

# A Design of Experiment Approach to Optimize an Image Analysis Protocol for Drug Screening

Antonella Lanati, Cecilia Poli, Massimo Imberti, Andrea Menegon,  
and Fabio Grohovaz

**Abstract** The Design of Experiment, DoE, was applied to support the development of an innovative optical platform for ion channel drug screening. In this work, DoE was exploited to investigate a set of software parameters instead of process variables, an approach that has been only rarely explored. In particular, it was used to define a standard analytical configuration for a MatLab-based image analysis software that has been developed in the laboratory to extract information from images acquired under the drug screening conditions. Since the choice of the type of analysis and filtering, as well as their interactions, was known to affect the final result, the aim was to identify a robust set of conditions in order to obtain reliable concentration-response (sigmoidal) curves in an automated way. We considered five parameters as factors (all qualitative) and two characteristics of the sigmoidal curve as reference outputs. A first DoE screening was performed to reduce the number of needed levels for one factor (an unconventional approach) and a second optimization study to define the best configuration setting. Image stacks from three different experimentation days were used for the analysis and modelled by blocks to investigate inter-day variations. The optimized set of parameters identified in this way was successfully validated on different cell lines exposed to their references drugs. Thanks to this study, we were able to: find the optimized configuration for the analysis, with a reduced number of trials compared to the classical “One Variable at A Time” approach; acquire information about the interactions of different analytical conditions as well as the inter-day influence; and, finally, obtain statistical evidence to make results more robust.

---

A. Lanati (✉) • C. Poli  
Valore Qualità, 27100 Pavia, Italy  
e-mail: [alanati@valorequalita.eu](mailto:alanati@valorequalita.eu)

M. Imberti  
Open Sistemi, 26100 Cremona, Italy

A. Menegon  
San Raffaele Scientific Institute, 20132 Milan, Italy

F. Grohovaz (✉)  
San Raffaele Scientific Institute, 20132 Milan, Italy

Vita-Salute San Raffaele University, 20132 Milan, Italy  
e-mail: [grohovaz.fabio@hsr.it](mailto:grohovaz.fabio@hsr.it)

**Keywords** Design of Experiment • Ion Channels • Drug Screening • Image analysis

## 1 Introduction

### 1.1 *Use of Design of Experiment in Research*

The design of experiment (DOE), also known as experimental design, was developed by Ronald Fisher in the early 1920s. It is a statistical method aimed to plan and analyze experiments, in order to extract the maximum amount of information with the fewest possible number of runs. It allows also to build regression models and to optimize the output by choosing proper variable settings. The traditional way of conducting experiments is intended to change One Variable at A Time (OVAT), thus accepting the risks to become trapped in a local optimum, missing the global optimum. The DoE allows changing simultaneously all the variables, helping in finding their best combination [6, 12, 14]. Moreover, it provides a regression model that, in the range of variables used to build it, can make predictions for values different from those used in the study (see: [1, 18]).

During the last decades, DOE has been successfully applied to optimize processes in chemistry and engineering [13] as well as in pharmaceutical and biopharmaceutical industry, both in development and production [2, 7, 9–11, 14, 22, 25, 26]. New applications are emerging in biomedical research, specifically in medium-high throughput assays and in the optimization of laboratory protocols [3, 4]. In particular, DoE has recently been used in drug screening, where progress in molecular biology and advanced technologies has given new opportunities to test large chemical libraries against biological targets. However, the introduction of combinatorial chemistry and high-throughput screening has not met the expectations, rather it has been accompanied by a decline in productivity [20]. This can be ascribed to a number of reasons, including the fact that the process of selection leaves behind many potentially interesting molecules [16, 23]. This has drawn the attention to cell-based assays and to more robust screening approaches in order to increase R&D efficiency/efficacy, and thus productivity. In this respect, attention has also grown toward methods for an efficient development and setup of the assays. In the last decades, there are several examples of the application of DoE in drug screening [5, 11, 15, 17, 24], mainly related to the optimization of biological and biochemical process condition. Other examples are oriented to optimization of data processing in metabolomics [8, 27]. The use of experimental design for optimizing software parameters is still poorly explored (e.g. [21]).

In this paper, we report the use of DoE for the fine set up of the analytical processing in a newly developed drug screening approach.

## 1.2 *Dedicated Image Analysis Software for a New Drug Screening Approach*

An innovative optical platform for ion channel drug screening, based on a proprietary approach, has been developed by a multidisciplinary team. Briefly, cells expressing the channel of interest are loaded with a fluorescent voltage-sensitive dye and the effect of a drug is revealed by the fluorescence values recorded before and during exposure to electrical stimulation (see EP2457088 patent for more details). Images acquired under the above conditions are processed by a MatLab-based Image Analysis (MaLIA; a program developed in the laboratory) that offers the possibility to employ different filters, parameters and types of analysis. Data representative of the cellular response to the electric pulses are used to extrapolate changes in resistance/conductance; these values are put in relation with increasing concentrations of the molecule of interest, thereby obtaining a typical sigmoidal Concentration-Response (CR) curve defined by a set of qualitative and quantitative parameters. In the course of the project, the MaLIA has progressively evolved, gaining in flexibility, to explore multiple analytical options. In this development phase, different analysis configurations were experimented: the parameter space was narrowed to a set of five, four of which varying between two values only. The final goal of the project was to define the optimized values of these parameters, in order to perform a standard analysis in full automation, without external, arbitrary interventions. To this end, we employed the DoE to evaluate the effects of different parameters/filters implemented in the MaLIA as well as their interactions.

## 2 The Design of Experiment (DoE) Method

The different values assumed by each *factor* (the experimental variables) are called *levels* (typically only two, codified as  $-1$  or  $+1$ ) and can be either qualitative or quantitative. The DoE allows evaluating both the influence of single variables (*main factors*) and the interplay among factors (*interactions*), i.e. when the effect of a factor depends on the level of one or more other factors. A specific combination of levels is called *treatment* (or *run*). Each treatment is evaluated in terms of *outputs* or *responses*, which are representative of the behaviour of the system. The magnitude of a change in response, when factors are varied, is called *effect*.

In order to achieve statistically relevant conclusions from experiments, it is necessary to adopt different statistical principles: *randomization* (i.e. scrambling the running order of treatments), *replication* (i.e. repeating each treatment twice or more) and *blocking* (i.e. modelling extraneous sources of variation as special variables). These principles minimize experimental bias that may mask the responses of the significant factors (see [18]).

The term *factorial design* identifies the most used set of treatments employed to investigate the effects of factors on a response. The simplest factorial design is called *two-level factorial design*, and is used in two forms: *full factorial* and *fractional factorial design*. Full factorial design requires  $2^K$  runs, where  $K$  is the number of factors, and generates, with the increasing number of factors, a considerable (even unmanageable) amount of runs. In case of many factors (e.g.  $>5$ ), we can reduce the number of requested runs, based on specific assumptions, such as ignoring interaction of more than three factors: in this way we perform a fractional factorial design. Experimental designs reduced to  $2^{K-1}$  runs are called Resolution V designs. In these designs, no main effect or two-factor interaction is aliased with any other main effect or two-factor interaction, but two-factor interactions are aliased with three-factor interactions. There is also the possibility to consider more than two levels for each factor, and to create a *general factorial design* [1].

After performing experiments according to the planned design, results are analyzed through a graphical interpretation (*Factorial Plots* and *Statistical Plots*) and a set of statistical parameters. The Factorial Plots include the *Main Effect Plot* and the *Interactions Plot*. The former is represented as a straight line: the slope indicates the direction while its magnitude the strength of the effect. On the other hand, the *Interactions Plot* shows how different combinations of factor settings affect the response: non-parallel lines show interaction between couple of factors. The *Normal Probability Plot* (one of the *Statistical Plots*) is a different representation of a distribution, with the cumulative percentage on the logarithmic Y-axis and the ordered values of the observations on the X-axis. In this representation, the Gaussian distribution appears as a straight line. It is used to check normality of the data and to find out the most significant ones: non-significant data are dispersed along a straight line, whereas significant data are apart. In the experimental design, the Normal Probability Plot is used to evaluate significance and normality of both main and interaction effects. The *Pareto Plot* (another *Statistical Plot*) displays the absolute values of main and interaction effects: a reference line shows statistically significant values ( $P < 0.05$ ). The *Normal Plot for Residuals* is conceptually the same as the one used for effects and interactions and estimates the difference (residuals) between actual and predicted values (calculated by the regression model obtained from the DoE analysis), to verify whether the data have a Gaussian distribution.

This analysis can be complemented by a number of statistical parameters, including a regression model describing each response as a function of the selected factors and information coming from the ANOVA analysis (see: [18]).

As a general approach, a *screening analysis* is first performed with less stringent conditions to identify the most significant factors. Subsequently, an *optimization analysis* is applied to a narrower set of factors to find the best condition that optimizes the output(s).

Along with the factorial designs, DoE offers a rich set of other designs, to suit most requirements. Few examples are:

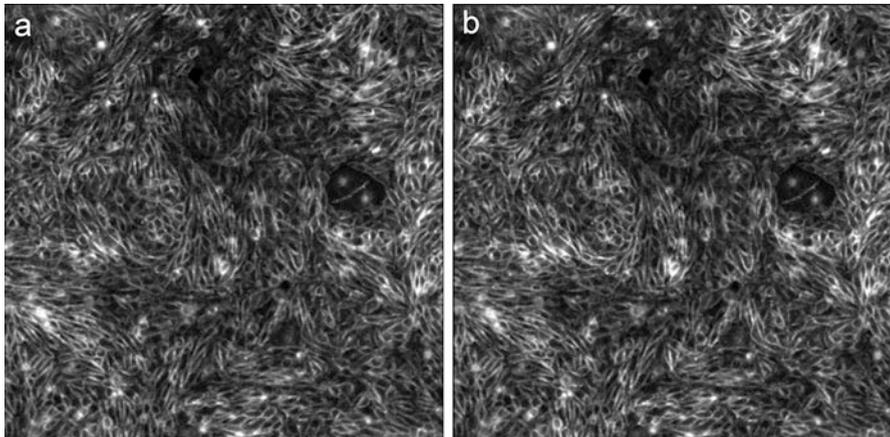
1. *Plackett–Burman* design, which evaluates the effects of main factors only, with a small set of runs. It is mainly used in the screening phase;

2. *Response surface* designs (e.g. Central Composite, Box–Behnken), which are used to identify points of absolute maximum, and to highlight possible nonlinearities (for quantitative factors only). They are mainly used in the optimization phase.
3. *Mixture design*, which is used when factors are components of a blend, and the output depends on their relative proportion.

### 3 Experimental Setup

Experiments were performed on a Chinese hamster ovary (CHO) cell line expressing the human transient receptor potential (TRPV1) channel (kindly provided by Axxam S.p.A) using capsaicin as reference agonist. CHO-TRPV1 cells were stained with a voltage sensitive dye (VSD; di-4-ANEPPS), and exposed to a square electric pulse. Local fluorescence values were measured before and during the pulse (Fig. 1, left and right, respectively) both in the absence and in the presence of capsaicin. As expected from the poor sensitivity of the VSD (8 % fluorescence variation/100 mV), changes are hardly appreciated at first sight and a sophisticated analysis is necessary to automatically isolate and evaluate subcellular responsive areas. Further details are available on the patent EP2457088 and will be reported in a full paper on this new approach (Menegon et al. in preparation).

Among the different types of DoE designs, we decided to use factorial designs for two reasons. On the one hand, we needed to evaluate second order interactions



**Fig. 1** CHO-TRPV1 cells images before (a) and during (b) exposure to an electrical square pulse. The signal (differences in fluorescence intensity in specific subcellular regions) is not easily appreciable without proper data processing

and Plackett–Burman was not suitable. On the other hand, our factors were typically at two levels, making inappropriate other analyses such as response surface designs.

The very same stack of images was processed many times with MaLIA, to cover all the combinations of parameters indicated by the experimental design. Randomization was not required, because no external bias factors could affect the running of the software analysis. For DoE analysis, we selected, among the parameters implemented in MaLIA, the following five factors (variables) that appeared to influence the output data:

- a. Binning: to reduce image noise by combining cluster of pixels into single pixels;
- b. Shape-mask (ShapeM): to select the membrane responsive areas;
- c. Minimum responses filtering (MinRespFilt): to discard signal values lying inside the noise range;
- d. Response calculation (RespCalc): Fold Change (FC) or Normalized Fold Change (NFC);
- e. Output data filtering (OutputDataFilt): pure statistical or functional (to exclude variations not coherent with the expected biological response)

We defined also two outputs to evaluate the influence of these parameters on CR curves:

1. R-squared (Rsq), as a measure of good fitting of the sigmoidal curve;
2. Top minus bottom (T-B), as the difference between highest and lowest values in the sigmoidal curve (a measure of the efficacy of tested drug).

Finally, in order to account for possible inter-day variations (due to biological variability and/or changes in the process), we repeated the same set of treatments on image stacks obtained in three different experimental days, and modelled each of these replications by blocks.

The MaLIA parameters are qualitative and at two levels only, with the sole exception of Binning that has three possible levels: for a full evaluation, a general factorial design with five factors should be employed. According to Anderson and Whitcomb (see Chap. 7, pp. 133–134): (1) a general factorial design is to be avoided when the number of factors increases (typically higher than 3), (2) a reduction of a general factorial design requires ad hoc elaboration. The same authors suggest making preliminary tests to attempt to reduce the analysis to a two-level factorial. In our case, a complete general factorial design (5 factors, one of them at 3 levels) would require  $2^{(5-1)} \times 3 = 48$  runs per replicate that, multiplied by the 3 foreseen replicates, give a total of 144 runs. As expected, the DoE software we use does not allow for reducing general factorial designs. In line with the suggestions of Anderson and Whitcomb, we evaluated the possibility to reduce the number of levels for Binning. Therefore, we first set an unconventional screening analysis, by considering the sole two factors directly involved in the extraction of data from images: Binning (three levels) and Shape Mask (two levels), by using a standard set of the other parameters. Thanks to the reduction of Binning to two levels, in the second analysis we were able to evaluate all factors at two levels with a fractional factorial design. In this way, it was possible to perform the analysis with only 6 + 16

runs for each replicate. Finally, we made a validation to verify: (1) that there is no significant interaction between Binning and the other factors; and (2) that the discarded Binning level was less suitable for optimal results.

We generated Normal Plot and Pareto Plot, to identify statistically significant factors and interactions, as well as the Main Effect Plot and the Interaction Plot to evaluate factors influence on each output. Goodness of fit was judged by the Residuals Plots and other statistical parameters. Blocks provided information about the influence of different experimental days (inter-day variations). All the DoE analysis was performed by Minitab, a statistics package developed at the Pennsylvania State University (Minitab Inc., State College, PA, USA).

## 4 Results of the First Analysis (Screening)

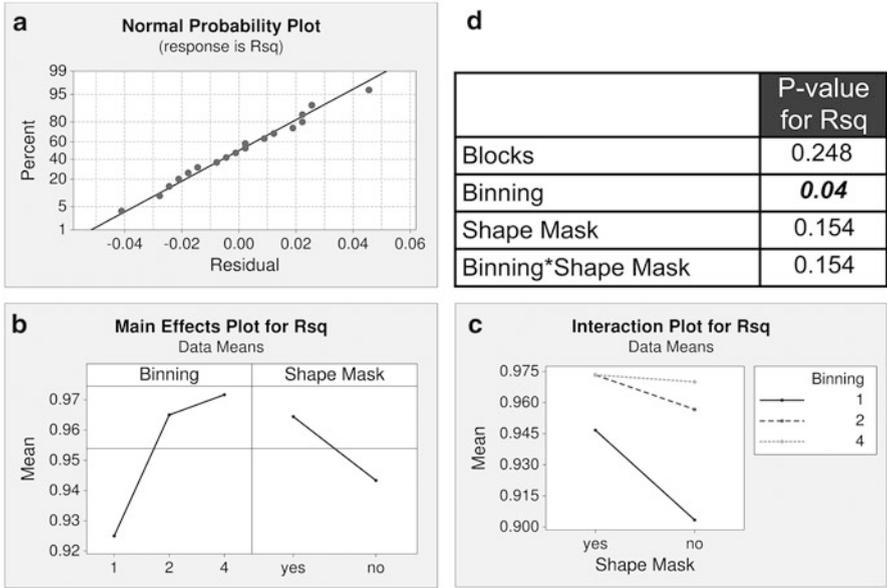
The first analysis was aimed to find the two most significant values out of the three possible levels of image Binning and was performed considering the Shape-Mask as the sole factor able to interact significantly with the Binning. In fact, only these two variables are directly related to the pixels of the image. Factors and their levels were as follows:

1. Binning ( $1 \times 1$ , i.e. no binning;  $2 \times 2$ ;  $4 \times 4$ ; referred to as 1, 2 and 4, respectively);
2. Shape-mask (yes; no).

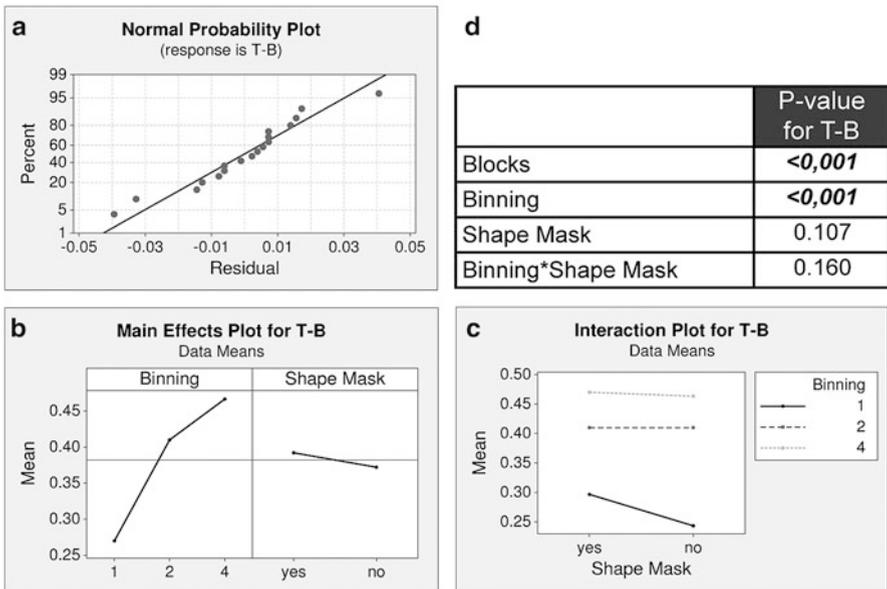
Because of the three-levels Binning factor, a General full factorial design was used (18 runs, 3 replicates). The screening analysis clearly demonstrates an interaction between Binning and Shape Mask on the Rsq output (Fig. 2) but not on the T-B output (Fig. 3).

Figure 2a indicates that residuals for Rsq are normally distributed—i.e. very close to the line representing the normal distribution—a condition necessary to proceed with a standard analysis without doing a variable transformation (see: [1]). The analysis shows a significantly lower Rsq for Binning 1 compared with Binning 2 and even more with Binning 4 ( $P = 0.04$ , Fig. 2d). An improvement in Rsq is observed when Shape Mask is applied (Fig. 2b). The interaction Plot (Fig. 2c) confirms that Binning 1 gives lower Rsq while Binning 2 and 4 show the best results. The influence of Shape Mask is maximal with Binning 1, moderate with 2 and negligible with 4. A  $P$ -value = 0.248 for the variable Blocks shows no influence of inter-day conditions for Rsq.

Figure 3a indicates that residuals are normally distributed also for T-B. The effect of Binning on the T-B output confirms Binning 1 as the worst condition, but also shows a trend, with Binning 4 better than 2 (see Fig. 3b); interestingly, Shape Mask has no influence on the T-B considered alone or even in combination with Binning as shown by the interaction plot (Fig. 3c), where lines are almost parallel. A  $P$ -value  $< 0.001$  for the variable Blocks indicates a significant influence of inter-day conditions on T-B. Overall, Binning was the sole significant factor ( $P < 0.01$ , Fig. 3d).



**Fig. 2** Minitab graphs of the screening analysis for Rsq: Normal probability plot for Residuals (a), Main Effect Plot (b) and Interaction Plot (c). The table in (d) shows P for the chosen factors and their interactions



**Fig. 3** Minitab graphs of the screening analysis for T-B: Normal probability plot for Residuals (a), Main Effect Plot (b) and Interaction Plot (c). The table in (d) shows P values for the chosen factors and their interactions

## 5 Results of the Second Analysis (Optimization)

The aim of the second analysis was to define an optimal parameter configuration, by considering the following factors/levels:

1. Binning (2; 4);
2. Shape-mask (yes; no);
3. Minimum response filtering (yes; no);
4. Output data filtering (Stat; Funct);
5. Response calculation (NFC; FC).

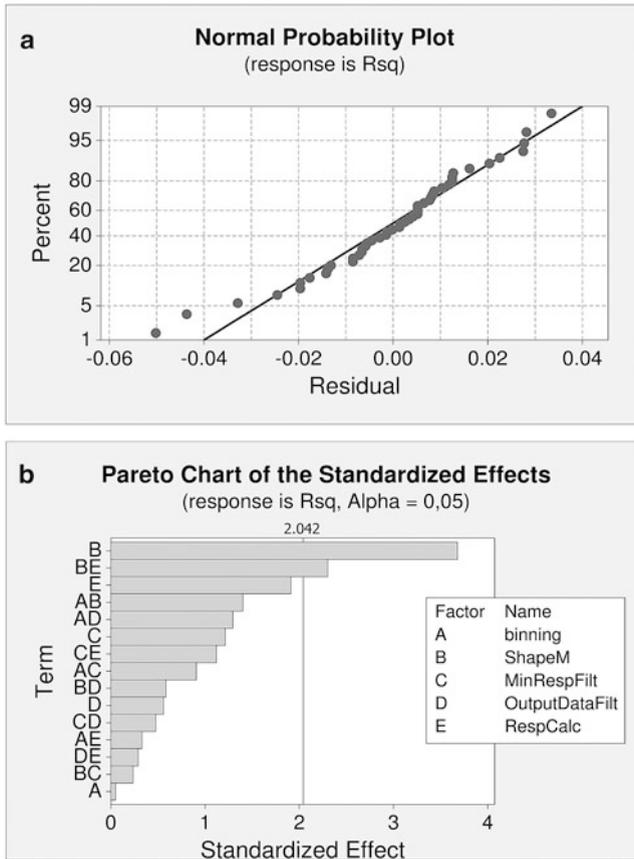
Under these conditions, a Full Factorial design would have required 32 runs per replicate, i.e. the same number of runs needed for an OVAT approach, however, with the advantage of providing information about interactions. Considering that we were interested also on the influence of inter-day variability, a minimum of 3 replicates (performed on image stacks produced in different days) had to be performed. This would have required a total of 96 runs. In order to reduce this number, we made the assumption that the interactions of the second order (i. e. interactions of two factors at a time) were sufficient for a correct approximation in our analysis, also considering that higher interactions (three factors at a time or more) are expected to be negligible in most cases (see [19]). Based on these considerations, we reduced the number of trials by employing the fractional factorial design with resolution V, which required 48 runs for 3 replicates, at the expenses of the assessment of third order interactions. Figure 4 shows the results of DoE Analysis for the Rsq output. Normal probability Plot (Fig. 4a) for Residuals show good fitting. The Pareto Chart of the Standardized Effects (Fig. 4b) indicates that the only statistically significant factor is Shape Mask ( $P = 0.001$ ) while the only significant interaction is Shape Mask with Response Calculation ( $P = 0.029$ ).

Taking into consideration the results shown in Fig. 5a, b, we can assume that, as far as Rsq is concerned, best results are obtained with: Shape-mask, Binning 4, no Minimum response filtering, Statistical Output data filtering and NFC Response calculation.

Similar analysis was performed considering T-B as the Output. Normal probability Plot (Fig. 6a) for residuals show good fitting. The Pareto chart of the standardized effects indicates that all the main factors, but Minimum response filtering, are statistically significant (Fig. 6b): Output data filtering ( $P < 0.001$ ); Binning ( $P < 0.001$ ); Shape-mask ( $P = 0.002$ ); and Response calculation ( $P = 0.004$ ). Minimum response filtering has a significant interaction with Output Data Filtering ( $P = 0.021$ ).

Considering the Main Effects Plot (Fig. 7a) and the Interaction Plot (Fig. 7b) for T-B, we can infer that best results are obtained with Shape-mask, Binning 4, Statistical Output data filtering, no Minimum response filtering and NFC Response calculation.

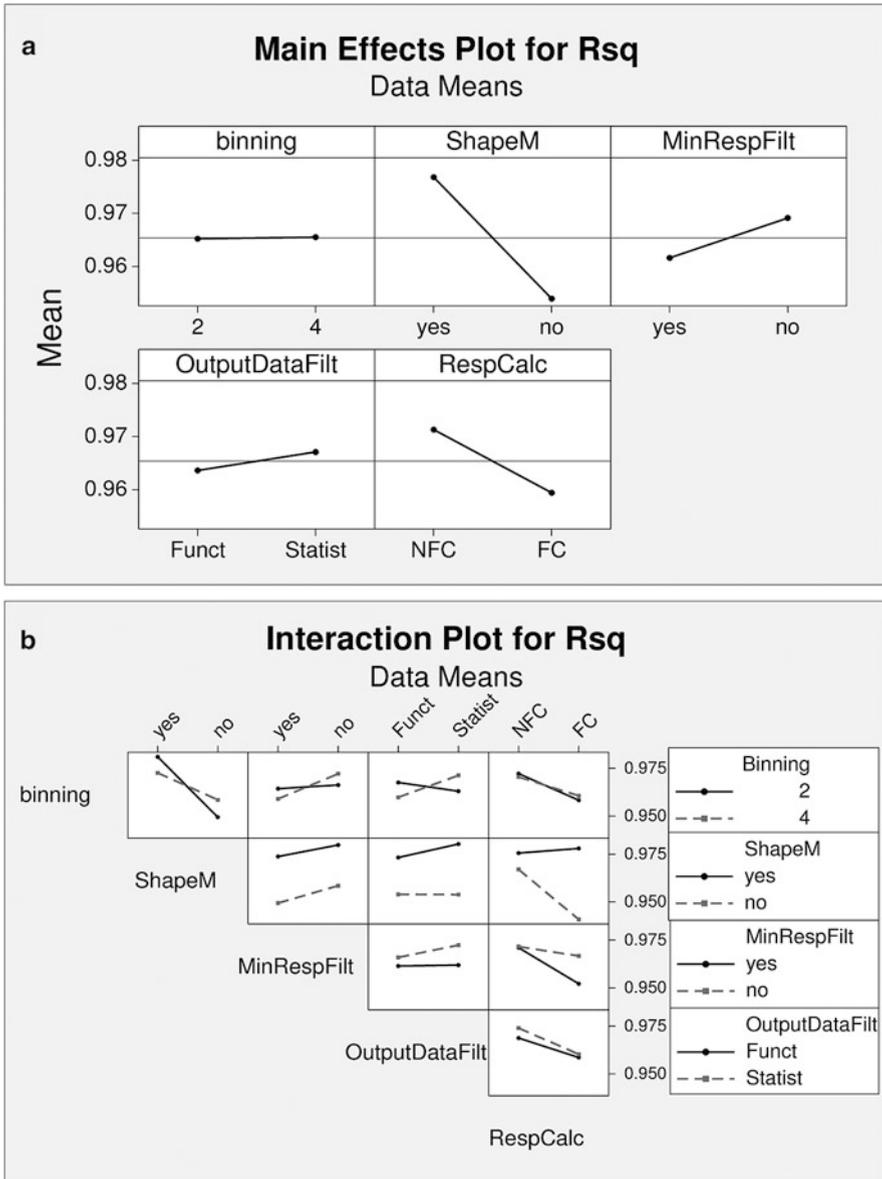
Based on the above results, we were able to define an optimized configuration (Table 1) and a suboptimal one (Table 2).



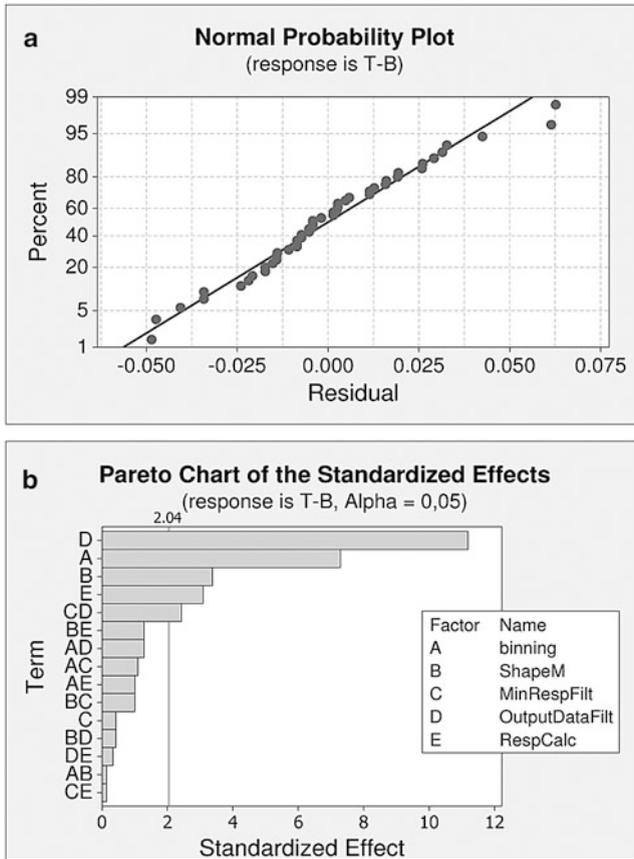
**Fig. 4** Minitab graphs for Rsq optimization analysis: the Normal Probability Plot for Residuals (a) indicates an suitable distribution of residuals; the Pareto of the Standardized Effects (b) indicates that there are only two significant effects (i. e. laying beyond the vertical line that marks the threshold for Alpha = 0.05): Shape Mask and the interaction between Shape Mask and Response Calculation

C-R curves were then calculated with both the optimized and the suboptimal set on the same data used for DoE analysis. Figure 8 illustrates an example in which the C-R curve obtained with the optimized set exhibits an Rsq value improved from 0.91 to 0.99 and a T-B value from 0.28 to 0.49, which represent a percent improvement (defined as  $(PS_{opt} - PS_{subopt})/PS_{subopt}$ , where PS = parameter set) of respectively +8.8 % (Rsq) and +75 % (T-B).

As a final consideration, P-value for BLOCKS showed an influence of inter-day conditions that is significant for T-B ( $P < 0.001$ ) but not for Rsq ( $P = 0.147$ ).



**Fig. 5** Minitab graphs for Rsq optimization analysis (factorial plots): the Main Effects Plot (a) confirms that the Shape Mask effect is the most important among single factors and that best results are obtained when the mask is applied: the Interaction Plot (b), shows the best combination for Shape Mask and Response Calculation (if ShapeM = yes, both values for RespCalc are suitable)

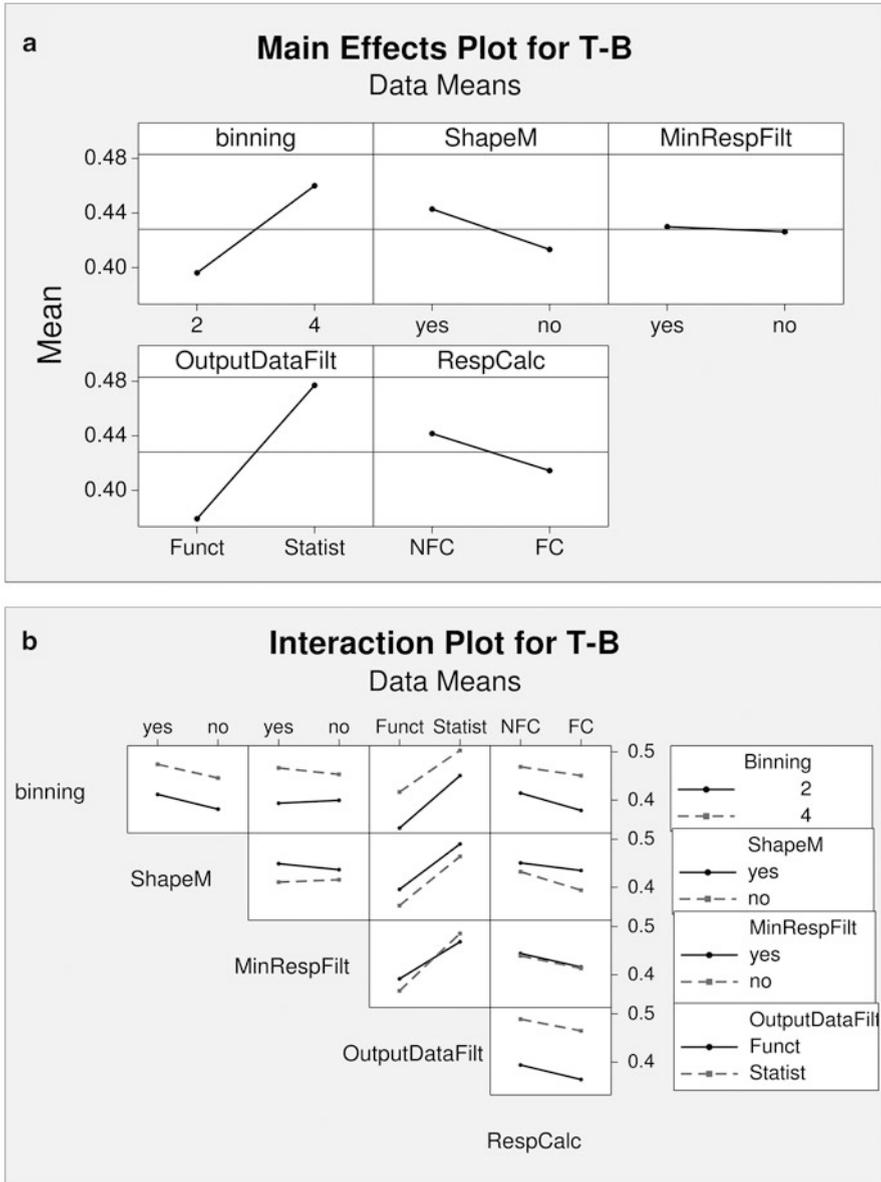


**Fig. 6** Minitab graphs for T-B optimization analysis: the Normal Probability Plot for Residuals (a) indicates a suitable distribution of residuals; the Pareto of the Standardized Effects (b) shows that all single factors, but the Minimum Response Filter (MinRespFilt), are significant, while only one interaction, the one between MinRespFilt and OutputDataFilt, lays beyond the vertical line (threshold for Alpha = 0.05)

## 6 Validation of Obtained Optimized Configuration

The optimized parameter configuration we obtained with the previous analysis was then validated.

As a first step, we verified the initial hypothesis that Binning had no significant interactions with factors other than ShapeM. Indeed, Figs. 4a and 6a show that the interactions between Binning and the other factors do not reach statistical significance. Of note, in the same Fig. 4a we can appreciate that ShapeM has a significant interaction with RespCalc, clearly indicating that it is not possible to separate the pixel-related factors from the others.



**Fig. 7** Minitab graphs for T-B optimization analysis (factorial plots): the Main Effects Plot (a) indicates the best values for the significant factors: Binning = 4, ShapeM = yes, OutputDataFilt = Statist and RespCalc = NFC. In the Interaction Plot (b), the value MinRespFilt = no together with OutputDataFilt = Statist are the significant interacting factors values that optimize the output T-B

**Table 1** Optimized parameter set for Rsq and T-B

Optimized configuration			
Factor	Optimized level for Rsq	Optimized level for T-B	Optimized choice
Binning	n.i.	4	4
ShapeMask	YES	YES	YES
Minimum response filtering	n.i.	NO	NO
Output data filtering	n.i.	Stat	Stat
Response calculation	NFC	NFC	NFC

*n.i.* not influent

**Table 2** Suboptimal parameter set for Rsq and T-B

Suboptimal configuration			
Factor	Suboptimal level for Rsq	Suboptimal level for T-B	Suboptimal choice
Binning	n.i	2	2
ShapeMask	NO	NO	NO
Minimum response filtering	n.i.	YES	YES
Output data filtering	n.i.	Funct	Funct
Response calculation	FC	FC	FC

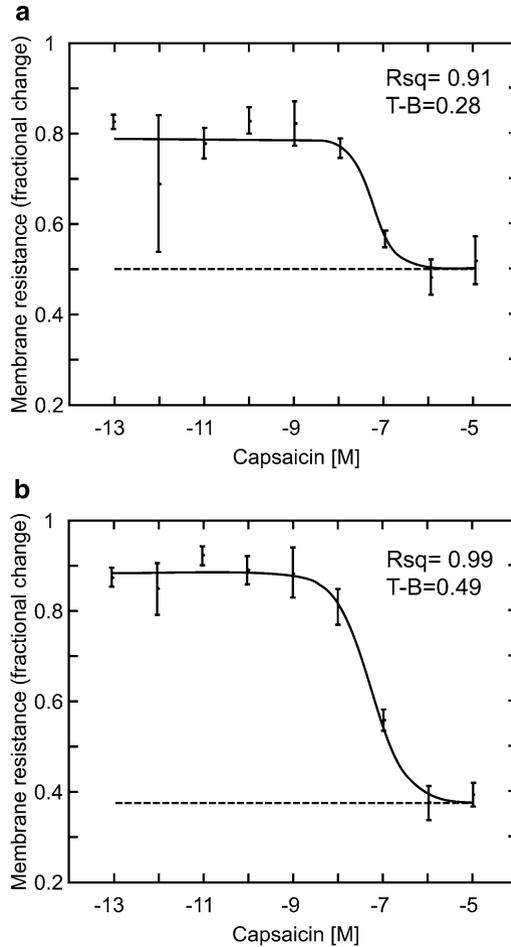
*n.i.* not influent

Afterwards, to validate the rejection of Binning = 1 in the first analysis, we ran MaLIA on 8 different image stacks with the same parameters employed in the optimized (Table 1) and suboptimal (Table 2) configurations, with the exception of Binning value set to 1. The substitution of Binning = 1 worsened the value of Rsq and T-B in both the optimized (−6 % and −30 %, respectively) and the suboptimal configuration (−17 % and −54 %, respectively). We can conclude that, as suggested by the experience during the development of the MaLIA program and assumed during the design of the first analysis, Binning = 1 minimized the overall performance. This ex post validation also confirms the validity of the assumptions we made in the first analysis of this unconventional DoE design.

Afterward, we produced C-R curves with both the optimized and the suboptimal sets on data from different experiments in order to validate the results in a wide range of cell and drug types (see Table 3).

Experimental data were randomly selected within a time interval of 2 years, representing five cellular lines exposed to their reference drugs. Two experiments for each cell line were considered. Such a wide time interval was used to take into account also changes due to the evolution of both biological protocols and screening processes.

The Rsq and the T-B values of the C-R curves obtained with the optimized and suboptimal parameters sets are compared in Fig. 9a, b and shown as percent variation  $((PS_{opt} - PS_{subopt})/PS_{subopt})$  in Fig. 9c. The charts clearly indicate that the optimized set consistently produces better CR curve: Rsq benefits of a slight



**Fig. 8** CR curves obtained with optimized and suboptimal parameter sets: the CR curve shows the fractional changes of the membrane resistance at different drug concentrations (log). The CR curves obtained with the suboptimal (a) and with the optimized (b) parameter sets, on the same images stack, are compared to put in evidence the marked improvement: R<sub>sq</sub> from 0.91 to 0.99, T-B from 0.28 to 0.49

improvement (up to 4.3 %), while T-B takes much more advantage (up to 90.4 %). The only exception is represented by an experiment (HEK-293 GABA-A exp. 1), in which R<sub>sq</sub> is lower (−1.2 %) with the optimized set, even though the T-B response maintains a positive performance of +9.3 %.

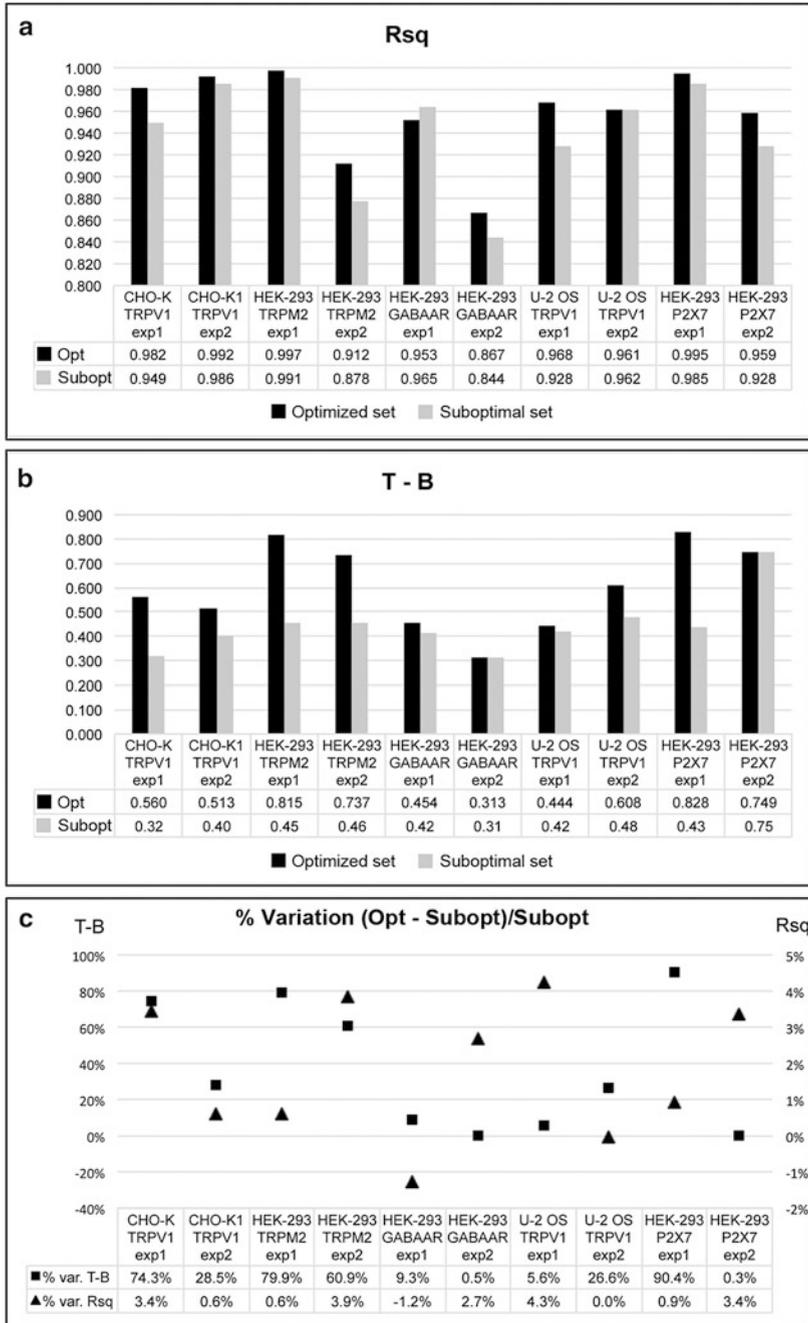


Fig. 9 (continued)

**Table 3** Pharmacological targets used to validate the optimized parameter set

Pharmacological target		Reference drug
CHO-K1/TRPV1	Chinese hamster ovary K1 cells expressing the transient receptor potential cation channel subfamily V member 1	Capsaicin
HEK-293/TRPM2	Human Embryonic Kidney 293 cells expressing the transient receptor potential cation channel, subfamily M, member 2	(agonist)
HEK-293/GABAAR	Human Embryonic Kidney 293 cells expressing the $\gamma$ -aminobutyric acid type A receptor	$\gamma$ -aminobutyric acid
U-2 OS/TRPV1	human osteosarcoma U2OS cell line expressing the transient receptor potential cation channel subfamily V member 1	Capsaicin
HEK-293/P2X7R	Human Embryonic Kidney 293 cells expressing the purinergic receptor subclass P2X7	BzATP, (2'/(3')-O-(4-Benzoylbenzoyl)adenosine-5'-triphosphate tri(triethylammonium)

## 7 Discussion and Conclusions

DoE was performed to optimize the set of analytical parameters used in a new drug screening procedure. This is a simple application of the method that provided useful results with good efficiency (time and resources vs results).

We adopted an unconventional DoE approach: the screening design, instead of being employed to reduce the number of factors, was used to reduce the number of levels of one of the factors, taking advantage of the specific knowledge of the image analysis process. This simplification made possible, in a second analysis, a direct and simpler comparison among all main parameters, thereby avoiding the more complex General Factorial design. This second design, which we called optimization in this study, was performed with a Resolution V Fractional Factorial design. A standard Response Surface design could not be employed since all factors are qualitative and intermediated values could not be envisaged (see: [18]). The choice of using a Fractional Factorial design, which ignores third order interactions, is largely supported by the evidence that results are little influenced by pairs of factors (interaction of the second order), validating the initial assumption of a negligible contribution of higher order interactions.

 **Fig. 9** Validation of the optimized set on different cells/drugs (see Table 2): Rsq (a) and T-B (b) values are always improved with the optimized set of parameters (*black columns*) rather than with the suboptimal one (*gray columns*), with a single exception (HEK-293 GABAAR exp.1,  $Rsq_{opt} < Rsq_{subopt}$ ). The percent variations are shown in (c), where Rsq variations refer to the *right y-axis*, while the T-B variations to the *left y-axis*

The unconventional choice of a general full factorial, followed by a fractional factorial design, allowed us to downsize the number of runs. A traditional general factorial, with 4 factors at 2 levels and 1 at 3 levels, would have required 48 ( $2^4 \times 3$ ) runs for each replicate. With our approach, we made 6 ( $3 \times 2$ ) runs for the first phase and 16 ( $2^{(5-1)}$ ) for the second analysis, with a total of 22 and a saving of 26 runs with respect to a single general factorial design. Considering 3 replicates, we saved 78 runs. Each MaLIA run (inputting data, setting parameters, waiting for analysis elaboration and collecting results) takes at list 7–10 min to a skilled operator. Accordingly, we saved up to 13 h on a total forecasted effort of 24 h, i.e. 54 % saving. Overall, by this DoE approach, we saved time, gaining more information.

Finally, the use of the Blocking reveals that the impact of the experimental day could not be neglected in this study, which embraces 2 years' work of development of the drug screening platform. Interestingly, only the T-B, but not the Rsq, was subjected to inter-day variation. This result can lead to two conclusions: first, this influence may deserve further analysis after final validation of the screening platform. Secondly, experimental data could have been transformed to correct for blocks, thus obtaining a result independent of day-to-day variability (see [1]; Chap. 2). However, the excellent validation of the optimized set of parameter demonstrates that the result is robust enough to make a more sophisticated analysis not necessary. Further investigation might involve a refinement of the quantitative thresholds used for some of the parameters (e.g. threshold for Minimum response filtering). If we consider the results in terms of the final application, the observed inter-day variation appears to reflect the process of optimization of the biological and biochemical conditions during the progressive development of the drug screening platform. On the other hand, they provide direct evidence that also in sub-optimal experimental conditions, the set of choice guarantees the best possible result. This evidence receives further confirmation by the fact that a consistent improvement was observed independently of the cell lines and drug type. Overall, this is an important prerequisite to consider this new approach for the study of different pharmacological targets, in an unbiased way and in an industrial context.

In conclusion, our work demonstrates that the application of DoE on the selection of software parameters, although still poorly exploited, can provide very useful results by reducing the number of trials compared to a complete OVAT approach. In this respect, it is worth noticing how a conscious introduction of constraints to reduce the degrees of interactions, along with a two-stage design, can greatly simplify the modelling and thus the obtaining of the result.

**Acknowledgments** The authors wish to thank Axxam S.p.A. for kindly providing the cell lines employed in this work. We greatly acknowledge the scientific contribution of Dr. Riccardo Fesce during the development of MaLIA. The project was developed in the R&D laboratory of the Advanced Light and Electron Microscopy BioImaging Center (Experimental Imaging Center-San Raffaele Scientific Institute) within the framework of the project “Optical Method for Ion Channel Screening”, “Progetto Metadistretti Tecnologici”, Regione Lombardia.

## References

1. Anderson, M.J., Whitcomb, P.J.: DOE Simplified. Practical Tools for Effective Experimentation. Productivity Press, New York (2007)
2. Beyer, H.G., Sendhoff, B.: Robust optimization—a comprehensive survey. *Comput. Methods Appl. Mech. Eng.* **196**, 3190–3218 (2007)
3. Bollin, F., Dechavanne, V., Chevalet, L.: Design of experiment in CHO and HEK transient transfection condition optimization. *Protein Expr. Purif.* **78**(1), 61–68 (2011)
4. Bongiovanni, A., Colotti, G., Liguori, G.L., Cirafici, A.M., Di Carlo, M., Digilio, F.A., Lacerra, G., Mascia, A., Lanati, A., Kisslinger, A.: Applying quality principles and project management methodologies in biomedical research: a public research network's case study. *Accred. Qual. Assur.* **20**(3), 203–213 (2015)
5. Chou, T.C.: Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol. Rev.* **58**, 621–681 (2006)
6. Dejaegher, B., Vander Heyden, Y.: Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. *J. Pharm. Biomed. Anal.* **56**, 141–158 (2011)
7. Ebrahimi-Najafabadi, H., Leardi, R., Jalali-Heravi, M.: Experimental design in analytical chemistry—part I: theory. *J. AOAC Int.* **97**(1), 3–11 (2014)
8. Eliasson, M., Rännar, S., Madsen, R., Donten, M.A., Marsden-Edwards, E., Moritz, T., Shockcor, J.P., Johansson, E., Trygg, J.: Strategy for optimizing LC-MS data processing in metabolomics: a design of experiments approach. *Anal. Chem.* **84**, 6869–6876 (2012)
9. Elliott, P., Billingham, S., Bi, J., Zhang, H.: Quality by design for biopharmaceuticals: a historical review and guide for implementation. *Pharm. Bioprocess.* **1**(1), 105–122 (2013)
10. Glasnov, T.N., Tye, H., Kappe, O.: Integration of high speed microwave chemistry and a statistical design of experiment' approach for the synthesis of the mitotic kinesin Eg5 inhibitor monastrol. *Tetrahedron* **64**, 2035–2041 (2008)
11. Gooding, O.W.: Process optimization using combinatorial design principles: parallel synthesis and design of experiment methods. *Curr. Opin. Chem. Biol.* **8**, 297–304 (2004)
12. Hibbert, D.B.: Experimental design in chromatography: a tutorial review. *J. Chromatogr. B* **910**, 2–13 (2012)
13. Ilzarbe, J., Alvarez, M.J., Viles, E., Tanco, M.: Practical applications of design of experiments in the field of engineering: a bibliographical review. *Qual. Reliab. Eng. Int.* **24**, 417–428 (2008)
14. Leardi, R.: Experimental design in chemistry: a tutorial. *Anal. Chim. Acta* **652**, 161–172 (2009)
15. Lutz, M.W., Menius, J.A., Choi, T.D., Gooding Laskody, R., Domanico, P.L., Goetz, A.S., Saussy, D.L.: Experimental design for high-throughput screening. *Drug Discov. Today* **1**(7), 277–286 (1996)
16. Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V., Hertzberg, R.P., Janzen, W.P., Paslay, J.W., Schopfer, U., Sittampalam, G.S.: Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **10**(3), 188–195 (2011)
17. Malo, N., Hanley, J.A., Carlile, G., Liu, J., Pelletier, J., Thomas, D., Nadon, R.: Experimental design and statistical methods for improved hit detection in high-throughput screening. *J. Biomol. Screen.* **15**(8) (2010)
18. Montgomery, D.C.: Design and Analysis of Experiments. Wiley, New York (1997)
19. Montgomery, D.C., Runger, G.C.: Applied Statistics and Probability for Engineers. Wiley, New York (2003)
20. Posner, B.A.: High-throughput screening-driven lead discovery: meeting the challenges of finding new therapeutics. *Curr. Opin. Drug Discov. Devel.* **8**(4), 487–494 (2005)
21. Pota, M., Pedone, A., Malavasi, G., Durante, C., Cocchi, M., Menziani, M.C.: Molecular dynamics simulations of sodium silicate glasses: optimization and limits of the computational procedure. *Comput. Mater. Sci.* **47**(3), 739–751 (2009)
22. Ranga, S., Jaimini, M., Sharma, S.K., Chauhan, B.S., Kumar, A.: A review on Design of Experiments (DOE). *Int. J. Pharm. Chem. Sci.* **3**(1), 216–224 (2014)

23. Scannell, J.W., Blanckley, A., Boldon, H., Warrington, B.: Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* **11**, 191–200 (2012)
24. Tye, H.: Application of statistical ‘design of experiments’ methods in drug discovery. *Drug Discov. Today* **9**(11), 485–491 (2004)
25. Weissman, S.A., Anderson, N.G.: Design of Experiments (DoE) and process optimization. A review of recent publications. *Org. Process. Res. Dev.* ACS Publications (2014)
26. Zang, C., Friswell, M.I., Mottershead, J.E.: A review of robust optimal design and its application in dynamics. *Comput. Struct.* **83**, 315–326 (2005)
27. Zheng, H., Clausen, M.R., Kastrup Dalsgaard, T., Mortensen, G., Bertram, H.C.: Time-saving design of experiment protocol for optimization of LC-MS data processing in metabolomic approaches. *Anal. Chem.* **85**, 7109–7116 (2013)