# Applying Design of Experiments Methodology to PEI Toxicity Assay on Neural Progenitor Cells

**Sara Mancinelli, Valeria Zazzu, Andrea Turcato, Giuseppina Lacerra, Filomena Anna Digilio, Anna Mascia, Marta Di Carlo, Anna Maria Cirafici, Antonella Bongiovanni, Gianni Colotti, Annamaria Kisslinger, Antonella Lanati, and Giovanna L. Liguori**

**Abstract** Design of Experiments (DoE) statistical methodology permits the simultaneous evaluation of the effects of different factors on experimental performance and the analysis of their interactions in order to identify their optimal combinations. Compared to classical approaches based on changing only one factor at a time (OFAT), DoE facilitates the exploration of a broader range of parameters combinations, as well as providing the possibility to select a limited number of combinations covering the whole frame. The advantage of DoE is to maximise the amount of information provided and to save both time and money. DoE has been primarily used in industry to maximise process robustness, but recently it has also been applied in biomedical research to different types of multivariable analyses,

---

The authors "Sara Mancinelli", "Valeria Zazzu" and "Andrea Turcato" contributed equally to this work.

The authors "Antonella Lanati" and "Giovanna L. Liguori" share senior co-authorship.

S. Mancinelli • V. Zazzu • G. Lacerra • G.L. Liguori (✉)
Institute of Genetics and Biophysics "A. Buzzati Traverso" (IGB), CNR, Naples, Italy
e-mail: giovanna.liguori@igb.cnr.it

A. Turcato • A. Lanati,
Valore Qualità, Pavia, Italy

F.A. Digilio
Institute of Biosciences and Bioresources (IBBR), CNR, Naples, Italy

A. Mascia • A.M. Cirafici • A. Kisslinger
Institute of Experimental Endocrinology and Oncology "G. Salvatore" (IEOS), CNR, Naples, Italy

M. Di Carlo • A. Bongiovanni
Institute of Biomedicine and Molecular Immunology "A. Monroy" (IBIM), CNR, Palermo, Italy

G. Colotti
Department of Biochemical Sciences, Institute of Molecular Biology and Pathology (IBPM), CNR, Sapienza University, Rome, Italy



V. Zazzu et al. (eds.), *Mathematical Models in Biology*,
DOI 10.1007/978-3-319-23497-7_4

from determination of the best cell media composition to the optimisation of entire multi-step laboratory protocols such as cell transfection.

Our case study is the optimisation of a transfection protocol for neural progenitor cell lines. These cells are very hard to transfect and are refractory to lipidic reagents, so we decided to set-up a protocol based on the non-lipidic Poliethylenimine (PEI) reagent. However, the effect of PEI toxicity on cells has to be correctly evaluated in the experimental design, since it can affect output computation. For this reason, we decided to apply DoE methodology to investigate the effect of PEI, both concentration and type, on cell viability and its interaction with other factors, such as DNA and cell density. The statistics-based DoE approach allowed us to express analytically the neural cell viability dependence on PEI amount/cell and efficiently identify the dose levels of PEI suitable for transfection experiments.

**Keywords** Design of experiments • PEI toxicity • Neural cell transfection • Factorial analysis

## 1 Design of Experiments as a Method for Protocol Optimization

Design of experiments (DoE) or experimental design is a methodology whose purpose is planning experiments and analysing their results by optimising the use of resources and time. At the same time, DoE is an effective tool for both maximising the amount of information and minimising the amount of data to be collected. In DoE, a process is seen as an input–output system with a measurable output that depends on the variations of multiple factors. Given a process or system, changes are made to the input variables and the effects on response variables (output) are measured. On this formal basis, factorial experimental designs allow the investigation of the effects of factors by varying them simultaneously instead of changing only one factor at a time (OFAT). OFAT is the most immediate approach to experimentation and is carried out by performing one or more tests for each value (level) of the independent variable (factor), leaving all the other conditions unchanged. The evaluation of the output effects induced by the variation of other factors should thus be obtained by repeating the same type of procedure for every single factor. Moreover, the evaluation of the effect of each factor, in correspondence with a precise combination of all the others, does not consider the interactions between the modelled factors. Therefore, a procedure based uniquely on the OFAT scientific method would omit the study of the effects of contemporary variations of two or more factors. Otherwise, a full OFAT model including all the possible interactions would require an uneven expense of time and resources. For this reason, choosing DoE methodology is the best option to optimise laboratory practices, such as transfection protocols, quickly and efficiently. A figurative representation of the two different methodological approaches is reported in Fig. 1.

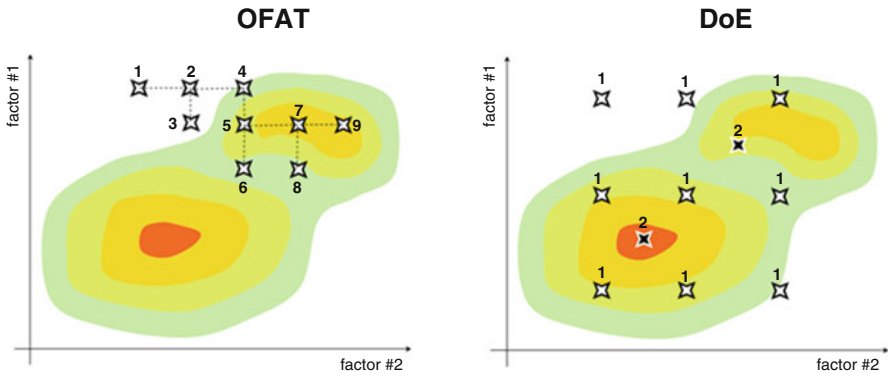


**Fig. 1** One factor at a time (OFAT) vs. Design of Experiments (DoE) methodology. Figurative representation of two different approaches to finding the optimal configuration of factors for the same process. System outputs are portrayed with a colour scale (from *white* to *red*) in a continuous bi-dimensional space that is to be explored by experimentation. The four-pointed stars represent all of the attempts made by researchers. Every attempt is a combination of two values, one for each factor, and the ordinal number associated with each attempt refers to the experimental step in which it is made. By varying one factor at a time, a stronger exploitation of attempts, steps and time is needed. Moreover, the results suffer from a higher risk of sub-optimality because of the possibility of arriving at a relative maximum without gaining a general understanding of the system

Candidate factors and their specific levels are selected, depending on the magnitude of their effect on the final result. The possibility of succeeding in arriving at a thorough comprehension of a biological process is strongly connected to the capability of identifying the most influential factors. DoE gives an estimate of the sensitivity of the output as a function of each factor, as well as of the combined effect of two or more factors, in a reduced number of trials (treatments or runs). A treatment corresponds to a determined set of factor levels, and the total number of runs depends on: (1) the experimental design, (2) the number of factors and (3) the replication factor of each experiment.

## 2 Main Aspects of DoE Methodology

DoE is mainly used for:

- Screening many factors and selecting the most relevant ones
- Discovering interactions among factors
- Executing an experiment lowering the risk of biases
- Verifying experimental assumptions and data consistency
- Analysing, interpreting and presenting the results
- Designing statistically robust protocols

- Establishing and maintaining Quality control, possibly by re-iteration of DoE and refinement of the underlying mathematical model

Some principles of DoE are Randomisation, Replication, Blocking, Orthogonality and Factorial experimentation. When applied, these characteristics/principles contribute to improve the robustness of scientific investigation and help researchers in developing experimental settings, so that successive trials can validate previous experimentations in a rigorous way.

Randomisation is accomplished by randomising the testing sequence. In this way, experimental results are protected against biases (e.g. temporal, order, operator-dependent).

Replication is a fundamental operation by which estimation precision is increased and uncontrollable noise is reduced at the same time. Signal-to-noise ratio is augmented by means of a replicate which is a complete repetition of the same experimental treatments and conditions, possibly in a randomised order. The higher costs due to an increase in the number of tests are thus balanced by a more accurate model parameters estimation.

Blocking improves accuracy by removing the effect of known nuisance factors when it is known that identical procedures are applied to each batch. The difference between two procedures is not influenced by the batch-to-batch differences. Blocking is a restriction of complete randomisation that, thanks to the subtraction of batch-to-batch variability from the "experimental error", increases estimation precision.

Orthogonality is used to generate results whose effects are uncorrelated and therefore can be more easily interpreted. The factors in an orthogonal experiment design are varied independently of each other. This makes it possible to summarise the collected data by taking differences of averages and to show main results graphically by using simple plots of suitably chosen sets of averages.

Finally, factorial experimentation requires that experimental designs include simultaneous, independent and orthogonal variations of all the factors. Since the total number of combinations increases exponentially with the number of factors studied, fractions of the full factorial design can also be constructed. The drawback of a reduction of tests in a fractional factorial design is the possibility of confounding between main effects and factors combinations effects.

Different experimental designs suitable for an experiment are: Plackett–Burman design, Box–Wilson (central composite) designs, Box–Behnken design, Factorial designs, equiradial designs (among them, Dohelert design), mixture designs and combined designs. The full factorial designs [3, 26] allow to estimate primary effects together with the effects of combinations of factors, called interactions, with a limited experimental and statistical complexity. Due to its experimental simplicity coupled with improved statistical efficiency [17], the full factorial design seems to be one of the most eligible approach in biological studies, in which the analysis of interactions between factors (e.g. genetic interactions, protein-protein interactions, gene-protein interactions) is becoming increasingly crucial. In this study we chose a full factorial design, constructing at least a twofold replicate design by keeping at

the same time the number of total runs as reasonably low as possible, without the risk of confounding effects. Non linearities were not studied in this initial phase, so a two-level full factorial approach was used.

In a two-level factorial screening experiment, every chosen factor varies between two levels: qualitative factors have two categorical values (e.g. low/high, A/B, left/right), while quantitative factors vary between two numeric values. Given this design, if N is the number of candidate factors, the number of requested different runs is $r = 2^N$. The number of different runs r must always be greater than the rank of the design matrix X that has to be estimated. Finally the number of total runs t is obtained by multiplying the number of treatments by the replication factor *k*, that is the number of times each single treatment is repeated, leading to $t = k*r = k*2^N$. The coefficient k is 1 if the experiment is not replicated. The treatments are then executed in a randomised order to avoid having uncontrolled variables (i.e. not modelled as factors) contribute to the repeatability variance, affecting the results in a systematic way [3, 17, 26]. This method relies upon the statistical estimation of parameters, which are factors with main effects and factorial interactions. Every parameter is estimated by a mathematical model whose aim is to explain the variability of the output by a combination of factor effects and their interactions, in the form $y = ax_1 + bx_2 + ..cx_1x_2 \ \ldots$ where $x_i$ are the modelled factors and a, b, c the parameters identified (e.g. *a*, *b*: main effects; *c*: factors interaction). The accuracy of parameter estimation is calculated by the coefficient of determination $R^2$, that is a measure of the percentage of data variability explained by the model and it is used in the multiple linear regression analysis.

An additional measure is adjusted $R^2$, that integrates knowledge on the number of modelled variables into a score for the goodness of fit. Its choice is suggested when two models with a different number of factors are compared.

The appropriateness of the estimated mathematical model can be effectively visualised by means of simple plots that show magnitude, whiteness and distribution of residuals. Residual distribution analysis is a way to visualize the mathematical model's fitness for the system under study. Residuals must be low in magnitude and distributed normally. A model's failure to fit must lead to a redefinition of factor levels or to revision of the design itself. An ANOVA analysis can be associated with the DoE analysis, enriching the statistical evaluation of the experiment.

## 3 DoE Applications from Industry to Biological Research

DoE methodology has been widely used in the field of industrial design for the development of processes in order to improve performance. In this field, the primary objectives of these experiments were to: (1) determine the most influential variables on the response, (2) increase product volume, (3) reduce variability. Factorial design has been used for the optimisation of protocols in a variety of industrial fields including manufacturing [10, 11, 13, 25], as well as for pharmaceutical studies

within the Quality-by-design approach to define the design space for standardised production processes (ICH Q8 2009; [28].

Recently, DoE has been playing an important role also in scientific research areas such as food science [15], chemistry [4] and engineering [22]. The application of DoE has brought very good results in biological fields, among them chromatography [17], metabolomics [29] and especially cellular biology and tissue engineering [2, 4, 7, 12, 18–21, 23, 24]. Optimising the conditions for a specific process in an OFAT manner is a time consuming operation and does not take into account interdependency between factors, which is likely to play a role in most biological processes. Moreover, since the definition of biological protocols has to deal with different environmental conditions, robust estimations of variable parameters and easy visualisation of results are needed to really understand the biological system under observation.

In cellular biology, DoE-based strategies have been applied to develop and optimise serum-free media for culturing mesenchymal stem cell (MSC) spheroids by systematically evaluating media mixtures and a panel of different components for their effects on cell proliferation [2]. Moreover, a factorial Quality-by-design approach has been combined with other approaches such as high-throughput mRNA profiling of a customised chondrogenesis-related gene set as a tool to study in vitro chondrogenesis of human bone marrow derived MSC. The analysis identified the best of the tested differentiation cocktails 21. Scientists have taken advantage of DoE methodology not just to screen different components of a culture medium, but also to apply it to optimise an entire protocol, such as specific cell line transfection and protein production, obtaining promising results. It has been shown that DoE significantly improves transfection efficiency by a global economy of materials and time [5]. Transfection is the transient or stable introduction of exogenous molecules and genetic material, DNA or RNA, into cultured mammalian cells and is commonly utilised in biological laboratories to study gene function, modulation of gene expression, biochemical mapping, mutational analysis, and protein production. No single delivery method or transfection reagent can be applied to all types of cells, and neural cells are among the most difficult cells to transfect [9, 20]. Very importantly, cellular cytotoxicity and transfection efficiencies vary dramatically depending on the reagent, protocol and cell type being utilised.

## 4 Our Case Study: The Set-Up and Optimisation of a Transfection Protocol for Neural Progenitor Cells

A groundbreaking project named Quality and Project Management OpenLab (qPMO), inspired by Quality and Project Management principles, has been implemented by a network of Italian National Research Council (CNR) Institutes with the aim of realising and disseminating within the scientific community an innovative way to plan and organise research activity (http://quality4lab.cnr.it) [6].

In this context, we selected the DoE model as a very interesting and promising methodology suitable for different kinds of scientific experiments, and we planned

to apply it to both simple assays and medium-high throughput experiments with the final aim of identifying a general guideline for the application of DoE to the set-up and optimisation of scientific protocols.

Our case study is the optimisation of a transfection protocol for neural progenitor for mes-c-myc A1 cell line to obtain a standardised and reproducible laboratory procedure. Mes-c-myc A1 cells are immortalised neural progenitor cells (NPCs) derived from mesencephalon of mouse embryos at 11 days of development which have the characteristics of self-renewal and multipotency [8]. It has been shown that mes-c-myc A1 cells, as is typical of all neural cells, are difficult to transfect and are low-responding to traditional lipidic transfection methods [20]. For this reason, Polyethyleneimine (PEI) was chosen as a transfection reagent because PEI is a cationic non-lipidic transfection reagent normally chosen to achieve higher transfection efficiencies in cell lines that are refractory to liposome-based transfection [19]. A number of PEI molecules have been described in detail with varying molecular size or structure: branched (B) PEI with an average molecular weight of 800 kDa (PEI800) and 25 kDa (PEI25) and a linear (L) form with an average molecular weight of 22 kDa (PEI22) with high transfection activity in vitro and in vivo [28].

Among the factors relevant for transfection efficiency, we selected three quantitative factors: (1) concentration of PEI, (2) DNA amount, (3) cell density and a qualitative one, the type of PEI, L (22 kDa) vs. B (25 kDa). PEI transfectant, as a polyethyleneimine molecule, binds DNA by the presence of nitrogen positive cations (N) in its structure that attract phosphate negative ions (P) of DNA. N/P ratios, depending on PEI amount, have a dramatic impact on transfection efficiency and cytotoxicity [14]. A high concentration of PEI provides a high number of nitrogen cations that can bind DNA efficiently, but which can also be toxic for cells. For this reason, it is important to investigate the correct proportions of PEI and DNA amount, depending of cell line and cell density. Likewise, PEI concentration as well as PEI structure play an important role in both transfection efficiency and cytotoxicity. B conformation of PEI provides a large amount of amine groups that bind DNA more efficiently with respect to the L type. B-PEIs have stronger binding affinity which can condense DNA more efficiently, but they have a less effective release, leading to reduced transfection efficiency. B-PEIs are also more toxic, reducing the viability of cells for transgene expression [19].

To evaluate the percentage of transfected cells, we used ImageJ software to minimise error and variation in downstream analysis [1]. A DNA plasmid containing a green fluorescent protein (GFP) reporter was chosen for transfection to help the output calculation: cells expressing pIRES-EGFP (transfected cells) were visualised directly by fluorescence microscope after PFA 4 % fixation and Hoechst counterstaining for nuclei (all cells). ImageJ java-based image processing software was used for images processing. A plug-in was created to automate and standardize cell counting of total cells (blue labelled nuclei). GFP positive cells (Green labelled,

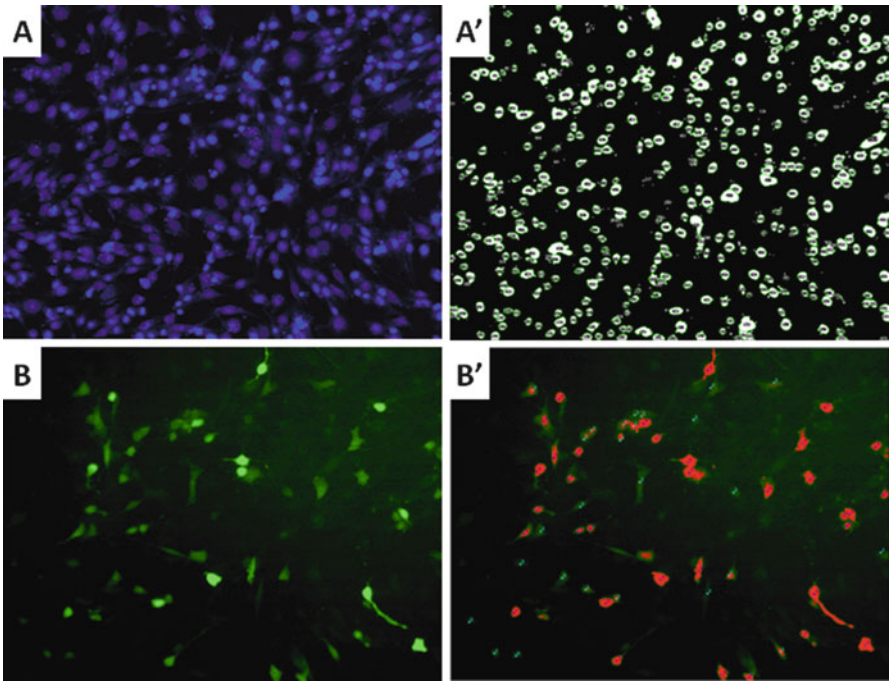


**Fig. 2** Transfection efficiency calculation through ImageJ software. Transfected Mes-c-myc A1 cells were visualized by fluorescence microscope. (**A**) Hoechst-labelled mes-c-myc A1 cells (all cells) and (**B**) cells expressing GFP (transfected cells). (**A'**) a plug-in was created using ImageJ java-based image processing program to automate and standardize cell counting of total cells. GFP positive cells (transfected cells) were counted by summing the number of cells captured by the threshold (*red cells*, **B'**) and the cells presenting a weak signal undetectable for the threshold calculated by the software (**B**, **B'**)

transfected cells) were counted by summing the number of cells captured by the default threshold calculated by the software (red cells) and the cells presenting a weak signal undetectable for the threshold (Fig. 2).

## 5 DoE Methodology Applied to Toxicity Assay

In transfection experiments an important issue to consider is the cell toxicity of the transfectant agent. In our case, for example, a low transfection efficiency could be due to two opposite conditions, an inadequate amount of transfectant and/or DNA or too much transfectant, which can be toxic for the cells.

For this reason, we decided to apply DoE methodology to evaluate cell viability depending on PEI concentration and PEI type in the culture medium, which are the most critical factors, and their interactions with other important factors, such as

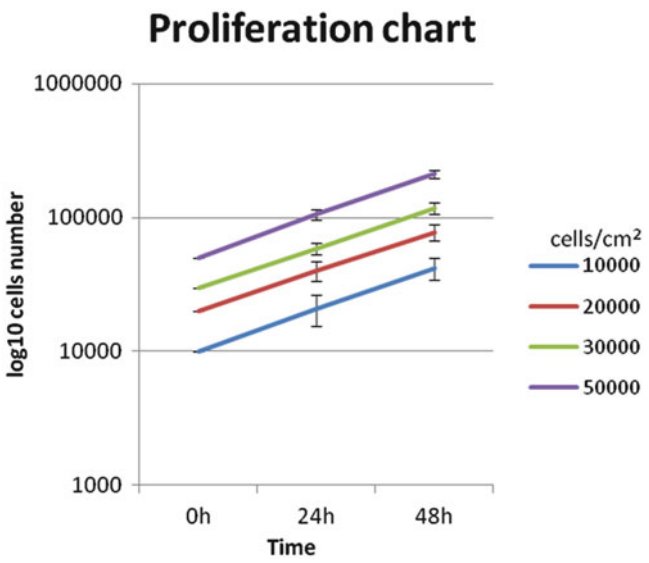


**Fig. 3** Proliferation Chart. Proliferation trend the of mes-c-myc-A1 cell line related to the number of cells seeded. The four lines are parallel, indicating that the proliferation rate does not vary with the increase of cell density

DNA concentration and cell density. Before that, we verified that between 10,000 and 50,000 cells/cm$^2$ the proliferation rate was constant and corresponded to the expected one [8]. In this interval, the cells are in a logarithmic phase of proliferation (Fig. 3), and this behaviour is fundamental for DNA uptake efficiency.

Here we show a flowchart describing every experimental and analytical phase of our DoE approach (Fig. 4). Once the calculation of the output (cell viability) was defined, factors and their levels were chosen, the design of the experiment was generated by means of Minitab® Statistical Software and stored in a worksheet. Cells were seeded into 24-well plates (each well is approximately 2 cm$^2$) the day before the treatment. The treatment was performed by adding to cells a solution made with an appropriate PEI amount, with and without DNA, according to the design created, in a finale volume of 100 μl of MEM/F12 culture medium without antibiotic and serum. The incubation lasted over night at 37 °C in 5 % $CO_2$. The next day, culture medium containing the testing amount of PEI and DNA was replaced with fresh complete culture medium. Twenty-four hours after the treatment, viability (output) was calculated as percentage of alive cells respect to the total ones. Alive and dead cells were calculated as follows: culture medium from each condition was collected in different 1.5 ml tubes in order to draw up dead cells in suspension. Next, attached-alive cells were treated with 0.1 % Trypsin for 1 min at 37 °C; thereafter, cells were collected together with their correspondent cells suspension in the same tube. Subsequently, each tube was centrifuged at 900 times gravity for 3 min and the pellet was re-suspended in 200–400 μl of culture medium. At the end, 10 μl of cell suspension were mixed with 10 μl of Trypan Blue (Dye of dead

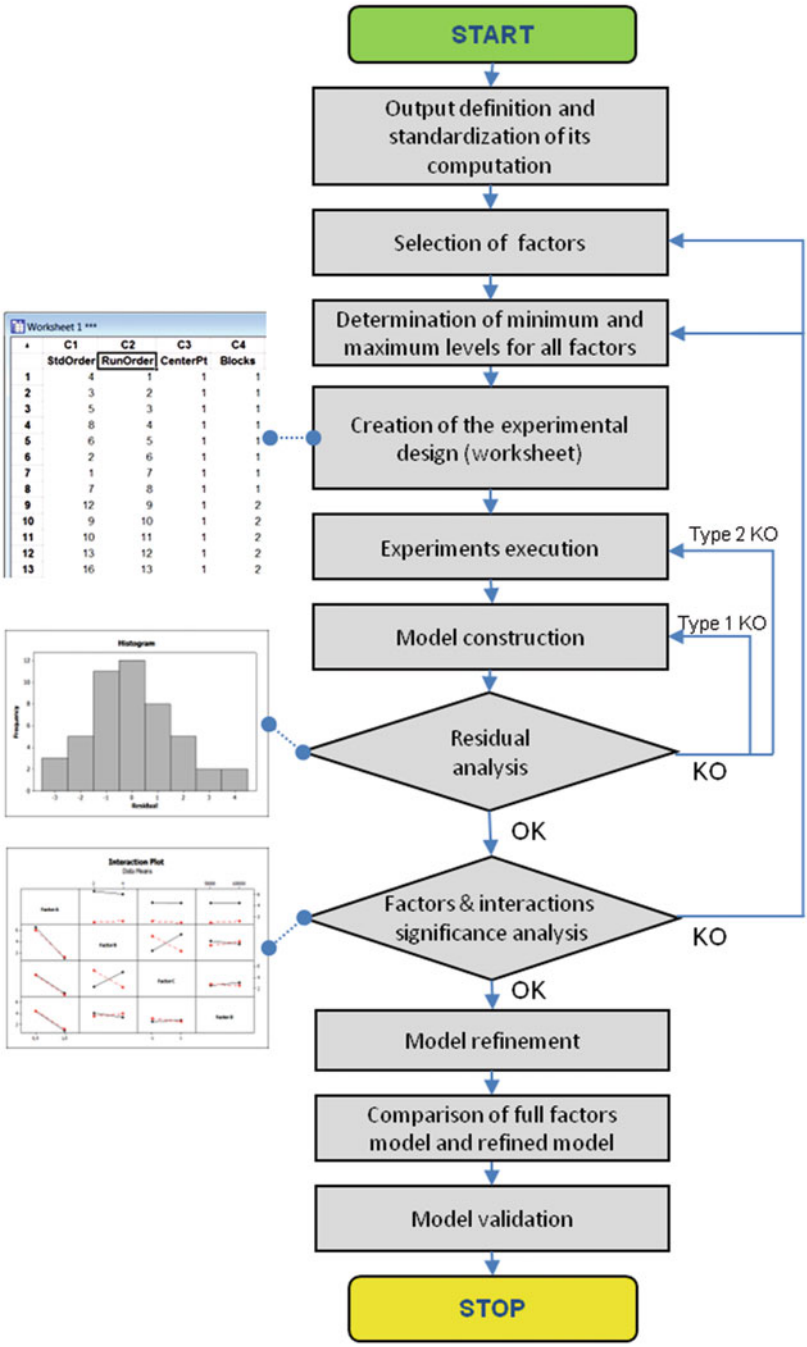


**Fig. 4** (continued)

cells) and placed into a haemocytometer. Alive (unstained) and dead cells (stained) were counted under a microscope at 10× magnification. The number of total cells for each well was calculated by summing up alive and dead cells. After collecting data, a linear model was fitted to the data and graphs were generated to evaluate the effects. Residuals plot graphs allowed the evaluation of how accurately the model (model with full factors and interactions) fitted the data. Main Effects and Interaction Effects Plots were analysed to understand the effects on the response of each factor and their interactions. Subsequently the full factor model was compared with a refined model that included only significant factors and interactions.

The model chosen was validated by random experiments using different PEI concentrations and cell densities. After the executions of these experiments, we calculate the amount of PEI/cell used in each condition. Once calculated the output of the random conditions experimented (cell viability) we plotted the results together with the refined model to verify the accuracy of the model constructed by DoE approach.

## *5.1 Factorial Designs and Residual Analysis*

Different open source and commercial software are available to create factorial designs and analyse the responses: e.g. Minitab® Statistical Software, DOE++, Design-ease, JMP, Develve, MaxStat professional, MacANOVA. Among them we chose Minitab® Statistical Software which offers four types of experimental designs: factorial, response surface, mixture and Taguchi (robust). By default, Minitab® Statistical Software randomises the run order of all design types. Randomisation helps to ensure that the model meets certain statistical assumptions and allows the reduction of the effects of factors not included in the study. Four factors were put under study: (1) PEI concentration, (2) PEI type, (3) presence or absence of DNA and (4) cell density. We chose a *two level* (each factor varies between two levels) *full* (all combinations are included; no reduction of the design) *factorial design* (all factors are varied at the same time) to analyse the effects of the four factors considered on the output. The execution order of each treatment was performed according to a 'RunOrder' column, which insures the correct randomisation of runs (16 combinations in duplicate). The worksheet and the rough results obtained are shown in Fig. 5.

**Fig. 4** DoE experimental flowchart. Schematic representation of the experimental workflow for DoE statistical analysis. To create the experimental design and analyse the results, Minitab® Statistical Software was used. *Rectangles* represent processes; *rhombuses* represent flow checkpoints. Two main checkpoint are considered: Residual Analysis and Factors and interactions significance analysis. KO indicate that checkpoint has not been overcome. Type 1 KO: residuals have not a normal distribution with mean 0. Type 2 KO: residual analysis shows presence of bias or outliers. At the end of the analysis the refined model was validated

**a** Two level full factorial design worksheet for Toxicity assay

| StdOrder | RunOrder | DNA (0,5µg/ml) | PEI type | PEI conc. (mg/L) | Cell Density (cells/cm²) | Viability (%) |
|---|---|---|---|---|---|---|
| 2 | 1 | yes | B | 6 | 25000 | 85.81 |
| 12 | 2 | yes | L | 6 | 50000 | 100.00 |
| 1 | 3 | no | B | 6 | 25000 | 97.73 |
| 18 | 4 | yes | B | 6 | 25000 | 85.90 |
| 9 | 5 | no | B | 6 | 50000 | 97.34 |
| 22 | 6 | yes | B | 18 | 25000 | 5.88 |
| 31 | 7 | no | L | 18 | 50000 | 61.32 |
| 26 | 8 | yes | B | 6 | 50000 | 91.36 |
| 32 | 9 | yes | L | 18 | 50000 | 40.50 |
| 24 | 10 | yes | L | 18 | 25000 | 15.44 |
| 19 | 11 | no | L | 6 | 25000 | 100.00 |
| 16 | 12 | yes | L | 18 | 50000 | 45.42 |
| 10 | 13 | yes | B | 6 | 50000 | 88.27 |
| 8 | 14 | yes | L | 18 | 25000 | 25.97 |
| 20 | 15 | yes | L | 6 | 25000 | 96.94 |
| 3 | 16 | no | L | 6 | 25000 | 99.48 |
| 30 | 17 | yes | B | 18 | 50000 | 37.35 |
| 27 | 18 | no | L | 6 | 50000 | 100.00 |
| 5 | 19 | no | B | 18 | 25000 | 7.94 |
| 11 | 20 | no | L | 6 | 50000 | 98.48 |
| 15 | 21 | no | L | 18 | 50000 | 26.47 |
| 29 | 22 | no | B | 18 | 50000 | 37.00 |
| 14 | 23 | yes | B | 18 | 50000 | 42.45 |
| 21 | 24 | no | B | 18 | 25000 | 0.00 |
| 17 | 25 | no | B | 6 | 25000 | 100.00 |
| 28 | 26 | yes | L | 6 | 50000 | 100.00 |
| 7 | 27 | no | L | 18 | 25000 | 27.60 |
| 23 | 28 | no | L | 18 | 25000 | 32.52 |
| 4 | 29 | yes | L | 6 | 25000 | 94.39 |
| 25 | 30 | no | B | 6 | 50000 | 93.45 |
| 6 | 31 | yes | B | 18 | 25000 | 4.83 |
| 13 | 32 | no | B | 18 | 50000 | 26.94 |

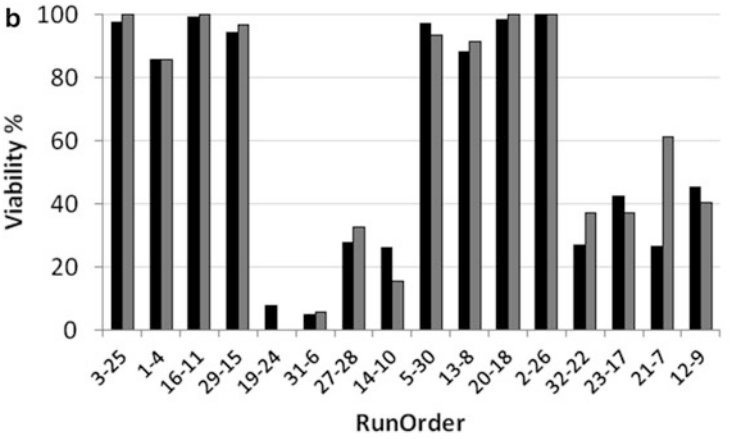


**Fig. 5** (continued)

For the statistical analysis of the results, a regressive linear model was fitted to the observations and residuals were analysed (Fig. 6). The distribution appeared to fit the linear model: the Histogram, the graph that correlates the residuals with their frequency, appears to be approximately symmetric and bell-shaped, confirming the normal distribution of residuals (Fig. 6A). The Versus Order plot showed randomly scattered residuals with the absence of significant patterns in the distribution: this demonstrates the time independence of residuals, non-constant variance and missing higher-order terms (Fig. 6B). Once the normality distribution of the residuals was determined, the analysis of factor interactions was performed.

## *5.2 Factors and Interactions Significance Analysis*

To verify the significance of the factors and their interactions, a Pareto Chart was generated in which any effect extending beyond the reference red line was significant at the default level of 0.05 (Fig. 6C). As shown by the graphs, three different factors were found to be significant (PEI type, PEI concentration and cell density) and only one two-factor interaction (between PEI concentration and cell density). The presence of DNA did not influence the response. This observation might support the idea that negatively charged DNA molecules do not balance positive charges of PEI transfectant, affecting toxicity in some way.

To interpret the results, Main Effects Plots and an Interactions Plot were analysed (Fig. 7). A main Effects Plot shows the one-factor effect called main effect (Fig. 7A). The horizontal line, corresponding to about 60 (60 % of cell viability), represents the mean of the response of all the runs. The line for DNA confirmed what was shown in the Pareto Chart, that is, in both conditions of absence (no) or presence (yes) of DNA the mean of all conditions was approximately 60 %, resulting in a line with slope close to 0 that indicates the non-influence of the factor. The most important factor was PEI amount. The effect of this factor was the line with the highest slope. In detail, all runs in which 6 mg/L of PEI transfectant were utilised showed a higher cell viability (close to 100 %, total cell viability) with respect to the condition with 18 mg/L. PEI type and cell density have a similarly low slope, indicating that these factors had a comparable small effect on the response. The PEI type plot shows that the mean output of all the runs performed with L-PEI is higher if compared to the one of all the runs performed with B-PEI, confirming the lower

**Fig. 5** Toxicity assay performed with a two-level full factorial experimental design. (**A**) Worksheet created by Minitab® Statistical Software representing combinations of tested factors and levels. Standard order indicates the order in which combinations are generated according to the design chosen. Run order corresponds to the randomisation of generated combinations. DNA, PEI type, PEI concentration and Cell density are the factors analysed. Viability represents the calculated output. (**B**) Graphical representation of the results, in which the two series represent the two replicates (*black*, replicate 1; *gray*, replicate 2)

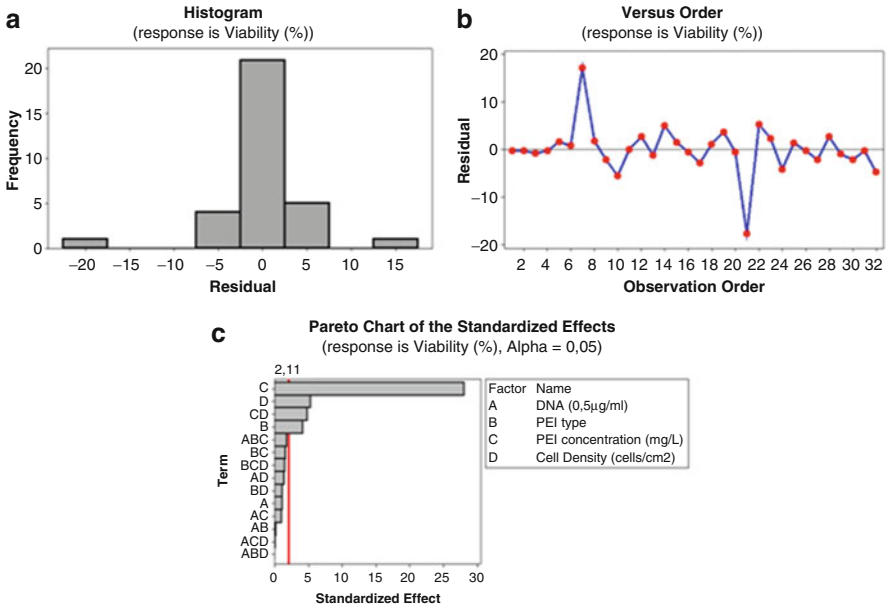


**Fig. 6** Residual and distribution analysis for Toxicity assay. (**A**) Histogram shows that the distribution of the residuals in the experiment is normal. (**B**) Versus order plot illustrates that observation order of the residual is well randomised, due to the absence of a repetitive pattern in the plot. (**C**) Pareto Chart of the Standardized Effects graphically represents the significant factors and interactions. Any significant factor or interaction is characterised by columns that extend beyond the *red line*; the greater the distance from the line, the higher the influence of that factor. Significant factors and interactions were, in order: PEI concentration (factor C), cell density (factor D), interaction between PEI concentration and cell density (CD) and PEI type (factor B)

toxicity of linear conformation of the PEI molecule with respect to the branched one [19]. In the cell density plot, all runs with the factor of 25,000 cells/cm$^2$ showed a lower cell viability with respect to runs with 50,000 cells/cm$^2$, demonstrating cell viability depended on the amount of transfectant per cell. This analysis confirmed PEI concentration as the most critical parameter and identified PEI type and cell density as influencing factors.

To identify significant interactions among factors, the Interactions Plot showing the effect of multiple factors was also analysed (Fig. 7B). Evaluating interactions is extremely important because an interaction can magnify or diminish main effects. All the interactions were not significant because the two lines had the same slope, except for the plot showing the interaction between PEI (mean of the values of both L and B) concentration and cell density. In this case, the two different levels of PEI concentration exhibited two different behaviours: at the lower PEI concentration (black line), cell viability was always approximately 100 % for both cell density conditions, while with the higher PEI concentration (red line), cell viability was higher when cell density was 50,000 cells/cm$^2$ (around 40–50 %) and

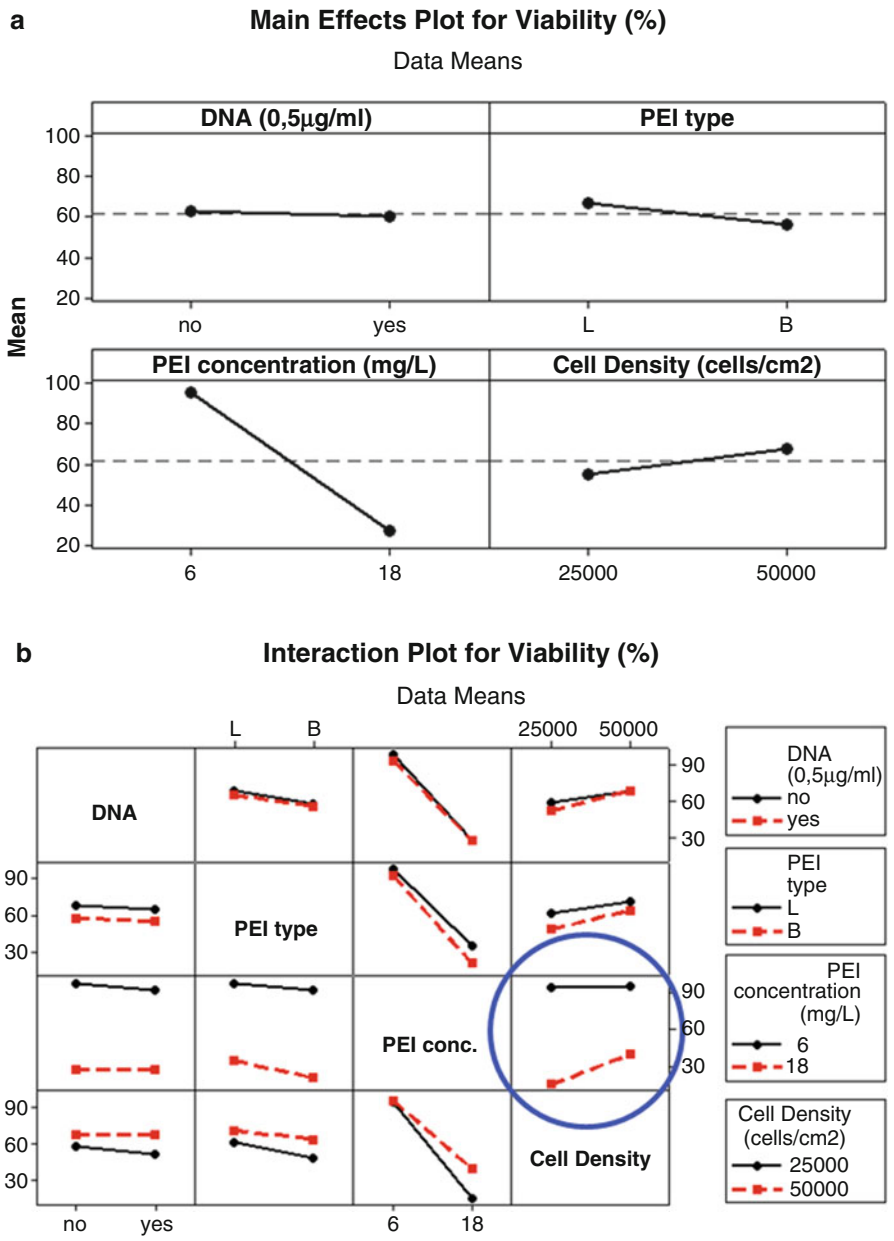


**Fig. 7** Factorial analysis for Toxicity assay. (**A**) The Main effect plot shows the mean of outputs in correspondence with the different levels of each of the four factors. The *dotted line* corresponds to a viability of 60 %. The plot shows that PEI concentration is the factor with the strongest effect on cell viability, due to the higher slope of the line connecting the two analysed levels. (**B**) The Interaction plot is represented by a matrix showing interactions among factors. Similar slope of the lines indicates no significant interaction. The only significant interaction is the one between PEI concentration and cell density, highlighted in the *circle*

lower when cell density was 25,000 cells/cm$^2$ (approximately 0 %, corresponding to full mortality). These data show that the minimum amount of PEI used (6 mg/L) never affected cell viability, while the maximum amount (18 mg/L) decreased cell viability between 0 and 50 %, depending on cell density.

### *5.3 Model Refining and Validation*

A refined model was constructed discarding all the not significant main effects and interactions: only the significant factors PEI type (B), PEI concentration (C), cells density (D) and the interaction between these last two factors (C*D) were saved for the model refinement. Subsequently, the full factors model and the refined model were compared. Since full factors model had given as a fitness measure $R^2$ adjusted of 96.24 % and the refined model gave $R^2$ adjusted of 95.96 %, the loss of explained variance was considered minimal and the refined model was chosen for further investigations. The analysis of $R^2$, including the adjustment of extra explanatory variables, gave us additional evidence to exclude from the analysis not significant factors, such as DNA amount, and irrelevant factors interactions.

## 6 Conclusions

The DoE approach applied to our Toxicity assay let us clearly determine which factors most influenced the output of the process under study: PEI concentration and cell density. Moreover, the simultaneous variation of these factors and the subsequent statistical analysis let us identify a significant interaction between PEI concentration and cell density, unmasking the real significant factor influencing cell viability, that is PEI amount per cell. By relating cell viability measured in the Toxicity assay to the amount of PEI, both B and L, per cell tested (PEI pg/cell) we could finally generate a refined model (Fig. 8). To validate this model, we determined cell viability at different levels of PEI pg/cell in the examined interval. All the conditions tested were reasonably close to the line, indicating that the model generated by the factorial analysis had good accuracy and could be used to predict cell viability variation connected to the dose of PEI per cell. The Toxicity assay let us select the upper value of PEI pg/cell to be tested in following transfection experiments, in order to obtain the maximum transfection efficiency avoiding extreme cell death. Thus, we decided to set the maximum level at a concentration of PEI in the culture medium of 12 mg/L, corresponding to 30 PEI pg/cell (cell density 50,000 cells/cm$^2$), which would not reduce cell viability more than 50 %.

In summary, taking advantage of statistics-based factorial experimental design we could express analytically mes-c-myc A1 cell viability dependence on PEI amount per cell. Application of DoE allowed to determine the maximum amount

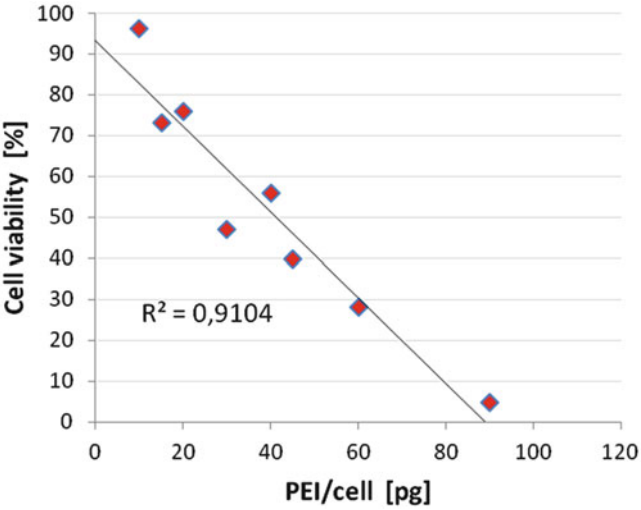


**Fig. 8** Dependence of cell viability on PEI amount/cell. A polynomial model was fitted with Toxicity assay viability results, $y = -1.05x + 93.19$ where x: PEI/cell, y: viability ($R^2 = 0.9104$). *Red rhombuses* correspond to experimental conditions chosen to test the factorial analysis

of PEI usable in mes-c-myc A1 cells transfection experiments coupled to the higher cell viability. The main advantages of using DoE were saving time and resources for the complete experimental plan (leading to efficiency), the evaluation of both main and interaction effects of the selected factors in an easy graphical way, and statistical information that allows data to be robust and reliable (both of which lead to effectiveness). Nowadays, with an increasing scientific competition asking researchers to produce in a short time reliable and reproducible results, our data support the application of DoE to scientific studies for obtaining the best results and optimising the use of resources.

**Acknowledgments** We would like to thank Umberto di Porzio and Giancarlo Bellenchi (Institute of Genetics and Biophysics 'Adriano Buzzati-Traverso') for helpful suggestions and for providing us with the mes-c-myc A1 cells, Salvatore Pulcrano and Valerio Piscopo (Institute of Genetics and Biophysics 'Adriano Buzzati-Traverso') for providing A1 cell culturing protocols, Genesia Manganelli and Emilia Giorgio (Institute of Genetics and Biophysics 'Adriano Buzzati-Traverso') for helping set up the preliminary assays, and Teresa Nutile (Institute of Genetics and Biophysics 'Adriano Buzzati-Traverso') for her statistical support. Portions of information contained in this work are printed with permission of Minitab Inc. All such material remains the exclusive property and copyright of Minitab Inc. All rights reserved. We also thank Richard E. Burket for editing and English revision.

This work was supported by grants from the Ministero dell'Economia (Ministry of Economics and Finance in Italy, CNR FaReBio di Qualità, qPMO Project), the Ministero Istruzione Università Ricerca (Medical Research in Italy RBNE08LN4P_002) and 'Fondazione con il Sud' (2011-PDR-13) to G.L.L.

## References


1. Abramoff, M.D., Magalhaes, P.J., Ram, S.J.: Image processing with ImageJ. Biophoton. Int. **11**, 36–42 (2004)
2. Alimperti, S., Lei, P., Wen, Y., Tian, J., Campbell, A.M.: Serum-free spheroid suspension culture maintains mesenchymal stem cell proliferation and differentiation potential. Biotechnol. Prog. **30**, 974–983 (2014)
3. Anderson, M.J., Whitcomb, P.J.: DOE Simplified. Practical tools for effective experimentation. Productivity Press, New York (2007)
4. Bezerra, M.A., Santelli, R.E., Oliveira, E.P., Villar, L.S., Escaleira, L.A.: Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta **76**, 965–977 (2008)
5. Bollin, F., Dechavanne, V., Chevalet, L.: Design of experiment in CHO and HEK transient transfection condition optimization. Protein Expr. Purif. **78**, 61–68 (2011)
6. Bongiovanni, A., Colotti, G., Liguori, G.L., Di Carlo, M., Digilio, F.A., Lacerra, G., Mascia, A., Cirafici, A.M., Barra, 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**, 203–213 (2015)
7. Chen, Y., Bloemen, V., Impens, S., Moesen, M., Luyten, F.P., Schrooten, J.: Characterization and optimization of cell seeding in scaffolds by factorial design: quality by design approach for skeletal tissue engineering. Tissue Eng. Part C Methods **17**, 1211–1221 (2011)
8. Colucci-D'Amato, G.L., Tino, A., Pernas-Alonso, R., ffrench-Mullen, J.M., di Porzio, U.: Neuronal and glial properties coexist in a novel mouse CNS immortalized cell line. Exp. Cell Res. **252**, 383–391 (1999)
9. Dalby, B., Cates, S., Harris, A., Ohki, E.C., Tilkins, M.L., Price, P.J., Ciccarone V.C.: Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. Methods. **33**, 95–103 (2004)
10. Dhas, E.R., Dhas, J.H.: A review on optimization of welding process. Procedia Eng. **38**, 544–554 (2012)
11. Dirion, B., Cobb, Z., Schillinger, E., Andersson, L.I., Sellergren, B.: Water-compatible molecularly imprinted polymers obtained via high-throughput synthesis and experimental design. J. Am. Chem. Soc. **125**, 15101–15109 (2003)
12. Enochson, L., Brittberg, M., Lindahl, A.: Optimization of a chondrogenic medium through the use of factorial design of experiments. Biores Open Access **1**, 306–313 (2012)
13. Fei, N.C., Mehat, N.M., Kamaruddin, S.: Practical applications of Taguchi method for optimization of processing parameters for plastic injection moulding: a retrospective review. ISRN Ind. Eng. **2013**, 1–11 (2013)
14. Florea, B.I., Meaney, C., Junginger, H.E., Borchard, G.: Transfection efficiency and toxicity of polyethylenimine in differentiated Calu-3 and nondifferentiated COS-1 cell cultures. AAPS PharmSci. **4**, E12 (2002)
15. Gacula, M.C.: The design of experiments for shelf life study. J. Food Sci. **40**, 399–403 (2006)
16. Green, S., Liu, P.Y., O'Sullivan, J.: Factorial design considerations. JCO **16**, 3424–3430 (2002)
17. Hibbert, D.B.: Experimental design in chromatography: a tutorial review. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. **910**, 2–13 (2012)
18. Hubert, C., Lebrun, P., Houari, S., Ziemons, E., Rozet, E., Hubert, P.: Improvement of a stability-indicating method by Quality-by-Design versus Quality-by-Testing: a case of a learning process. J. Pharm. Biomed. Anal. **88**, 401–409 (2014)
19. Hsu, C.Y., Uludağ, H.: A simple and rapid nonviral approach to efficiently transfect primary tissue–derived cells using polyethylenimine. Nat. Protoc. **7**, 935–945 (2012)
20. Karra, D., Dahm, R.: Transfection techniques for neuronal cells. J. Neurosci. **30**, 6171–6177 (2010)
21. Knöspel, F., Schindler, R.K., Lübberstedt, M., Petzolt, S., Gerlach, J.C., Zeilinger, K.: Optimization of a serum-free culture medium for mouse embryonic stem cells using design of experiments (DoE) methodology. Cytotechnology **62**, 557–571 (2010)

22. 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)
23. Jakobsen, R.B., Østrup, E., Zhang, X., Mikkelsen, T.S., Brinchmann, J.E.: Analysis of the effects of five factors relevant to in vitro chondrogenesis of human mesenchymal stem cells using factorial design and high throughput mRNA-profiling. PLoS One **9**, e96615 (2014)
24. Liu, G., Kawaguchi, H., Ogasawara, T., Asawa, Y., Kishimoto, J., Takahashi, T., Chung, U.I., Yamaoka, H., Asato, H., Nakamura, K., Takato, T., Hoshi, K.: Optimal combination of soluble factors for tissue engineering of permanent cartilage from cultured human chondrocytes. J. Biol. Chem. **282**, 20407–20415 (2007)
25. Mandenius, C.F., Brundin, A.: Biocatalists and bioreactor design—bioprocess optimization using design-of-experiments methodology. Biotechnol. Prog. **24**, 1191–1203 (2008)
26. Montgomery, D.C.: Design and Analysis of Experiments. Wiley, New York (1997)
27. Weissman, S.A., Anderson, N.G.: Design of Experiments (DoE) and Process Optimization. A Review of Recent Publications. Organic Process Research & Development. ACS Publications (2014)
28. Wightman, L., Kircheis, R., Rössler, V., Carotta, S., Ruzicka, R., Kursa, M., Wagner, E.: Different behavior of branched and linear polyethylenimine for gene delivery in vitro and in vivo. J. Gene Med. **3**(4), 362–372 (2001)
29. Zheng, H., Clausen, M.R., Dalsgaard, T.K., 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)