published study on the HPLC separation of levofloxacin and ciprofloxacinis is particularly interesting because 6 factors have been successfully optimized via BBD (52).

Run	X 1	X2	X3
1	-1	-1	0
2	1	-1	0
3	-1	1	0
4	1	1	0
5	-1	0	-1
6	1	0	-1
7	-1	0	1
8	1	0	1
9	0	-1	-1
10	0	1	-1
11	0	-1	1
12	0	1	1
13	0	0	0

Table VExperimental matrix in Box-Behnken design for 3 factorsTabela VMatrica eksperimenata za Boks-Benkenov dizajn za 3 faktora



Figure 6. (a) Box-Behnken design and (b) Doehlert design for 3 factorsSlika 6. (a) Boks-Benkenov dizajn i (b) Dolertov dizajn za ispitivanje 3 faktora

Doehlert design

Doehlert design (uniform shell design) is a non-rotatable design that allows the analyst to choose a different number of levels for different factors. Factors that are considered more important than others can be assigned a higher number of levels. Also, it is a beneficial property in case of practical limitations (e.g., limited resources) (25). In the latter situation, it is possible to simply "move" to another part of the experimental domain. Nevertheless, this is not a good approach if it omits important data from the experimental domain. Doehlert design requires $n^2 + n + C_p$ experiments (where, *n* is the number of factors and $C_p is$ the number of replicates in the central point). Thus, it is a more economical design than CCD or BBD. Graphical representation of Doehlert design for three factors is represented in Figure 6b.

Literature survey revealed that Doehlert design was used to optimize mobile phase composition and mobile phase flow rate in most LC studies (53-56).

D-optimal design

The experimental limitations can be an obstacle to the application of symmetrical optimisation designs. In this case, D-optimal can be a valuable soultion since it works well with both symmetric and assymetric experimental regions. The particular benefits of this design lie in its ability to estimate quantitative and qualitative factors and to predefine the mathematical model. The minimum number of experiments in the D-optimal design is equal to the number of coefficients in the postulated model (usually quadratic model without interaction terms). The set of experiments to be performed is selected from the total number of possible combinations of factor levels. This information-rich subset is a combination of experiments with the maximal determinant for $X^T X$ (X is model matrix and X^T is its transpose). When the maximal determinant is accomplished, the precision of the mathematical model coefficients is assured even before the experimentation and statistical evaluation is performed (12). This is a prominent characteristic of D-optimal design. The information-rich subset of data formation (possible factor combinations)

would be time-consuming and hard task if all the possible subsets would be evaluated and compared. In order to simplify this step and, at the same time, provide precise and accurate model coefficients, different algorithms are used. Some of them are Fedorov and Mitchel (9, 12, 16).

A more comprehensive study on D-optimal design can be found in (25). The use of D-optimal design in HPLC method optimization was reported in (57, 58).

Mixture designs

Mixture designs are used to optimize mixtures such as mobile phases in HPLC. The limitation of the design is reflected in fact that the sum of all components of the mixture must be 1 (100%) at each point of the experimental domain. Thus, the obtained mathematical model does not contain the intercept term, because it is not possible to obtain a mixture in which the content of all constituents is 0%. The terms of the mathematical model are in poor correlation with the factor effects and should not be used for their assessment. Graphical representation of a three-components mixture is an equilateral triangle, whose vertices refer to single components, sides to the binary mixtures and internal points to the ternary mixtures (26). Mixture designs were used with great advantage in the optimization of the mobile phase composition, as specified in (59, 60).

Other designs

Taguchi design is usually constructed as PBD. It comprises inner and outer arrays which can be used to evaluate the experimental condition that are less likely to be affected by environmental factors (61). Nowadays, Taguchi design is preferentially used as a screening tool in AQbD-based LC studies (62, 63). Sequential designs are intended for use in preliminary studies when the analyst isn't familiar with the part of experimental domain that should be examined (7). Supersatured designs are designs that are more economical than FrFD or PBD because they can estimate at least *Nss* factors with *Nss* experiments. However, these designs are rarely used in the screening phase, given that all main effects are confounded (25). Nested designs are considered the most suitable designs for ruggedness testing (15). Split-plot designs have not yet been applied in HPLC analysis of pharmaceuticals (7).

Regulatory aspects related to DoE in HPLC analysis

Quality-by-Design (QbD) is a specific concept used in drug development with the ultimate goal of incorporating quality into the final product. Due to the significant role that analytical procedures play in the estimation of that quality, the extension of QbD paradigm to the development of pharmaceutical methods has been promoted through several regulatory documents, namely, ICH Q8 (64), Q9 (65), Q10 (66) and Q11 (67). In a relatively new guideline, published by the FDA in 2015 (68), it is stated that robustness of analytical method for drugs and biologics should be evaluated by a systematic approach that includes initial risk assessment and multivariate experimental approach, such as DoE.

However, this guideline is made as a set of recommendations and does not provide mandatory legal responsibilities.

The QbD principles applied to the development of analytical methods are denoted as Analytical QbD (AQbD). The product of AQbD approach is a highly robust, fully understood method that constantly provides performance that has been predefined (69, 70). AQbD key components include the identification of Analytical Target Profile (ATP), Critical Method Attributes (CMA) with risk assessment, Critical Proces Parameters (CPP), method optimization, definition of Design Space (DS), control strategy and Continuous Method Monitoring (CMM). Although several statistical tools can be employed to support AQbD, as explained in reference (4), DoE methodology is recognized as the most valuable one.

Literature survey revealed tremendous number of HPLC methods that were developed using DoE-based AQbD approach in recent years (23, 62, 69, 71-73). In most studies, CCD and D-optimal designs were favorably used to optimize HPLC systems with 2 critical method parameters (CMPs) and BBD in case of 3 or 4 CMPs.

The definition of optimal separation conditions satisfying predefined criteria in DoE-based multiresponse HPLC method development is usually achieved by contour plot or response surface overlay. The shortage of such methodology lies in the fact that the risk of obtaining undesired results due to errors of calculating coefficients of mathematical models and the capability of retention prediction in every point of experimental region resulting from mathematical models and are not evaluated. Additional risk calulations by means of uniform error distribution equal to the calculated standard error and its addition to the estimates of model coefficients as well as the propagation of prediction errors to the predefined quality criterion using Monte Carlo simulations is considered as a more AQbD compliant strategy (69).

Contemporary applications of DoE in HPLC

The development of a stability-indicating LC method is a challenging process, due to the use of complex, multicomponent samples. These samples are formed after the exposure of the API(s) and/or pharmaceutical product to various stress conditions. So-called stress samples contain several unknown degradation products apart from API(s). In this regard, the complexity of the method development process is attributed to the fact that considerable number of chromatographic parameters has to be fine-tuned in order to achieve optimal separation of all analytes (8, 70). Nowadays, the DoE-based AQbD approach is favorably used in the development of stability-indicating LC methods. For instance, in (74) development of stability-indicating method for cloxacillin was supported by PBD (screening of CMPs) and combined mixture-process variable design (optimization of 3 the most significant CMPs). In (69) BBD-based AQbD approach was used in the development of micellar HPLC method for the analysis of cilazapril and hydrochlorothiazide in dosage form. A detailed overview of DoE-AQbD supported development of stability-indicating HPLC methods can be found in (70).

Hyphenation of HPLC with mass spectrometry (MS) has made one of the major breakthroughs in pharmaceutical research. However, the separation and detection of compounds have become trickier due to the larger set of instrumental parameters that has to be optimized. To address this problem in LC-MS analysis of genotoxic impurity of interest, Székely et al. (75) applied the DoE. All significant factors, including mobile phase flow rate, gradient, injection volume (LC parameters), collision energy and cone voltage (MS parameters) were examined at three levels using CCD. The newest research articles are reporting application of Design Space accompanied with PBD and CCD for the development of LC-MS method for simultaneous analysis of five statins (76).

Modified LC systems often represent a technique of choice for the separation of complex pharmaceutical mixtures or API(s) of highly similar retention behavior. Linear or quadratic polynomial models usually fail to explain the retention behavior of analytes in HPLC systems of greater complexity. In such case, it might be useful to apply artificial neural networks (ANN) as a modeling technique. ANN is great tool for modeling nonlinear patterns between variables. Novotná etal. (77) used DoE (CCD) in combination with ANN to achieve optimal ion-pair HPLC separation of neuroprotective peptides. More recently, Maljurić et al. (78) developed a green LC method for the separation of selected antipsychotics by performing ANN-QSRR (Quantitative Structure-Retention Relationship) study. The green LC concept was introduced by partial replacement of toxic solvent acetonitrile with cyclodextrin in the mobile phase. QSRR studies were used for prediction of retention and optimization of separation conditions. Since QSRR studies relay on physical-chemical properties of analytes represented by molecular descriptors, the molecular interactions as well as molecular mechanisms of separations in given separation system could also be assessed. Having in mind notable complexity of system cyclodextrin-modified LC and the data set comprising of analyte specific, instrumental and chromatographic factors, the combination of CCD-based data collection, ANN modeling technique and QSRR as chemometric data processing tool successfully contributed to thorough understanding of analytes' retention behavior in cyclodextrinmodified LC system.

Conclusion

DoE methodology has shown overall advantage over the OFAT approach. In the future, regulatory requirements for method development and validation in pharmaceutical analysis are expected to be stricter. Thus, the use of DoE methodologies will be the only valid way to develop efficient and robust methods. Suitable experimental design should be chosen on a case-by-case basis, taking into account advantages, disadvantages and special features provided by a particularl design, as explained in this manuscript.

Acknowledgment

This study has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, contract number 451-03-9/2021-14/200161.

References

- 1. Ahuja S. Overview of modern pharmaceutical analysis. Sep. Sci. Technol. 2011;10:1-9.
- 2. Watson D. Pharmaceutical Analysis 5th Ed: A Textbook for Pharmacy Students and Pharmaceutical Chemists. Edinburgh: Churchill Livingstone; 2020; p. 480.
- 3. Lobrutto R, Patel T. Method validation in Kazakevich Y, Lobrutto R,(ed). HPLC for pharmaceutical scientists. 1st ed. John Wiley & Sons, Inc. 2007; p. 455-502.
- 4. Vogt FG, Kord AS. Development of quality-by-design analytical methods. J Pharm Sci. 2011;100(3):797-812.
- 5. Mattrey FT, et al. Current challenges and future prospects in chromatographic method development for pharmaceutical research. TrAC-Trend Anal Chem. 2017;95:36-46.
- 6. Baczek T. Computer-assisted optimization of liquid chromatography separations of drugs and related substances. Curr Pharm Anal. 2008;4(3):151-161.
- Ganorkar SB, Shirkhedkar AA. Design of experiments in liquid chromatography (HPLC) analysis of pharmaceuticals: Analytics, applications, implications and future prospects. Rev Anal Chem. 2017;36(3):1-43.
- Sahu PK, et al. An overview of experimental designs in HPLC method development and validation. J. Pharm. Biomed. Anal. 2018;147:590-611.
- 9. Komsta Ł, et al. J. Chemometrics in chromatography. CRC Press. 2018.
- 10. Stojanović B. Factorial-based designs in liquid chromatography. Chromatographia. 2013;76(5-6):227-240.
- 11. Hibbert DB. Experimental design in chromatography: a tutorial review. J Chromatogr B. 2012;910:2-13.
- 12. Leardi R. Experimental design in chemistry: A tutorial. Anal Chim Acta. 2009;652(1-2):161-172.
- Hadjmohammadi MR, Ebrahimi P. Optimization of the separation of anticonvulsant agents in mixed micellar liquid chromatography by experimental design and regression models. Anal Chim Acta. 2004;516(1-2):141-148.
- 14. Vemić A, et al. Chaotropic agents in liquid chromatographic method development for the simultaneous analysis of levodopa, carbidopa, entacapone and their impurities. J Pharm Biomed Anal. 2013;77:9-15.
- 15. Dejaegher B, Heyden YV. Ruggedness and robustness testing. J Chromatogr A. 2007;1158:138-157.
- Heyden VY, et al. Guidance for robustness/ruggedness tests in method validation. J Pharm Biomed Sci. 2001;24:723-753.
- 17. Debrus B, et al. Application of new methodologies based on design of experiments, independent component analysis and design space for robust optimization in liquid chromatography. Anal Chim Acta. 2011;691(1-2):33-42.

- Ferreira SLC, et al. Statistical designs and response surface techniques for the optimization of chromatographic systems. J Chromatogr A. 2007;1158(1-2):2-14.
- 19. Bezzera MA, et al. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta. 2008;76: 965-977.
- Attimarad M, et al. Validation of rapid RP-HPLC method for concurrent quantification of amlodipine and celecoxib in pure and formulation using an experimental design. Microchem J. 2020;152:104365.
- Dalvi AV, et al. Design of experiments-based RP-HPLC bioanalytical method development for estimation of Rufinamide in rat plasma and brain and its application in pharmacokinetic study. J Chromatogr B. 2018;1102:74-82.
- Panda SS, et al. Analytical procedure development: Concept to application for chemometry based ultrafast LC estimation of pimavanserin in pharmaceuticals. J Liq Chromatogr R T. 2020;43(3-4):118-130.
- 23. Yabré M, et al. Development of a green HPLC method for the analysis of artesunate and amodiaquine impurities using Quality by Design. J Pharm Biomed Anal. 2020; 190: 113507.
- 24. Brereton RG. Chemometrics-data analysis for the laboratory and chemical Plant. Chichester: Wiley. 2003; p. 489.
- 25. Dejaegher B, Heyden YV. Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. J Pharm Biomed Sci. 2011;56:141-158.
- 26. Leardi R. Experimental design. In: Data Handling in Science and Technology. Elsevier. 2013. p. 9-53.
- 27. Vujić Z, et al. Simultaneous analysis of irbesartan and hydrochlorothiazide: an improved HPLC method with the aid of a chemometric protocol. Molecules. 2012;17(3):3461-3474.
- 28. Jain A, et al. Application of chemometric approach for QbD-enabled development and validation of an RP-HPLC method for estimation of methotrexate. J Liq Chromatogr R T. 2019;42(15-16):502-512.
- 29. Beg S, et al. QbD-driven development and validation of an efficient bioanalytical UPLC method for estimation of olmesartan medoxomil. J Liq Chromatogr R T. 2016;39:587-597.
- Habib AA, Hammad SF, Megahed SM, Kamal AH. Innovative Quality by Design Approach for Development of Green Micellar HPLC Method for Simultaneous Determination of Atorvastatin and Amlodipine. Chromatographia. 2020;83(10):1221-1231.
- 31. Vermeij TAC, Edelbroek PM. Robust isocratic high performance liquid chromatographic method for simultaneous determination of seven antiepileptic drugs including lamotrigine, oxcarbazepine and zonisamide in serum after solid-phase extraction. J Chromatogr B. 2007;857(1):40-46.
- Mendia OG, et al. Efficient method development and validation for the determination of cardiovascular drugs in human plasma by SPE–UHPLC–PDA–FLD. Chromatographia. 2017;80(4):605-615.
- 33. Armstrong, N. Anthony. Pharmaceutical experimental design and interpretation. CRC Press, 2006.
- Nguyet ANM, Tallieu L, Plaizier JV, Massart DL, Heyden YV. Validation of an HPLC method on short columns to assay ketoconazole and formaldehyde in shampoo. J Pharm Biomed Anal. 2003;32(1):1-19.
- 35. Van Nederkassel AM, et al. Fast separations on monolithic silica columns: method transfer, robustness and column ageing for some case studies. J Pharm Biomed Anal. 2003;32(2):233-49.

- Song Q, Putcha L. Quantitation of promethazine and metabolites in urine samples using on-line solid-phase extraction and column-switching. J Chromatogr B. 2001;763(1-2):9-20.
- 37. Sun SW, Hsiu-Ting. Validated HPLC method for determination of sennosides A and B in senna tablets. J Pharm. Biomed. Anal. 2002;29(5):881-894.
- Kojic-Marinkovic S, Tumbas M, Rakic T, Jancic-Stojanovic B. Plackett-Burman design in robustness testing of liquid chromatographic method for determination of sodium-valproat. Arh Farm. 2014;64(2):128–43.
- Mitrović M, et al. Analytical quality by design development of an ecologically acceptable enantioselective HPLC method for timolol maleate enantiomeric purity testing on ovomucoid chiral stationary phase. J Pharm Biomed Anal. 2020;180:113034.
- 40. Abdel-Moety EM, Ezzat M, et al. A combined approach of green chemistry and Quality-by-Design for sustainable and robust analysis of two newly introduced pharmaceutical formulations treating benign prostate hyperplasia. Microcheml J. 2021;160:105711.
- 41. Beg S, et al. Applications of Monte-Carlo simulation and chemometric techniques for development of bioanalytical liquid chromatography method for estimation of rosuvastatin calcium. J Liq Chromatogr R T. 2017;40(18):907-920.
- 42. Aboushady D, Parr MK, Hanafi RS. Quality-by-Design Is a Tool for Quality Assurance in the Assessment of Enantioseparation of a Model Active Pharmaceutical Ingredient. Pharmaceuticals. 2020;13(11):364.
- 43. Nemutlu E, et al. Simultaneous multiresponse optimization of an HPLC method to separate seven cephalosporins in plasma and amniotic fluid: application to validation and quantification of cefepime, cefixime and cefoperazone. Talanta. 2009;80(1):117-126.
- 44. Dinç-zor Ş, et al. Chemometric optimization of an HPLC method for the simultaneous analysis of a multi component drug product by the help of central composite design. Microchem J. 2020;152:104322.
- 45. Quiming NS, et al. Chromatographic behavior of uric acid and methyl uric acids on a diol column in HILIC. Chromatographia. 2008;67(7):507-515.
- 46. Li P, et al. Optimizing ultraperformance liquid chromatographic analysis of 10 diterpenoid compounds in Salvia miltiorrhiza using central composite design. J Agr Food Chem. 2008;56(4):1164-1171.
- 47. Sivakumar T, et al. Multi-criteria decision making approach and experimental design as chemometric tools to optimize HPLC separation of domperidone and pantoprazole. J Pharm Biomed Anal. 2007;43(5):1842-1848.
- 48. Attimarad M, et al. Development and Validation of Rapid RP-HPLC and Green Second-Derivative UV Spectroscopic Methods for Simultaneous Quantification of Metformin and Remogliflozin in Formulation Using Experimental Design. Separations. 2020;7(4):59.
- 49. Dragomiroiu GTAB, et al. The development and validation of a rapid HPLC method for determination of piroxicam. Farmacia. 2015;63(1):123-31.
- Chaudhari SR, Shirkhedkar AA. Design of experiment avenue for development and validation of RP-HPLC-PDA method for determination of apremilast in bulk and in in-house tablet formulation. J Anal Sci Tech. 2019;10(1):1-9.

- Maljurić N, et al. A new strategy for development of eco-friendly RP-HPLC method using Corona Charged Aerosol Detector and its application for simultaneous analysis of risperidone and its related impurities. Microchem J. 2020;153:104394.
- 52. Czyrski A, Sznura J. The application of Box-Behnken-Design in the optimization of HPLC separation of fluoroquinolones. Sci Rep. 2019;9(1):1-10.
- 53. Jebali S, et al. Application of Factorial and Doehlert Designs for the Optimization of the Simultaneous Separation and Determination of Antimigraine Drugs in Pharmaceutical Formulations by RP-HPLC-UV. Int J Anal Chem. 2019; DOI: 10.1155/2019/9685750
- 54. Bonfilio R, et al. Multivariate optimization and validation of an analytical methodology by RP-HPLC for the determination of losartan potassium in capsules. Talanta. 2009;80:236-41.
- 55. Lafossas C, et al. Analysis of the retention of tetracyclines on reversed-phase columns: Chemometrics, design of experiments and quantitative structure-property relationship (QSPR) study for interpretation and optimization. Talanta. 2019;198:550-559.
- 56. Araujo P, Janagap S. Doehlert uniform shell designs and chromatography. J Chromatogr B. 2012;910:14-21.
- 57. Herrero A, et al. A new multiresponse optimization approach in combination with a D-Optimal experimental design for the determination of biogenic amines in fish by HPLC-FLD. Anal Chim Acta. 2016;945:31-38.
- 58. Golabchifar A, et al. Optimization of the simultaneous determination of imatinib and its major metabolite, CGP74588, in human plasma by a rapid HPLC method using D-optimal experimental design. Talanta. 2011;85(5):2320-2329.
- 59. Vanbel PF, Tilquin BL, Schoenmakers PJ. Criteria for optimizing the separation of target analytes in complex chromatograms. Chemometr Intell Lab. 1996;35(1):67-86.
- 60. Imam SS, et al. Optimization of mobile phase by 3 2-mixture design for the validation and quantification of risperidone in bulk and pharmaceutical formulations using RP-HPLC. Anal Method. 2014;6(1):282-288.
- 61. Deming SN, Morgan SL. Experimental Design: A Chemometric Approach 2nd ed. Elsevier Science; 1993: p. 436.
- Saini Sumant, et al. QbD-steered development and validation of an RP-HPLC method for quantification of ferulic acid: Rational application of chemometric tools. J Chromatogr B. 2020;1155:122300. DOI: 10.1016/j.jchromb.2020.122300
- 63. Beg S, et al. Development and validation of a stability-indicating liquid chromatographic method for estimating olmesartan medoxomil using quality by design. J Chromatogr Sci. 2015;53(7):1048-1059.
- 64. ICH Harmonised Tripartite Guideline: Pharmaceutical Development Q8(R2), current Step 4 version; International Conference on Harmonisation: Geneva, 2009.
- 65. ICH Harmonised Tripartite Guideline: Quality Risk Management Q9, current Step 4 version; International Conference on Harmonisation: Geneva, 2005.
- 66. ICH Harmonised Tripartite Guideline: Pharmaceutical Quality System Q10, current Step 4 version; International Conference on Harmonisation: Geneva, 2008.
- 67. ICH Harmonised Tripartite Guideline: Development and Manufacture of Drug Substances Q11, current Step 4 version; International Conference on Harmonisation: Geneva, 2012.

- Analytical Procedures and Methods Validation for Drugs and Biologics Guidance for Industry; U.S. Food and Drug Administration Silver Spring, MD, 2015.
- Otašević B, et al. Comparison of AQbD and grid point search methodology in the development of micellar HPLC method for the analysis of cilazapril and hydrochlorothiazide dosage form stability. Microchem J. 2019;145:655-663.
- 70. Tome T, et al. Development and optimization of liquid chromatography analytical methods by using AQbD principles: Overview and recent advances. Org Process Res Dev. 2019;23(9):1784-1802.
- 71. Sylvester B, et al. A Quality by Design (QbD) approach to the development of a gradient highperformance liquid chromatography for the simultaneous assay of curcuminoids and doxorubicin from long-circulating liposomes. J Pharm Biomed Anal. 2018;158:395-404.
- Patel MN, Kothari CS. Multivariate approaches for simultaneous determination of avanafil and dapoxetine by UV chemometrics and HPLC-QbD in binary mixtures and pharmaceutical product. J AOAC Int. 2016;99(3):649-663.
- 73. Panda SS, et al. Analytical eco-scale and quality by design-oriented liquid chromatography method for simultaneous quantification of metoprolol succinate, telmisartan, and cilnidipine in their fixed-dose combination. Sep Sci Plus. 2021;4(3):128-136.
- 74. Zhang X, Hu C. Application of quality by design concept to develop a dual gradient elution stabilityindicating method for cloxacillin forced degradation studies using combined mixture-process variable models. J Chromatogr A. 2017;1514:44-53.
- 75. Székely Gy, et al. Design of experiments as a tool for LC–MS/MS method development for the trace analysis of the potentially genotoxic 4-dimethylaminopyridine impurity in glucocorticoids. J Pharm Biomed Anal. 2012;70:251-258.
- 76. Belmir H, Abbouriche A, Bennamara A, Saffaj T, Ihssane B. Using Design Space and Response Surface Methodology for developing a liquid chromatography method for simultaneous determination of five statins in pharmaceutical form. Acta Chromatogr. 2021;33:345-353.
- 77. Novotná K, Havliš J, Havel J. Optimisation of high performance liquid chromatography separation of neuroprotective peptides: fractional experimental designs combined with artificial neural networks. J Chromatogr A. 2005;1096(1-2):50-57.
- 78. Maljurić N, et al. Quantitative structure–retention relationship modeling of selected antipsychotics and their impurities in green liquid chromatography using cyclodextrin mobile phases. Anal Bioanal Chem. 2018;410(10):2533-2550.

Primena eksperimentalnog dizajna za razdvajanje lekova HPLC metodom

Jevrem Stojanović, Jovana Krmar, Ana Protić¹, Bojana Svrkota, Nevena Đajić, Biljana Otašević*

Univerzitet u Beogradu – Farmaceutski fakultet, Katedra za analitiku lekova, Vojvode Stepe 450, 11221 Belgrade, Serbia

*Autor za korespondenciju: Biljana Otašević, e-mail: biljana.otasevic@pharmacy.bg.ac.rs

Kratak sadržaj

Dizajn eksperimenata (DoE) je nezaobilazan alat u savremenoj analizi lekova budući da istovremeno balansira niz hromatografskih parametara kako bi se osiguralo optimalno razdvajanje u tečnoj hromatografiji pod visokim pritiskom (HPLC). Prikazana je teorijska osnova DoE i data su detaljna uputstva za njegovu primenu u HPLC ispitivanjima u farmaciji. Naročito se govori o klasifikaciji brojnih tipova dizajna i njihovim mogućnostima za racionalizaciju različitih faza tokom procesa razvoja HPLC metode, kao što su izbor najuticajnijih faktora, optimizacija faktora i procena robusnosti metode. Dodatno, sumirana je primena DoE kao sastavnog dela koncepta ugradnje kvaliteta u proizvod u domenu razvoja analitičkih metoda (AQbD) zasnovanih na HPLC tehnici. Takođe su prikazana i nedavna dostignuća u primeni DoE u razvoju LC metoda koje su pogodne za ispitivanje stabilnosti lekova, kao i LC-MS metoda. U budućoj perspektivi, preporučeno je izvođenje ispitivanja kvantitativnog odnosa između strukture i retencionog ponašanja (QSRR) analita u HPLC sistemu na osnovu podataka dobijenih primenom DoE.

Ključne reči: eksperimentalni dizajn, analitika lekova, tečna hromatografija pod visokim pritiskom, razvoj i validacija analitičke metode