Design of Experiments for the Development of Injectable Drug Products

5

Dhawal Chobisa

Abstract

Design of experiments (DoE) is a widely used statistical tool for planning experiments, collecting and analyzing data, and drawing valid conclusions. This chapter describes the basics of DoE, types of DoE designs, and rationale for the selection of a design. Applications of DoE in the development of pharmaceutical drug products are discussed with emphasis on injectable drug products. Also, a practical case study of the development of a nanoemulsion product is discussed in detail.

Keywords

Design of experiments (DoE) · Injectable drug products · Screening and optimization designs · DoE in the pharmaceutical drug products · Nanoemulsion · Design analysis · Residual and influence diagnostics

5.1 Introduction

The objective of the development of pharmaceutical drug products is to deliver safe and efficacious medicines to the patients reliably. It is crucial to ensure the desired drug product quality reproducibly. Failure to achieve the quality of drug products can lead to severe safety concerns and suboptimal therapeutic benefits for the patients. In early 2000, Quality by Design (QbD) based drug product development

D. Chobisa (✉)
Department of Industrial and Physical Pharmacy, College of Pharmacy, Purdue University, West Lafayette, IN, USA

Integrated Product Development Organization, Innovation Plaza, Dr. Reddy's Laboratories, Hyderabad, India

© Springer Nature Singapore Pte Ltd. 2021

S. Beg (ed.), *Design of Experiments for Pharmaceutical Product Development*,
https://doi.org/10.1007/978-981-33-4351-1_5

was adopted by the United States Food and Drug Administration (US-FDA) and other regulatory authorities to ensure the quality of the products [1–3].

Pharmaceutical drug products can be administered by various routes such as oral, topical, and injectable administration. The safety margin is very narrow with injectable (e.g. intravenous) dosage forms as they bypass the absorption step and quickly access the systemic circulation. QbD based development of injectable drug products involves the designing of drug delivery systems such as liposomes, polymeric nanoparticles, lyophilized powder, solutions, suspensions, and emulsions, selection of excipients, and product composition. Also, robust manufacturing processes and analytical methods are crucial to develop quality drug products. The understanding of unit operations such as sterilization is a unique challenge faced during the development of injectable dosage forms. Various six-sigma tools, including design of experiments (DoE), risk assessment, critical to quality (CTQ), affinity diagram, quality function deployment (QFD), failure mode and effects analysis (FMEA), statistical analysis, process capability analysis, control strategy, etc. are useful at different stages of product development. DoE is one of the most widely used tools for formulation and analytical method development, process optimization, and process validation [4, 5].

DoE is a systematic statistical approach that allows the evaluation of the impact of change in multiple input variables, known as factors, within the boundary of experimental design, on the output variables, known as responses. DoE is a robust data collection (by designing and conducting experiments) and analytical tool (by analysis and inferring data). The stochastic models, developed based on specific factors combination and obtained results, are used for identifying the effects of factors on responses and help understand the nature of interactions between two or more factors. DoE is useful for obtaining the "true optimum" (design space) with minimum possible experiments leading to faster and cost-effective product and process development. Also, an essential advantage of using DoE is that the experimenters can quantify the interactions between factors which cannot be determined with traditional one factor at a time (OFAT) approach [6–9].

5.2 Basics of DoE

This section describes basic concepts used frequently in DoE, which are essential to understand before discussing various types of experimental designs, the rationale of design selection, and case studies.

5.2.1 Randomization

R. A. Fisher introduced the concept of randomization in experimental designs in 1925. Randomized experiments are considered as "gold standards" for inferring unambiguous and valid conclusions from statistical data [10]. Systematic (but not randomized) experiments lead to judgment bias and inaccurate interpretation of the

data. Also, non-randomized experiments are vulnerable to confounding or hidden variables, also known as lurking variables, which vary with time. Examples of lurking variables include a change in temperature of different shelves in lyophilizer during series of experiments, the machine heats up, change in experimenter, temperature or humidity changes, etc. Randomization does not mean that experiments to be performed in any order occur to the experimenter; it must be a physical experimental process [11]. Methods for randomization include simple randomization (flipping a coin, throwing dice, and randomly select a card from a shuffled deck), block randomization (grouping in equal sample size), and stratified randomization (randomization in a way that controls and balances the effect of covariates) [12].

Randomization serves the following purposes [13].

- No selective bias to the results of experiments.
- Accurate and unbiased estimation of error effects.
- Ensures that the error effects are statistically independent.

5.2.2 Blocking

Blocking is a mathematical technique of removing variations associated with a known change during the experiment. For example, if two different batches (or lots) of a surfactant is needed to prepare an emulsion product; the change in manufacturer batch (or lot) of surfactant might affect the properties of the emulsion. Performing the experiments with two different blocks (each block of experiments with one batch of surfactant) normalizes the effect caused by batches of surfactant. Blocking helps to reduce variability due to known reasons for experiments that may take several days, may involve different experimenters, and may subject to known changes in experimental conditions. However, the experimenter should be careful during block selection. Blocking should not be applied to a factor if the experimenter is interested in studying the effect of that factor on the response. For example, in the above emulsion experiments, blocking should not be used on surfactant if the experimenter is interested in evaluating the effect of different batches of surfactant on emulsion properties. Blocking can be applied to more than one factor during the experiments [6, 14].

Blocking serves the following purposes [14].

- Ensures that the blocked variable does not spoil the evaluation of other variables.
- Precise estimation of an experimental error (by removing the effect of a blocked variable from error calculation).
- In a few cases, it is possible to measure the effect of the blocked variable on the response.

5.2.3 Orthogonality

Independent variables (factors) affecting the dependent variable (response) are orthogonal if they are not correlated. For example, the concentration of the oil phase and concentration of surfactant are two orthogonal independent variables affecting the stability of the globule size distribution of an emulsion.

Orthogonality is an indicator of the independence of factors. In the DoE layout, each column is representative of a factor. It is important to estimate the effect of a factor (and interactions) independently without the interference of other factors. Orthogonality ensures independent estimation of the effect of a factor [15, 16].

5.2.4 Replication

Replication can be defined as the repetition of the same set of experimental conditions more than once. All similar experiments are known as replicates. The variability in the response for a similar set of experimental conditions indicates that the source must be something other than the factors controlled during the experiments. The objectives of replicate experiments are to determine the experimental error and reduce the bias due to uncontrolled variables. It increases the signal to noise (S/N) ratio if the noise is due to uncontrolled variables. The experimental error can also be determined if the process is in statistical control for a time. The standard error of mean (SEM) (standard deviation of the theoretical distribution of the sample means) can be expressed by $\sqrt{(s/n)}$, where s stands for the standard deviation (measure of dispersion of individual values) and n stands for the number of samples. It is desirable to have a higher number of samples and a lower standard deviation to achieve lower SEM. Replicates can increase the number of samples, whereas blocking helps decrease the standard deviation [17].

5.2.5 Confounding/Aliasing

Confounding or aliasing refers to the inability of clean estimation of effects and interactions. Effects that cannot be estimated independently of each other are considered confounded or aliased. It is the price experimenter pays with the fractional factorial designs because experiments for all combinations of factor levels are not possible with a reduced number of experiments, i.e., not enough degree of freedom. For example, if the estimate of effect X_4 in four-factor experiment is $(X_4 + X_1X_2X_3)$, then the main effect X_4 is aliased with 3-way interaction $X_1X_2X_3$. It cannot be concluded whether the significant effect, if any, is due to X_4 alone, interaction $X_1X_2X_3$, or both [18, 19].

Confounding is undesirable. However, it is not practically possible to perform experiments for all combinations of factor levels at industrial settings due to time and cost constraints. Confounding is a decision for an experimenter to make for knowingly confound (higher order) interactions with main effects while generating the

experimental design. The good news for pharmaceutical scientists is that the higher-order interactions (3 factor interactions (3FI) and above), generally, have been observed to be insignificant in most cases [18, 19].

5.2.6 Resolution

The resolution of an experimental design refers to the degree of confounding, i.e., the degree to which the main effects are confounded with 2 or 3 or higher factor interactions. The number of resolution of design indicates interactions confounded with main effects; for example, resolution III means the main effects are confounded with 2-factor interactions (2FI). Similarly, in resolution IV designs, main effects are confounded with 3-factor interactions (3FI), and 2FI are confounded with 2FI. Generally, 2FI has a significant effect on responses. It is advisable to choose designs with a higher resolution. Resolution V or higher designs are good for characterization, and resolution IV designs are adequate for screening purposes. The resolution III designs should only be used for ruggedness testing and comparisons. The resolution term is not applicable to full factorial designs as they do not have a confounding effect [7, 20].

5.2.7 Model

A model is a mathematical relationship such as equations and formula constructed using statistical methods that relate changes in one or more factors to the changes in responses. Based on the nature of collected data, different models might be helpful such as linear models, interaction models, quadratic models with curvature in one or more variables, cubic models, etc.

Caution: Before we move forward to the types and selection of experimental designs, it is essential to understand what DoE can give us. *The design of experiments is not the panacea*. Statistical modeling works best with a sound scientific approach. Understanding the scope of experiments, selection of appropriate factors and levels, and suitable DoE design is the key to successful DoE. Factors and levels (operating ranges within the experimental boundary) should be selected based on scientific rationale. A few pre-DoE experiments might help to decide the factors and ranges to use in DoE. Also, tools such as a cause-and-effect relationship (fishbone diagram) and risk analysis are useful for the selection of factors.

5.3 Types and Criteria for Selection of an Experimental Design

5.3.1 Types of Experimental Designs

DoE designs include factorial and fractional designs, Placket–Burman, Taguchi design, response surface methods, etc. Following are the types of DoE designs categorized according to the objective of the experiments [21].

5.3.1.1 Comparative Objective

If the primary objective is to identify whether a factor, out of several studied factors, is significant. Randomized block designs are useful for comparative purposes.

5.3.1.2 Screening Objective

If the primary objective is to screen out a few vital factors out of several investigated factors that affect the responses. The selected important factors can be studied further for optimization. Fractional factorial designs, Placket–Burman, Taguchi designs, etc. are useful for screening purposes.

5.3.1.3 Optimization Objective

The optimization objective includes identification and quantification of the main effects and higher-order interactions, resulting in a design space. Several product and process development experiments have the goal of optimization to ensure reproducible quality. Various designs are used for optimization purposes, such as response surface methods (RSM) designs, including central composite, Box–Behnken, and optimal designs. Mixture designs such as simplex lattice, simplex centroid, and optimal designs are useful if the factors are proportions of a mixture. Mixture designs are used to find out the optimum composition/fraction of factors to achieve desired responses.

5.3.2 Rationale for Selection of an Experimental Design

The selection of an experimental design depends on parameters such as the objective of the experiments, number of factors to be investigated, available resources such as feasibility of maximum number experiments, time, material, cost, etc. A higher number of experiments provide more information, but in most practical cases, it is not feasible to invest more resources and a long time. Faster development of quality drug products is the key.

Full factorial designs are useful if the number of factors to be studied is less than 5. In case of factorial designs, the number of experiments is determined by $2^K$where K = number of factors. Full factorial design for 3, 4, and 5 factors suggest 8, 16, and 32 experiments, respectively. The number of experiments increases exponentially as the number of factors increases, for example, 64 (6 factors), 128 (7 factors), 256 (8 factors), and so on, which is cumbersome and time-consuming. Screening designs are useful to select a few important factors out of many. Fractional factorial,

minimum run screening, Plackett–Burman, and Taguchi designs are useful for screening (number of factors 5 or more). Response surface designs are used for characterization and optimization. These designs can be applied using the selected important factors, generally but not limited to 2–4, from the screening studies [14].

5.4 Case Study—Screening DoE for a Sterile Nanoemulsion Product

5.4.1 Introduction

5.4.1.1 Product

ABC (1% w/v) nanoemulsion.

5.4.1.2 Objective

The objective of the screening design was to find out the most important factors affecting the responses, i.e., selection of a vital few factors from the trivial many. Selected factors were studied for optimization purposes.

5.4.1.3 Factors

Factors were selected based on domain knowledge and a risk assessment (Table 5.1).

5.4.1.4 Responses

Responses are the critical quality attributes of the drug product.

- Globule size distribution (Z-average and PDI).
- In vitro drug release (IVR) at 1 h, 6 h, and 12 h.

Table 5.1 Factors for screening DoE

Factors	Factor names	Units	Type	Subtype	Minimum	Maximum
A	Particle size of drug (D50)	μ	Numeric	Continuous	10	30
B	Viscosity of oil phase[a]	N s/m^2	Numeric	Continuous	0.550	0.750
C	Homogenization temperature	^{0}C	Numeric	Continuous	30	60
D	Homogenization pressure	Bar	Numeric	Continuous	200	1000
E	Homogenization time	Minutes	Numeric	Continuous	5	25
F	Preservative	Wt. %	Numeric	Continuous	0.1	2

[a]Oil phase refers to the mixture of mineral oil, drug, and surfactant

Table 5.2 Estimated and confounded terms of screening DoE design

Estimated term	Confounded terms
A	A + BCE + DEF
B	B + ACE + CDF
C	C + ABE + BDF
D	D + AEF + BCF
E	E + ABC + ADF
F	F + ADE + BCD
AB	AB + CE
AC	AC + BE
AD	AD + EF
AE	AE + BC + DF
AF	AF + DE
BD	BD + CF
BF	BF + CD
ABD	ABD + ACF + BEF + CDE
ABF	ABF + ACD + BDE + CEF

5.4.2 Experimental Design

5.4.2.1 Selection of Design

A fractional factorial design (2^{6-2}, resolution IV; software—Design-Expert®Version 12.0.9.0, Stat-Ease Inc.) was selected. Resolution IV design allows clean estimation of main effects. The 2FI confound with other 2FI, which might not be a concern for screening purpose. Table 5.2 shows the confounded terms of the designs, which show that the main effects are confounded with 3FI (insignificant in most cases). Minimum run screening design (a set of two-level designs) can be used if it is necessary to reduce the experimental runs. However, minimum run designs are extremely sensitive to missing data. Even one missing data reduce the resolution of design to III, which means the main effects will be confounded with 2FI [22].

The power of the design (the ability of the design to detect the significant effects) should be more than 80% for practical purposes. It is advisable to decide the S/N ratio based on desirable difference to detect (signal) and the variability in the measurements of responses (noise).

5.4.2.2 Design Layout

Table 5.3 shows the design layout with factors combination and responses obtained after the experiments. Experiments were performed in a randomized manner to avoid bias.

5.4.2.3 Design Summary

Two-level factorial design with reduced 3FI without center points or blocks was built. Tables 5.4 and 5.5 show the summary and descriptive statistics of factors and responses, respectively.

Table 5.3 Design layout with responses

		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Response 1	Response 2	Response 3	Response 4	Response 5
Std	Run	A:Particle size of drug (D50)	B: Viscosity of oil phase	C: Homogenization temperature	D: Homogenization pressure	E: Homogenization time	F: Preservative	Z-average	PDI	IVR 1 h	IVR 6 h	IVR 12 h
		Micron	N s/m^2	Degree C	Bar	Minutes	Wt. %	Nm		%	%	%
2	1	30	0.55	30	200	25	0.1	610	0.17	22.1	57.8	82.3
1	2	10	0.55	30	200	5	0.1	829	0.27	14.4	50.1	78.8
8	3	30	0.75	60	200	25	0.1	550	0.12	24.2	58.8	87.7
10	4	30	0.55	30	1000	25	2	715	0.2	18.6	54.6	79.7
7	5	10	0.75	60	200	5	0.1	598	0.13	23.1	59.9	86.5
13	6	10	0.55	60	1000	25	0.1	239	0.07	29.5	65.4	93.7
12	7	30	0.75	30	1000	5	0.1	777	0.25	17.2	53.1	79
6	8	30	0.55	60	200	5	2	576	0.16	24.1	51.8	76.2
15	9	10	0.75	60	1000	5	2	269	0.09	27	62.8	90
4	10	30	0.75	30	200	5	2	832	0.26	15.5	48.8	72.8
5	11	10	0.55	60	200	25	2	544	0.15	25.9	57.7	82.2
9	12	10	0.55	30	1000	5	2	781	0.22	16.2	46.6	68.5
11	13	10	0.75	30	1000	25	0.1	722	0.23	17.9	45.9	70.1
16	14	30	0.75	60	1000	25	2	245	0.1	29.2	66.2	92.4
14	15	30	0.55	60	1000	5	0.1	275	0.12	30.1	63.4	94.5
3	16	10	0.75	30	200	25	2	617	0.18	22.8	44.8	70.9

Table 5.4 Summary of factors with associated descriptive statistics

Factor	Name	Units	Type	Minimum	Maximum	Coded Low	Coded High	Mean	Std. Dev.
A	Particle size of drug (D50)	Micron	Numeric	10.00	30.00	−1 ↔ 10.00	+1 ↔ 30.00	20.00	10.33
B	Viscosity of oil phase	N s/m2	Numeric	0.5500	0.7500	−1 ↔ 0.55	+1 ↔ 0.75	0.6500	0.1033
C	Homogenization temperature	Degree C	Numeric	30.00	60.00	−1 ↔ 30.00	+1 ↔ 60.00	45.00	15.49
D	Homogenization pressure	Bar	Numeric	200.00	1000.00	−1 ↔ 200.00	+1 ↔ 1000.00	600.00	413.12
E	Homogenization time	Minutes	Numeric	5.00	25.00	−1 ↔ 5.00	+1 ↔ 25.00	15.00	10.33
F	Preservative	Wt. %	Numeric	0.1000	2.00	−1 ↔ 0.10	+1 ↔ 2.00	1.05	0.9812

Table 5.5 Summary of responses with associated descriptive statistics

Response	Name	Units	Observations	Analysis	Minimum	Maximum	Mean	Std. Dev.	Ratio	Transform	Model
R1	Z-average	Nm	16.00	Factorial	239	832	573.69	210.64	3.48	None	Mean
R2	PDI		16.00	Factorial	0.07	0.27	0.1700	0.0632	3.86	None	Mean
R3	IVR 1 h	%	16.00	Factorial	14.4	30.1	22.36	5.22	2.09	None	Mean
R4	IVR 6 h	%	16.00	Factorial	44.8	66.2	55.48	7.07	1.48	None	Mean
R5	IVR 12 h	%	16.00	Factorial	68.5	94.5	81.58	8.58	1.38	None	Mean

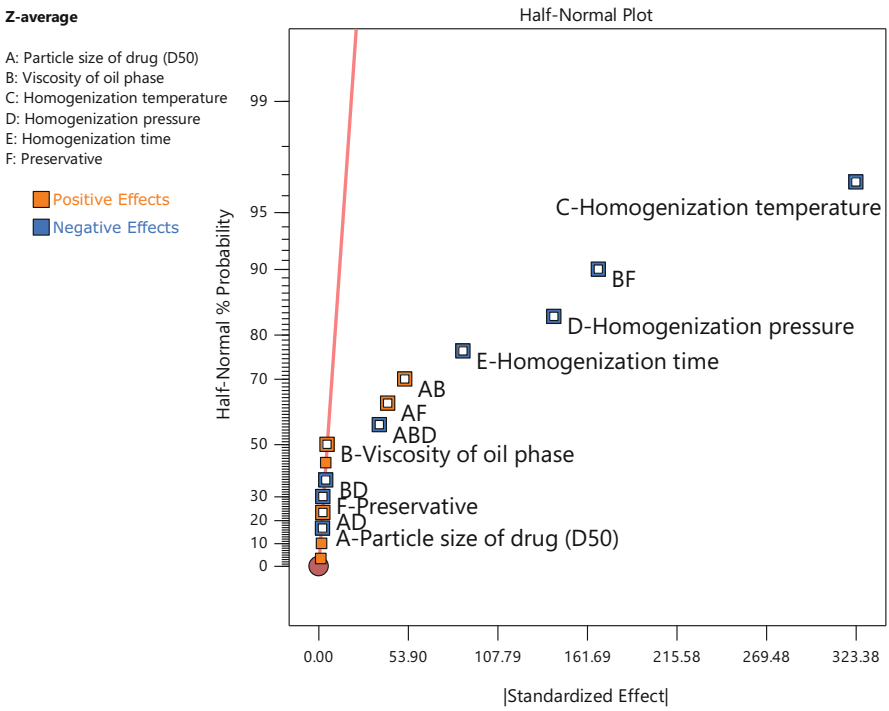

Fig. 5.1 Half-normal plot for respone 1(Z-average)

5.4.3 Effects Analysis

5.4.3.1 Response 1: Z-average

Half-normal plot: It assesses the relative significance of factors or interaction terms. It is a scale to determine the impact of factors or interaction terms on response. The terms that have more significant estimated effects appear away from the line in the right corner [23, 24]. Half-normal plot for a response (Fig. 5.1) shows that the main effects homogenization temperature (factor C), homogenization pressure (factor D), and homogenization time (factor E) have the larger effects. Also, interactions terms BF, AB, AF, and ABD are significant as they are away from the line. We can identify and focus on three main effects (C, D, and E) for further evaluation.

Normal plot: Normal probability plot of estimated effects is another tool to assess the relative impact/significance of factors or terms on response. The terms on the line have minimal effect, whereas terms on either side of the line represent the higher impact (greater the distance from the line, higher the impact) [25]. Figure 5.2 shows similar results as of the half-normal plot. Half-normal plot is another way of representing the normal plot with only positive values (conversion of estimated effects in absolute numbers).

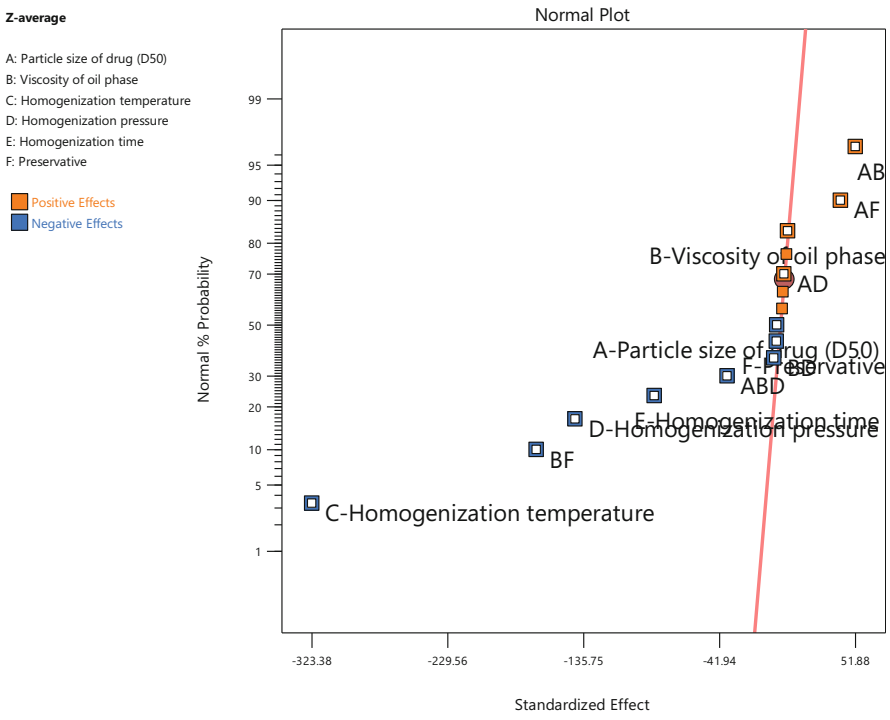

Fig. 5.2 Normal plot for response 1(Z-average)

Pareto chart: Pareto chart is a graphical way to present the selected model terms and their significance [26]. It contains two different t-limits (Bonferroni and standard t-limit). The values for both limits change based on the selected model terms (Fig. 5.3).

5.4.3.2 Response 2: PDI

Half-normal and normal plots: The half-normal and normal plots for the response PDI show that homogenization temperature (factor C) has a significant effect on PDI as compared to any other factor (Figs. 5.4 and 5.5).

Pareto chart: Fig. 5.6 shows that only homogenization temperature (factor C) exceeds the Bonferroni limit, which is more conservative than the standard t-critical. No other factor was found significant.

5.4.3.3 Response 3: IVR 1 h

Based on the effect's analysis by half-normal plot, normal plot, and pareto chart (data not shown), main effects homogenization temperature (factor C) and homogenization time (factor E) and the interaction term BF observed to be significant. Homogenization temperature (factor C) crossed the Bonferroni limit, whereas homogenization time (factor E) and BF crossed t-limit.

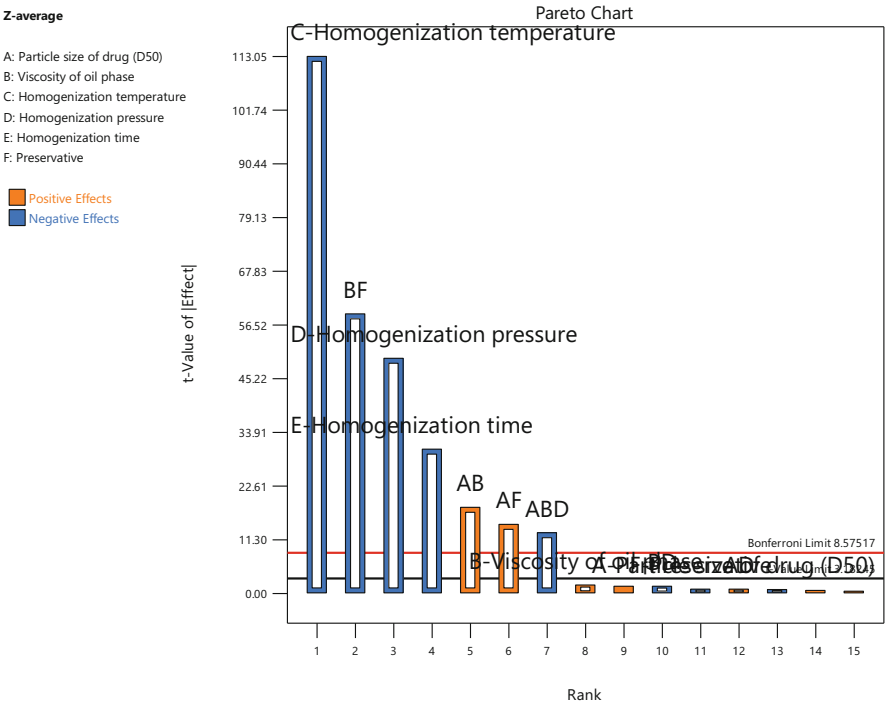

Fig. 5.3 Pareto chart for response 1 (Z-average)

5.4.3.4 Response 4: IVR 6 h

Based on the effect's analysis by half-normal plot, normal plot, and pareto chart (data not shown), only homogenization temperature (factor C) significantly affected (crossed the Bonferroni limit) the IVR at 6 h.

5.4.3.5 Response 5: IVR 12 h

Only homogenization temperature (factor C), like response 4, significantly affected (crossed the Bonferroni limit) the IVR at 12 h.

5.4.4 Analysis of Variance (ANOVA)

5.4.4.1 Response 1: Z-average

Table 5.6 shows the ANOVA for the selected model for the response Z-average. Based on such a high F-value (1694.27) and low p-value (<0.0001), it can be inferred that the selected model was significant. Model terms C, D, E, AB, AF, BF, and ABD were significant. R^2 of the model was 0.9999. Also, a good agreement between adjusted R^2 (0.9993) and predicted R^2 (0.9958) was observed (difference less than 0.2 between adjusted and predicted R^2).

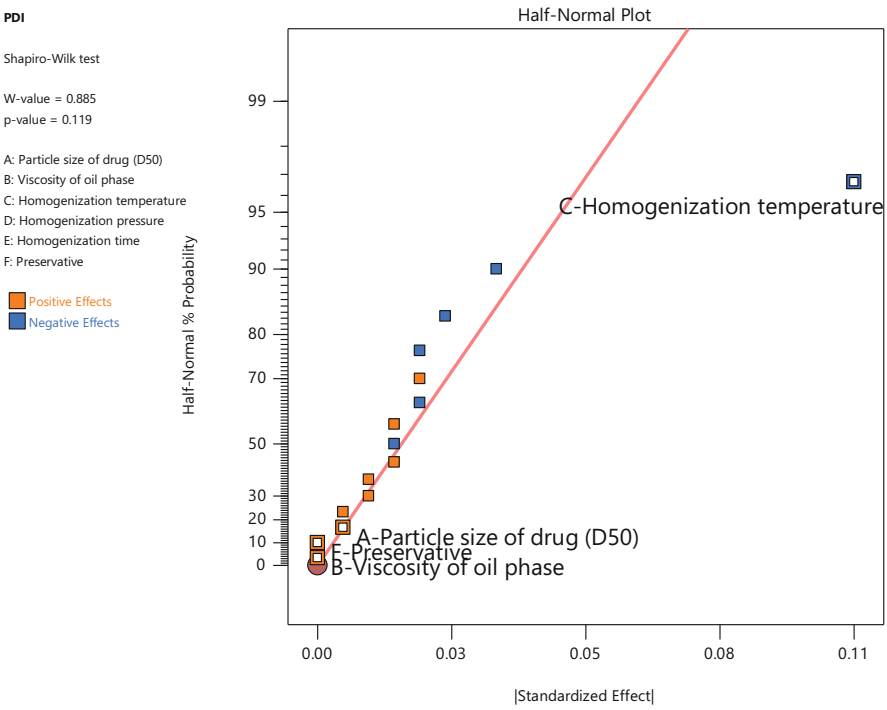

Fig. 5.4 Half-noraml plot for response 2 (PDI)

5.4.4.2 Response 2: PDI

ANOVA for response PDI shows that homogenization temperature (factor C) is significant. High F-value (7.69) and low p-value (0.0033) indicate that the model is significant (Table 5.7). The R^2 of the model was 0.7367. Also, a good agreement between adjusted R^2 (0.6409) and predicted R^2 (0.4429) was observed.

5.4.4.3 Response 3: IVR 1 h

Table 5.8 shows the ANOVA for the selected model for the response IVR 1 h. Based on F-value (14.86) and p-value (0.0002), the selected model was significant. Model terms C, E, and BF were significant. The R^2 of the model was 0.8814. Also, a good agreement between adjusted R^2 (0.8221) and predicted R^2 (0.6964) was observed.

5.4.4.4 Response 4: IVR 6 h

Homogenization temperature (factor C) was a significant term in the selected model (Table 5.9). The R^2 of the model was 0.5932. A good agreement between adjusted R^2 (0.5641) and predicted R^2 (0.4687) was observed.

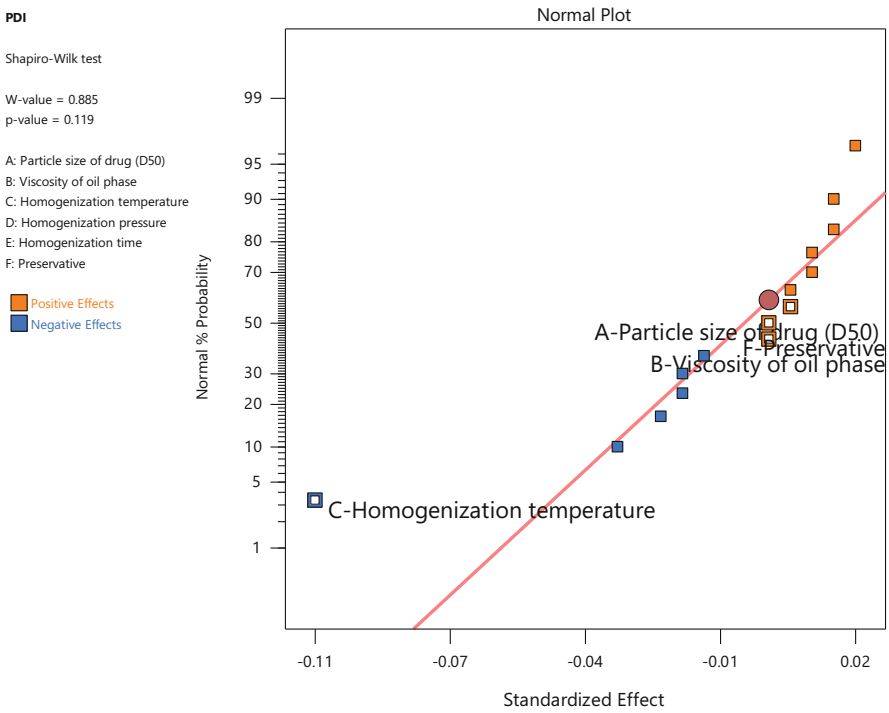

Fig. 5.5 Normal plot for response 2 (PDI)

5.4.4.5 Response 5: IVR 12 h

Like response 4, homogenization temperature (factor C) was a significant term in the selected model (Table 5.10). The R^2 of the model was 0.5780. A good agreement between adjusted R^2 (0.5479) and predicted R^2 (0.4488) was observed.

5.4.5 Diagnostics

Diagnostics play a vital role in verification, whether the selected regression model fits the data suitably and meet various assumptions. Various residual diagnostics and influence diagnostics are frequently used to test the appropriateness of the model. We will discuss the residual and influence diagnostics of the model selected for response 1 (Z-average) to avoid repetition. However, the diagnostic analysis should be performed for the models developed for all responses for practical purposes.

5.4.5.1 Residual Diagnostics

Analysis of residuals is an essential verification before concluding from the regression analysis. Diagnostics plots and residual analysis are used to detect problems associated with model analysis. The selected model, primarily linear, is reasonable if

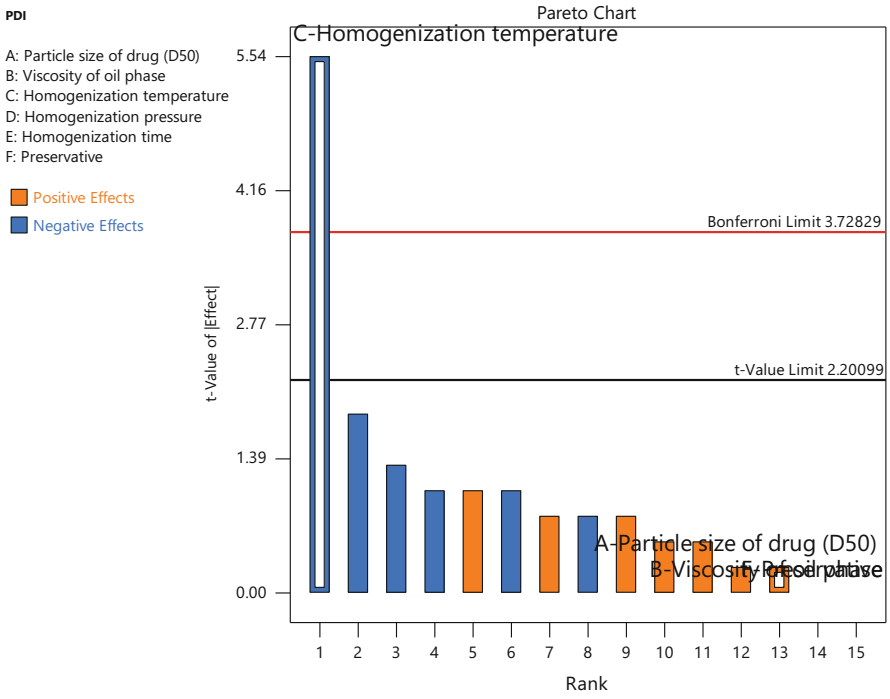

Fig. 5.6 Pareto chart for response 2 (PDI)

residuals have a normal distribution, constant variance, and independent of each other over time.

Normal Probability of Residuals

Figure 5.7 shows the normal distribution of the externally studentized residuals indicating that the selected model makes sense. Externally studentized residuals are used because of the higher sensitivity for the detection of problems. Moreover, each raw residual belongs to different populations and makes the interpretation difficult in both conditions (constant or variable variance). Studentized residuals calculation involves the deletion of an observation at a time and re-fitting the regression model on the remaining (n-1) observations followed by a comparison of observed and fitted values on the new model. Studentized residuals consider the standard deviation estimate and are thus more effective in detecting outliers. Normally distributed residuals follow a straight line. Any patterns in the normal probability plot of residuals suggest the superiority of alternative analysis, such as the transformation of the responses [27–30] and [35].

Table 5.6 ANOVA for response 1 (Z-average)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	6.654E+05	12	55452.10	1694.27	< 0.0001	significant
A-Particle size of drug (D50)	22.56	1	22.56	0.6894	0.4673	
B-Viscosity of oil phase	105.06	1	105.06	3.21	0.1711	
C-Homogenization temperature	4.183E+05	1	4.183E+05	12780.21	< 0.0001	
D-Homogenization pressure	80230.56	1	80230.56	2451.35	< 0.0001	
E-Homogenization time	30189.06	1	30189.06	922.39	< 0.0001	
F-Preservative	27.56	1	27.56	0.8421	0.4265	
AB	10764.06	1	10764.06	328.88	0.0004	
AD	27.56	1	27.56	0.8421	0.4265	
AF	6930.56	1	6930.56	211.75	0.0007	
BD	76.56	1	76.56	2.34	0.2236	
BF	1.134E+05	1	1.134E+05	3464.82	< 0.0001	
ABD	5365.56	1	5365.56	163.94	0.0010	
Residual	98.19	3	32.73			
Cor Total	6.655E+05	15				

Table 5.7 ANOVA for response 2 (PDI)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	0.0442	4	0.0110	7.69	0.0033	significant
A-Particle size of drug (D50)	0.0001	1	0.0001	0.0696	0.7968	
B-Viscosity of oil phase	1.388E-17	1	1.388E-17	9.662E-15	1.0000	
C-Homogenization temperature	0.0441	1	0.0441	30.70	0.0002	
F-Preservative	0.0000	1	0.0000	0.0000	1.0000	
Residual	0.0158	11	0.0014			
Cor Total	0.0600	15				

Residual Vs. Predicted

Residual vs. predicted plot is used to test the assumption of the constant variance of residuals. The random scatter of the residuals indicates the constant range of residuals, whereas patterns such as megaphone suggest a transformation of the data. Figure 5.8 shows a random scatter pattern of the residuals, indicating a constant variance [31].

Table 5.8 ANOVA for response 3 (IVR 1 h)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	359.60	5	71.92	14.86	0.0002	significant
B-Viscosity of oil phase	1.0000	1	1.0000	0.2067	0.6591	
C-Homogenization temperature	292.41	1	292.41	60.44	< 0.0001	
E-Homogenization time	31.92	1	31.92	6.60	0.0280	
F-Preservative	0.0400	1	0.0400	0.0083	0.9293	
BF	34.22	1	34.22	7.07	0.0239	
Residual	48.38	10	4.84			
Cor Total	407.98	15				

Table 5.9 ANOVA for response 4 (IVR 6 h)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	444.16	1	444.16	20.41	0.0005	significant
C-Homogenization temperature	444.16	1	444.16	20.41	0.0005	
Residual	304.59	14	21.76			
Cor Total	748.74	15				

Table 5.10 ANOVA for response 5 (IVR 12 h)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	638.83	1	638.83	19.17	0.0006	significant
C-Homogenization temperature	638.83	1	638.83	19.17	0.0006	
Residual	466.42	14	33.32			
Cor Total	1105.24	15				

Residual Vs. Run

Residual vs. run is a plot of residuals with run order of experiments. It checks for the effect of lurking variables on the outcomes. The lurking variable is an extraneous variable that can have a positive or negative correlation with both the dependent variable and the independent variable. A specific trend in the plot indicates the existence of a lurking variable over time. Figure 5.9 shows a random scatter that means no interference of the lurking variable [31].

Predicted Vs. Actual

The plot of predicted vs. actual responses tests the ability of the model to predict the responses accurately. A good correlation is an indication of the ability of a model to predict close to the actual values (Fig. 5.10).

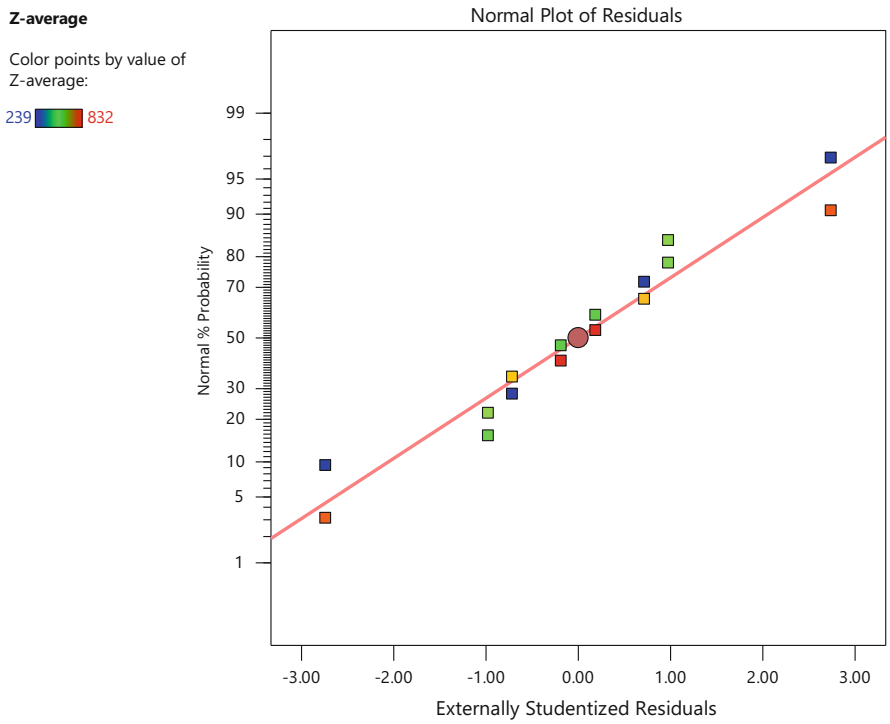

Fig. 5.7 Diagonastic plot and residual analysis: normal plot of residuals

Box–Cox Plot

A power transformation helps in reducing the anomalies such as non-normality and heteroscedasticity. Box–Cox transformation is a technique used for power transformations. Generally, statistical analysis and inference follow the assumption that data are normally distributed, have a common variance, and error structure is additive. However, if these assumptions are seriously violated, one may perform a power transformation and rebuild the model that has all essential aspects of the original model. Also, the new model satisfies all the assumptions. Box–Cox plot is a curve of the natural log of the sum of squares of residuals in which the minimum value indicates the lambda value. Power transformation is suggested based on lambda value. Lambda value of 1 (or any value for which 95% CI includes 1) does not require any transformation (Fig. 5.11). Other values of lambda such as 0.5 (square root), 0 (natural log), −0.5 (inverse square root), −1 (inverse), etc. suggest transformations [31, 32].

5.4.5.2 Influence Diagnostics

The identification of influential points is a critical aspect of regression diagnostics. It is essential to identify runs (observations) that have a high influence on the model

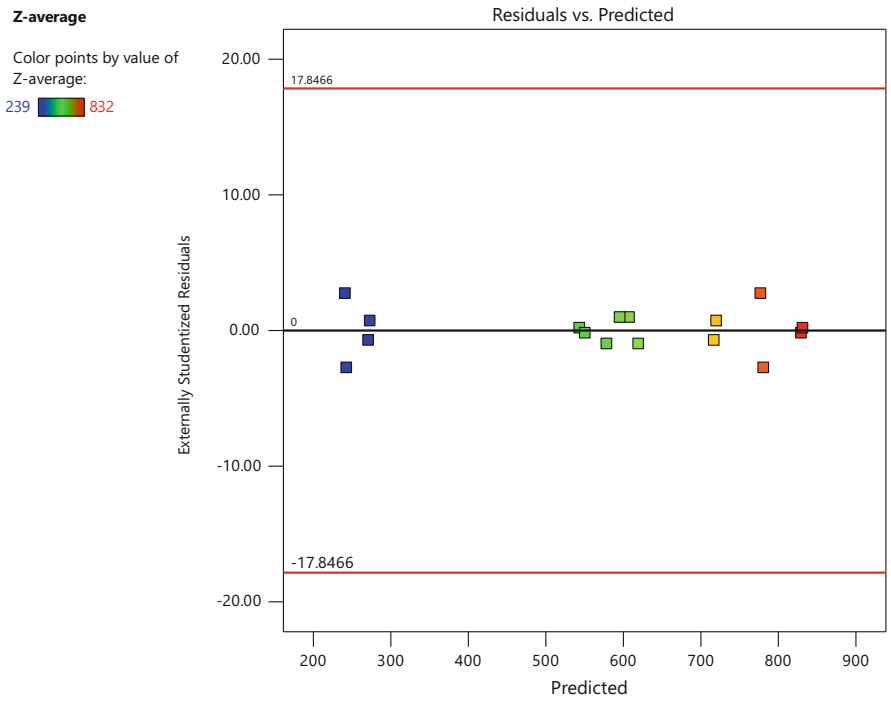

Fig. 5.8 Diagonastic plot and residual analysis: residual vs predicted

and the responses. Influence plots such as Cook's distance, leverage, DFFITS, and DFBETAS, provide a graphical measure of the influence of individual runs.

Cook's Distance or Cook's D (Di)

Cook's distance (Di) is used to identify the most influential run (or outlier) in regression analysis. A higher value of Di indicates a strong influence or a potential outlier. Generally, Di value more than one should be investigated, and value more than three represents an outlier. Di calculation includes rebuilding the regression model after removing i_{th} data point from the existing model and check for differences in predictions. Cook's distance within the limit indicates that no run is highly influential (Fig. 5.12) [33].

Leverage

Leverage is used to identify influential points and outliers by the distance of an observation point from the average predictor values. An observation point having a leverage of more than twice than the average is generally considered as high leverage. A high leverage point potentially has an impact on model fit; however, it does not necessarily mean that the point has a strong influence on the regression coefficient estimates. A higher distance from the predictor average for a point as compared to the other points can be situated in the same regression line. Therefore,

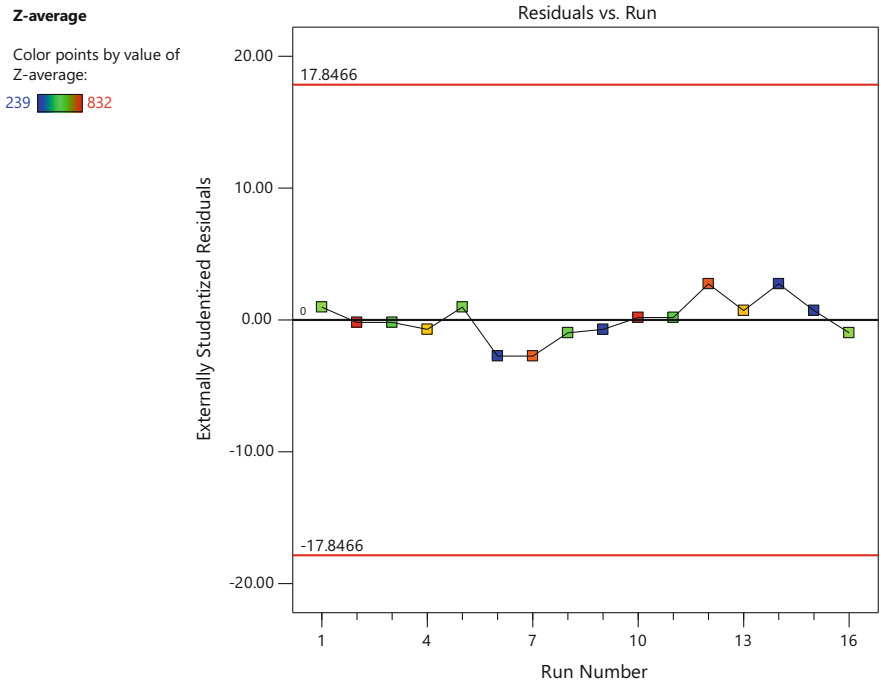

Fig. 5.9 Diagonastic plot and residual analysis: residual vs run

the evaluation of the discrepancy of the observation from other data might be helpful in addition to leverage value [34]. Figure 5.13 shows no high leverage run.

DFFITS, MDFFITS, and DFBETAS

In addition to the Cook's distance, several case-deletion diagnostics such as DFFITS, MDFFITS, and DFBETAS, are used in regression modeling. DFFITS measures the change in prediction value after removing the i_{th} point (i_{th} point not included in the model). MDFFITS is used when multiple points are removed. DFBETAS measures the change in coefficient estimate after removing the i_{th} point (i_{th} point not included in the model) [34].

Based on various residual and influence diagnostics, we can infer that the selected model is appropriate as residuals showed normal distribution, a constant variance, no transformation is required, and no high influence runs observed. Similar outcomes were observed for all the remaining models created for the other responses.

5.4.6 Summary of the Screening Design

The objective of the screening design of experiment was to identify critical factors affecting the critical quality attributes (CQA's) (responses) of the nanoemulsion. Six

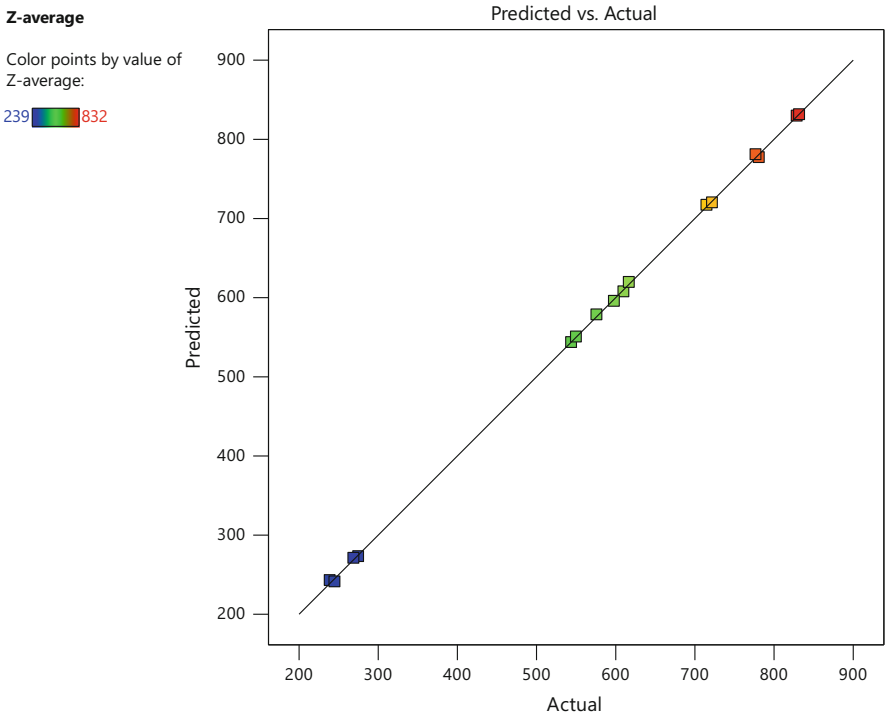

Fig. 5.10 Diagnostic plot and residual analysis: predicted vs actual

factors, namely, particle size of the drug, viscosity of oil phase, homogenization temperature, homogenization pressure, homogenization time, and preservative content were selected based on initial risk assessment. A fractional factorial design (2^{6-2}) of resolution IV was selected. Resolution IV design allows clean estimation of main effects. The two factors interactions confound with the other two factors interactions, but it might not be a concern for screening purpose. The factors can be prioritized for the optimization experiments based on the relative impact of the factors on responses. Globule size distribution (Z-average and PDI) and In vitro drug release (IVR) at 1 h, 6 h, and 12 h were selected as responses.

Based on the effects analysis and ANOVA for all responses, homogenization temperature (factor C) was found to be the most significant factor followed by homogenization time (factor E) and homogenization pressure (factor D). Also, a few interaction terms were significant. However, interaction terms were not given a due focus because the goal of screening design was to identify main factors affecting the CQA's of the nanoemulsion. Also, various residual and influence diagnostics showed that a reasonable regression model was built for all responses. Factors C, D, and E were selected for optimization studies. Optimization designs help define the design space (sweet spot), i.e., ranges of all factors that are suitable for achieving the desired qualities of the product.

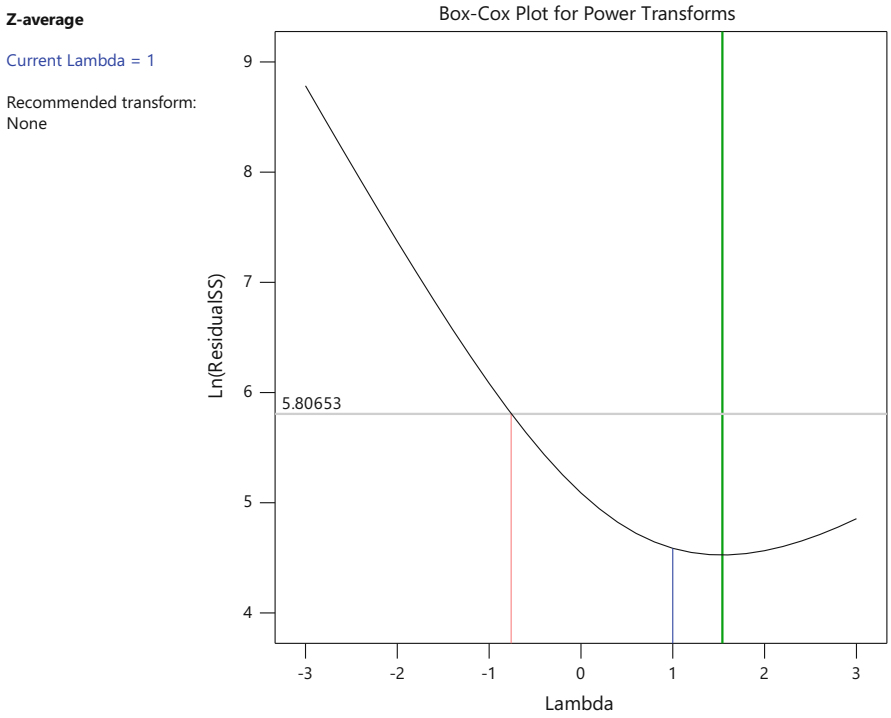

Fig. 5.11 Box-cox plot for power transforms

5.4.7 Connecting the Dots: The Science of Emulsion and Statistical Modeling

As we discussed earlier in the chapter, statistical modeling works best if complemented with sound scientific knowledge. Nanoemulsion can be defined as a kinetically stable system composed of oil and water, macroscopically homogeneous but heterogeneous at the microscopic scale. Based on type, oil (or water) droplets are distributed in the water (or oil) phase in o/w (or w/o) emulsion. The droplets of nanoemulsion, generally, have a submicron particle size (<1000 nm). The reduction and stabilization of particle size in the submicron size range require energy. The immiscibility of two phases is a result of high interfacial tension. Surfactants are used to reduce the interfacial tension and improve the stability of the emulsion. A negative change in Gibbs free energy upon mixing of two immiscible phases results in a stable emulsion. Gibbs free energy (ΔG) depends on enthalpy (ΔH), entropy (ΔS), temperature (T), interfacial tension (γ), and surface area (ΔA), given by the equation $\Delta G = \Delta H - T\Delta S + \gamma\Delta A$. Enthalpy change during the oil and water mixing is negligible. The entropy of mixing increases significantly with a decrease in particle size, i.e., $T\Delta S$ term dominates [35]. Higher temperature and

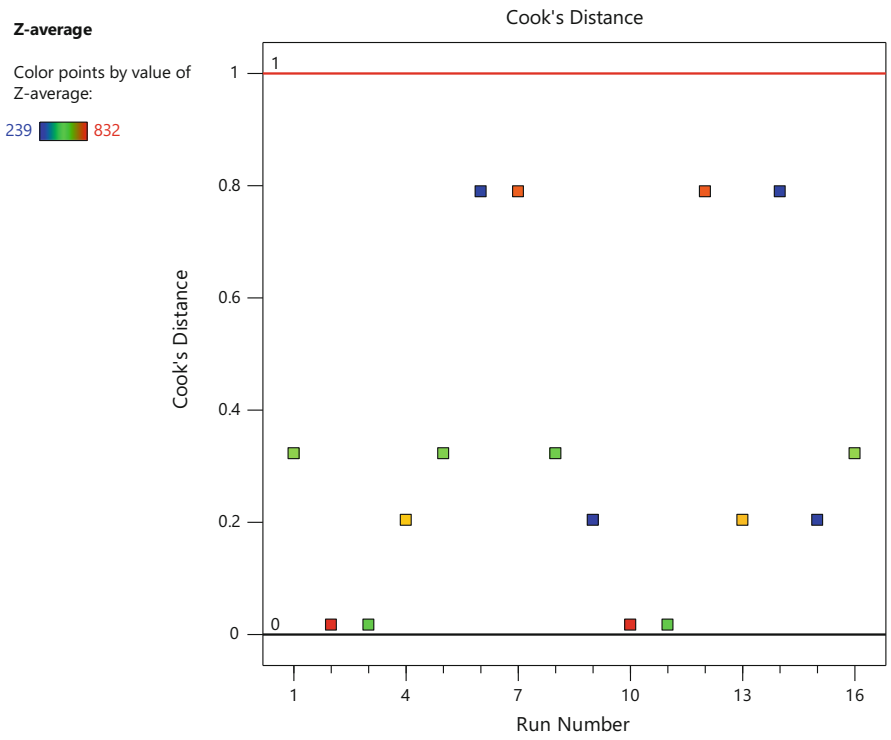

Fig. 5.12 Influence plots: Cook's distance

pressure during the homogenization process increase the energy input in the system and reduce the emulsion particle size. Moreover, the viscosity of the oil phase and surfactant decreases with an increase in temperature resulting in an efficient coating of surfactant on oil globules, making a stable emulsion. The screening design finds that parameters of the homogenization process significantly affected the quality of the nanoemulsion, which resonates with the theory of emulsion formation and stabilization.

5.5 Conclusion

DoE is an effective statistical tool to design the experiments in such a way that it ensures collecting the maximum information while minimizing the number of experiments. It helps in the analysis of the collected data and draws logical conclusions. DoE is used for screening and optimization purposes in various focus groups during injectable drug product development, including formulation, analytical, process scale-up, etc. DoE can detect and quantify interactions between factors. The traditional OFAT approach cannot detect interactions even with unlimited experiments. Interestingly, it has been observed many times during pharmaceutical

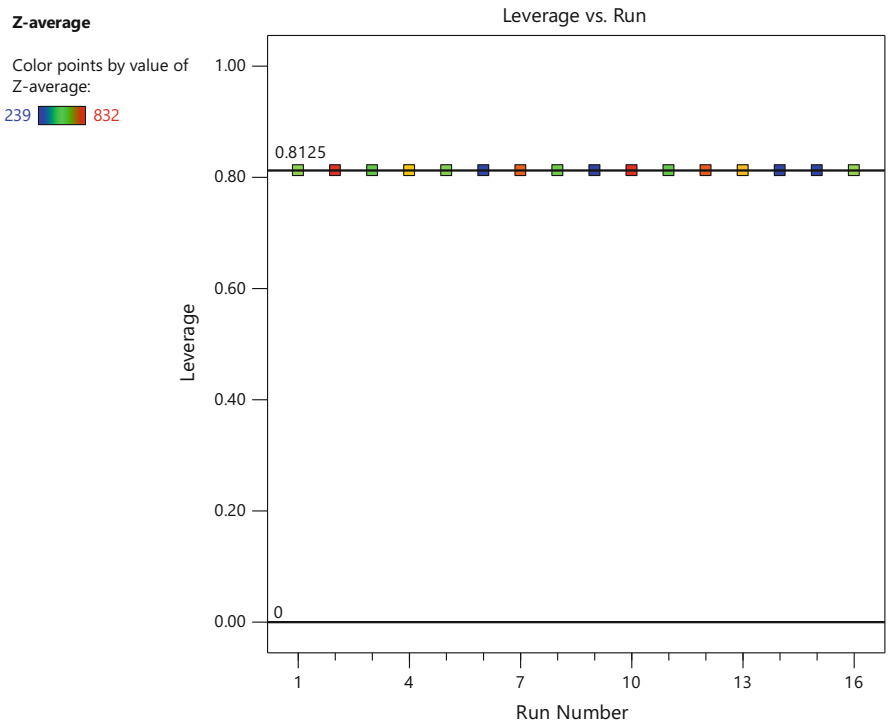

Fig. 5.13 Influence plots: Leverage vs Run

drug product development that interactions play a crucial role, sometimes exerting a more significant effect on responses than main effects. This chapter describes a few basic terminology and concepts which are used frequently in DoE. In addition, an outline of types of designs and criteria for selection of design were discussed. A case study of fractional factorial screening design for the development of a nanoemulsion product was discussed. Three out of six studied factors were found to be significant and considered for optimization studies.

Acknowledgment I want to thank Prof. Yoon Yeo for her support and suggestions in developing this book chapter.

References

1. Docarmo AM, Cunha-Filho MSS, Gelfuso GM, Gratieri T (2017) Evolution of quality on pharmaceutical design: regulatory requirement? Accred Qual Assur 22:199–205
2. Woodcock J, Woosley R (2008) The FDA critical path initiative and its influence on new drug development. Annu Rev Med 59:1–12
3. FDA (2020) Quality systems approach to pharmaceutical cGMP regulations

4. BEG S, Hasnain MS (2019) Pharmaceutical quality by design: principles and applications. Academic Press, Cambridge, MA
5. Chobisa D (2019) QbD-based development of pharmaceutical parenteral drug products: an overview. In: Pharmaceutical quality by design. Elsevier, Amsterdam
6. Anderson MJ, Whitcomb PJ (2016) DOE simplified: practical tools for effective experimentation. CRC Press, Boca Raton, FL
7. Fukuda IM, Pinto CFF, Moreira CDS, Saviano AM, Lourenço FR (2018) Design of experiments (DoE) applied to pharmaceutical and analytical quality by design (QbD). Brazil J Pharm Sci 54:e01006
8. Stavros NP, Colombo P, Colombo G, Rekkas DM (2017) Design of experiments (DoE) in pharmaceutical development. Drug Dev Ind Pharm 43:889–901
9. Shivhare M, McCreath G (2010) Practical considerations for DoE implementation in quality by design. J BioProcess Int 8:22–30
10. Alferes VR (2012) Methods of randomization in experimental design. Sage Publications, Thousand Oaks, CA
11. Hall NS (2007) RA fisher and his advocacy of randomization. J Hist Biol 40:295–325
12. Suresh K (2011) An overview of randomization techniques: an unbiased assessment of outcome in clinical research. J Hum Rep Sci 4:8
13. Kirk RE (2012) Experimental design. Wiley, Hoboken, NJ
14. Anderson MJ, Whitcomb PJ (2000) Design of experiments. In: Encyclopedia of chemical technology. Wiley, Hoboken, NJ, pp 1–22
15. Seberry J (2017) Orthogonal designs. Springer, Cham
16. Wang Y, Liu H, Cai Z, Zhou Y (2007) An orthogonal design based constrained evolutionary optimization algorithm. J Eng Optim 39:715–736
17. Pyzdek T, Keller PA (2003) Quality engineering handbook. CRC Press
18. Grömping U (2018) R package DoE. Base for factorial experiments. J Stat Softw 85:1–41
19. Minitab (2020) What are confounding and alias structure? [Online]. Minitab: Minitab. https://support.minitab.com/en-us/minitab/18/help-and-how-to/modelingstatistics/doe/supporting-topics/basics/what-are-confounding-and-alias-structure/. Accessed 7 Jan 2020
20. NIST Sematech (2020) A glossary of DOE terminology. https://www.itl.nist.gov/div898/handbook/pri/section7/pri7.htm#:~:text=Alias%3A%20When%20the%20estimate%20of,be%20aliased%20(see%20confounding)
21. Design of Experiments (2020) APB Consultant. https://isoconsultantpune.com/design-of-experiments/. Accessed 7 Aug 2020
22. Comparison of Factorial Designs (2020) Design-Expert, Stat-Ease, Inc: StatEase. Accessed 7 July 2020
23. Chung JH, Lim YB (2019) A modified LGB method for detecting significant effects based on a half-normal probability plot. J Korean Stat Soc 48:568–577
24. Moral RA, Hinde J, Demétrio CGB (2017) Half-Normal plots and Overdispersed models in R: the hnp package. J Stat Softw 81
25. Hazir E, Erdinler ES, Koc KH (2017) Optimization of CNC cutting parameters using design of experiment (DOE) and desirability function. J For Res 29:1423–1434
26. Harvey HB, Sotardi ST (2018) The pareto principle. J Am Coll Radiol 15:931
27. Ranganai E (2016) On studentized residuals in the quantile regression framework. Springerplus 5:1231
28. Kozak M, Piepho HP (2018) What's normal anyway? Residual plots are more telling than significance tests when checking ANOVA assumptions. J Agron Crop Sci 204:86–98
29. Vining G (2010) Technical advice: residual plots to check assumptions. Qual Eng 23:105–110
30. Studentized Residuals (2020) PennState Eberly college of science. https://online.stat.psu.edu/stat462/node/247/. Accessed 7 June 2020
31. Diagnostics Plots (2020) Design-Expert, Stat-Ease, Inc. Accessed 7 June 2020
32. Sakia RM (1992) The box-cox transformation technique: a review. J R Stat Soc: Ser D 41:169–178

33. Glen S 2016. *Cook's Distance / Cook's D: Definition, Interpretation* [Online]. https://www.statisticshowto.com/. Available: https://www.statisticshowto.com/cooks-distance/ [Accessed 7/6/2020 2020]
34. Liu X (2016) Residual and influence diagnostics. In: Methods and applications of longitudinal data analysis. Elsevier, Amsterdam
35. Moulik SP, Rakshit AK (2006) Physicochemisty and applications of microemulsions. J Surf Sci Technol 22:159