

Chapter 3

Process Design: Stage 1 of the FDA Process Validation Guidance

Keywords Critical process parameter (CPP) • Critical quality attribute (CQA) • Experimental design • Factorial design • Knowledge space • Normal operating range • Process capability • Process design • Process robustness • Quality by design • Quality target product profile (QTPP) • Region of goodness • Risk assessment

3.1 Introduction

This is the first of three chapters that describe statistical approaches related to the three stages of process validation described in the FDA Process Validation Guidance for Industry (2011). The three stages are

1. Process Design (Chap. 3),
2. Process Qualification (Chap. 4), and
3. Continued Process Verification (Chap. 5).

The three-stage process validation guidance aligns process validation activities to the product life cycle concept. Along with existing FDA guidance, it links the quality of the product with ensuring quality of the process, from product and process design through mature manufacturing. The FDA process validation guidance supports process improvement and innovation through sound science and includes concepts from other FDA supported guidance, including the International Conference on Harmonization (ICH) chapters Q8(R2) Pharmaceutical Development (2009), Q9 Quality Risk Management (2005), and Q10 Pharmaceutical Quality System (2008).

A goal of quality assurance is to produce a product that is fit for its intended use. Very broadly, within the guidance, process validation is defined as the collection and evaluation of knowledge, from the process design stage through commercial production, which establishes scientific evidence that a process consistently delivers quality product. This knowledge and understanding is the basis for establishing an approach to control a manufacturing process that results in products with the desired quality attributes. Across the three stages, the statistics contribution is to iteratively work on understanding sources of variability and the impact of

variability on the process and product attributes, build quality into the product and process, and detect and control the variation in a manner commensurate with the product risk and the patient needs.

This three-step approach assumes the following conditions:

1. Quality, safety, and efficacy are designed or built into the product.
2. Quality cannot be adequately assured merely by in-process and finished product inspection or testing.
3. Each step of a manufacturing process is controlled to ensure that the finished product meets all quality attribute specifications.

In slightly more detail, the three stages are

Stage 1—Process Design: The commercial manufacturing process is defined during this stage based on knowledge gained through development and scale-up activities. The knowledge can be of several forms: fundamental science, mechanistic or physics-based models, data-driven models based on previous compounds, and experimental understanding of the product being developed. In the process design stage various tools are employed to understand inputs to the process (parameters and material attributes) and their effect on the outputs (quality attributes). Throughout this development stage, decisions are made on how to establish and control the process to ensure quality in the output. This design stage can be in accordance with ICH Q8(R2) and ICH Q11 (2012) and as such may be a key change in the focus of activity for many companies.

Stage 2—Process Qualification: Following a process design stage where sufficient understanding has been gained to provide a high degree of assurance in the manufacturing process, the process design is evaluated to determine if the process is capable of reproducible commercial manufacturing. This stage has two elements: (1) design of the facility and qualification of the equipment and utilities and (2) process performance qualification (PPQ). This later element was historically called process validation, and most often conducted by executing three lots within predetermined limits.

Stage 3—Continued Process Verification: After establishing and confirming the process, manufacturers should maintain the process in a state of control over the life of the process, even as materials, equipment, production environment, personnel, and manufacturing procedures change. The goal of the third validation stage is continued assurance that the process remains in a state of control (the validated state) during commercial manufacturing. Systems for detecting departures from the product quality are helpful to accomplish this goal. Data are collected and analyzed to demonstrate that the production process remains in a state of control and to identify any opportunities for improvement.

Stages 2 and 3 are discussed in Chaps. 4 and 5, respectively.

3.2 More on PV-1

Inspection is too late, the quality or lack thereof, is already in the product. Inspection does not improve the quality, nor guarantee quality. As Harold F. Dodge said, “You cannot inspect quality into a product.”

“Quality cannot be inspected into a product or service; it must be built into it.” W.E. Deming in *Out of the Crisis* (2000).

Joseph M. Juran, renowned quality guru, characterized the development process as a hatchery for new quality issues and coined the term “quality by design” to describe the comprehensive discipline required to pro-actively establish quality (Juran 1992).

The pharmaceutical industry has traditionally been highly dependent on end-product testing and inspection. However, this has changed and continues to develop. The appropriate balance of a holistic quality approach versus end-product testing is now common across the industry. Concepts from ICH Q8–Q11 and the FDA guidance for process validation facilitate a move from an inspection-based to a design-based system. The focus of PV-1 is to design a product and associated processing by identifying and controlling process inputs so that the resulting output is of acceptable quality (defined as “what the patient needs”) and well-controlled. The result of PV-1 is to create a manufacturing process with an appropriate risk-based control strategy.

The PV-1 process progresses in the following manner:

1. Develop a Quality Target Product Profile (QTPP).
2. Iteratively design the active pharmaceutical ingredient (API) process, formulation, analytical methods, and final drug product process to achieve the QTPP.
3. Define the “region of goodness” for each process and process input.
4. Determine critical parameters and propose a control strategy.
5. Transition to Manufacturing, PV-2.

The next sections of this chapter overview and connect statistically related tools used in process design. These tools are used to identify good and flexible operating regions, help to determine critical parameters, and propose a control strategy.

Figure 3.1 presents terminology and provides a high level summary of the QbD development process. When starting to develop any product or process, the unknown is called the “unexplored space.” The subset labeled “knowledge space” consists of prior learnings on similar products, first principles understanding of the present process, and empirical information from experiments and other data analyses. After assessing what is known and unknown, the task is to identify and prioritize the knowledge necessary to produce a high quality, safe, and efficacious product. Risk assessment helps in the prioritization and both statistical and non-statistical tools are used to obtain the knowledge. Following an iterative development cycle, the knowledge and scientific experience might lead to several defined regions:

1. A process set point, where if needed, represents where the process is nominally operated.

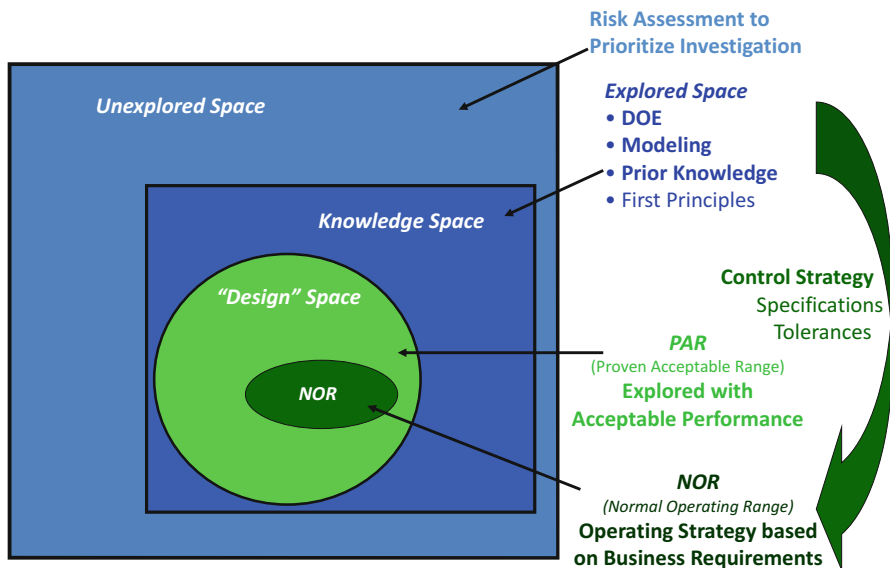


Fig. 3.1 Structure of PV-1 spaces and terminology

2. A normal operating range (NOR) which accounts for variability in the set point.
3. A proven acceptable range (PAR) is a region of goodness which allows for variability in incoming raw materials or otherwise permits flexibility in assuring quality.

Based on knowledge gained through development, parameters and process elements which must be controlled and monitored are identified via the "control strategy." This control strategy allows the manufacturing process to stay within a region of goodness.

The following definitions are useful in navigating the PV-1 landscape.

1. **Attribute:** A characteristic or inherent property of a feature. This term is used in two contexts. The first is as a reference to raw material or excipient features, called material attributes. The other is in reference to the features of the drug substance or drug product. These attributes are significant in defining product safety and efficacy, and are termed critical quality attributes.
2. **Control Strategy:** A planned set of controls, derived from current product and process understanding that ensures a consistent level of process performance and product quality. The controls can include parameters and attributes related to drug substance and drug product materials, facility and equipment operating conditions, in-process controls, finished product specifications, and the associated methods and frequency of monitoring and control (ICH Q10 2008).
3. **Critical Process Parameter (CPP):** A process parameter whose variability has an impact on a critical quality attribute and must be monitored or controlled to ensure the process produces the desired quality (ICH Q8(R2)).

4. **Critical Quality Attribute (CQA):** A physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality (ICH Q8(R2)).
5. **Design Space:** The multidimensional combinations and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality. Movement of a process within the design space is not considered to be a change. Movement out of the design space is considered to be a change and would normally initiate a regulatory postapproval change process. Design space is proposed by the applicant and is subject to regulatory assessment and approval (ICH Q8(R2)).
6. **Knowledge Management:** Systematic approach to acquiring, analyzing, storing, and disseminating information related to products, manufacturing processes, and components. Sources of knowledge include prior knowledge (public domain or internally documented), pharmaceutical development studies, technology transfer activities, process validation studies over the product life cycle, manufacturing experience, innovation, continual improvement, and change management activities (ICH Q10).
7. **Life cycle:** All phases in the life of a product from the initial development through marketing until the product's discontinuation (ICH Q8(R2)).
8. **Normal Operating Range (NOR):** A defined range within (or equal to) the Proven Acceptable Range. It defines the standard target and range under which a process operates.
9. **Parameter:** A measurable or quantifiable characteristic of a system or process ([ASTM E2363](#)).
10. **Process Design (PV-1):** Defining the commercial manufacturing process based on knowledge gained through development and scale-up activities.
11. **Process Qualification (PV-2):** Confirming that the manufacturing process as designed is capable of reproducible commercial manufacturing.
12. **Process Validation:** The collection and evaluation of data, from PV-1 through PV-3, which establishes scientific evidence that a process is capable of consistently delivering quality products.
13. **Process Capability:** Ability of a process to manufacture and fulfill product requirements. In statistical terms, process capability is measured by comparing the variability and targeting of each attribute to its required specification. The capability is summarized by a numerical index C_{pk} (see Chap. 5 for information on this topic). A process must demonstrate a state of statistical control for process capability to be meaningful.
14. **Process Parameter:** A process variable (e.g., temperature, compression force) or input to a process that has the potential to be changed and may impact the process output. To ensure the output meets the specification, ranges of process parameter values are controlled using operating limits.
15. **Process Robustness:** The ability of a manufacturing process to tolerate the variability of raw materials, process equipment, operating conditions, environmental conditions, and human factors. Robustness is an attribute of both

- process and product design (Glodek et al. 2006). Robustness increases with the ability of a process to tolerate variability without negative impact on quality.
16. Proven Acceptable Range: A characterized range of a process parameter for which operation within this range, while keeping other parameters constant, will result in producing a material meeting relevant quality criteria (ICH Q8 (R2)).
 17. Quality: The suitability of either a drug substance or drug product for its intended use. This term includes such attributes as the identity, strength, and purity (ICH Q6A 1999), ICH Q8(R2)).
 18. Quality by Design: A systematic approach to process development that begins with predefined objectives and emphasizes product and process understanding based on sound science and quality risk management (ICH Q8(R2)).
 19. Quality Risk Management: A systematic process for the assessment, control, communication, and review of risks to the quality of the drug (medicinal) product across the product life cycle (ICH Q9).
 20. Quality Target Product Profile (QTPP, pronounced Q-tip): A prospective summary of the quality characteristics of a drug product that ideally will be achieved to ensure the desired quality, taking into account safety and efficacy of the drug product (ICH Q8(R2)).
 21. Risk: The combination of the probability of occurrence of harm and the severity of that harm (ICH Q9).
 22. Risk Assessment: A systematic process of organizing information to support a risk decision to be made within a risk management process. It consists of the identification of hazards and the analysis and evaluation of risks associated with exposure to those hazards (ICH Q9).
 23. State of Control: A condition in which the set of controls consistently provides assurance of continued process performance and product quality (ICH Q10).

3.3 Iterative Process Design

Once the QTPP has been developed, product and process design can begin. Process design is the activity of defining the commercial manufacturing process that will be reflected in master production and control records. The goal of this stage is to design a process suitable for routine commercial manufacturing that can consistently deliver a product that meets the acceptance criteria of its quality attributes.

Process design is iterative and can include all processes associated with the product: API process, formulation, analytical methods, and final product processes. Not all processes are developed in the same manner. For example, the API synthesis process is not developed in the same fashion as a formulation process or an analytical method.

Step 1: Form a team.

A systematic team-based approach to development is the most efficient manner to develop a robust process. This team should include expertise from a variety of disciplines including process engineering, industrial pharmacy, analytical chemistry, microbiology, statistics, manufacturing, and quality assurance.

Step 2: Define the process.

Typically, a manufacturing process is defined by a series of unit operations or a series of synthesis steps. Prior to initiation of any studies, the team needs to agree which unit operations, reactions, or steps are included in the process. To aid in this definition, the team creates a map or flowchart of the process.

A process is a combination of people, machines, methods, measurement systems, environment, and raw materials that produces the intended output. Figure 3.2 displays a process flow diagram for a dry granulation process. Once the process has been defined, meaningful groupings of the unit operations are developed to form the basis for experimentation. Figure 3.3 provides a schematic of these grouping or “focus areas.” The parameters (inputs) and attributes (outputs) for each focus area are discussed and studied in detail. The team discusses a focus area and identifies the attributes and parameters that could potentially affect each attribute. Figure 3.4 provides an Ishikawa diagram (also known as a cause and effect or fishbone diagram) which is helpful in mapping potential sources of parameter variability by categories (e.g., machine, method, manpower, and environment) that could influence attributes.

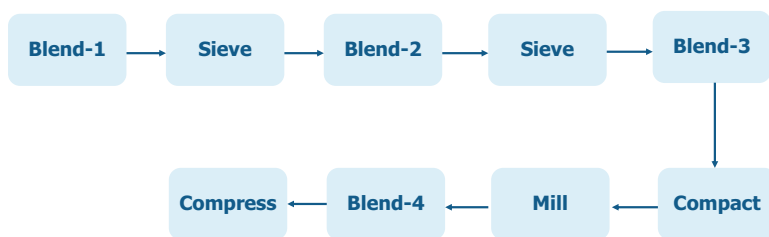


Fig. 3.2 Process flow diagram for a dry granulated product

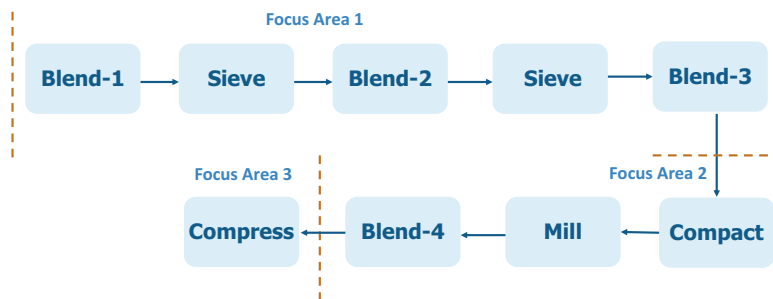


Fig. 3.3 Process map with experimental focus areas

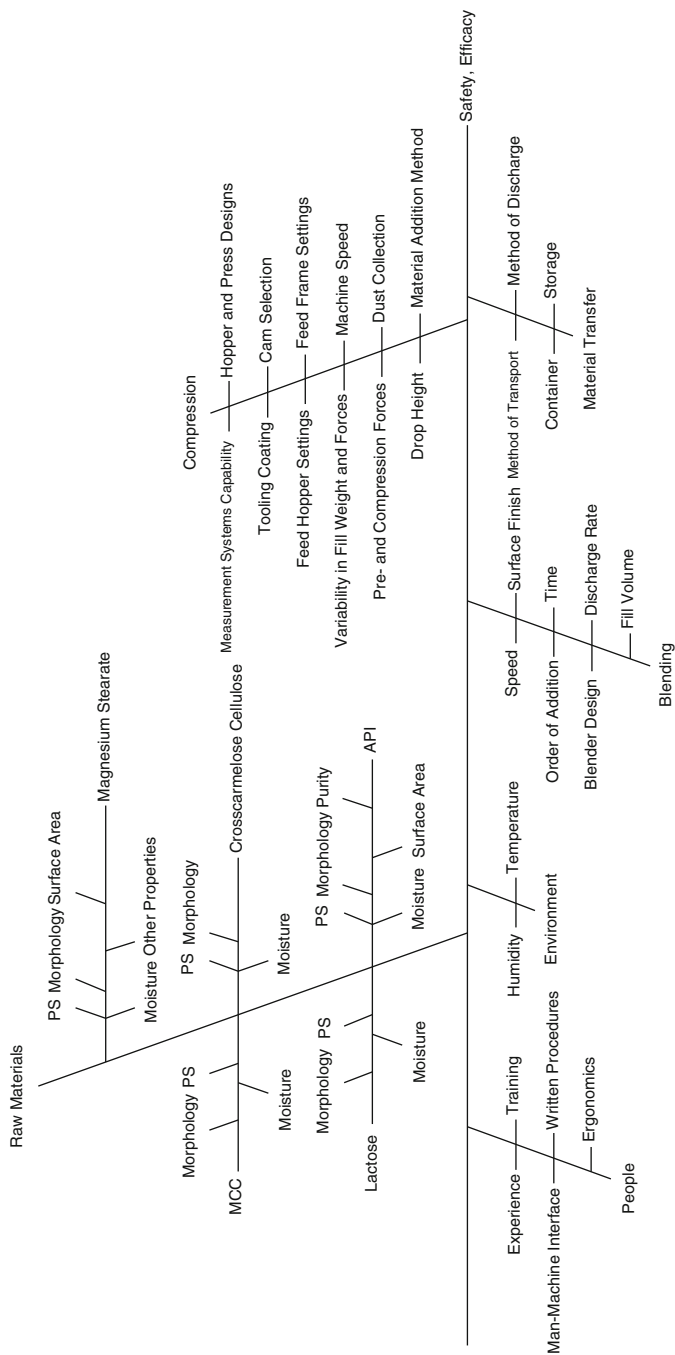


Fig. 3.4 Ishikawa cause and effect (fishbone) diagram (Glodek et al. 2006)

Table 3.1 Risk assessment matrix

Attribute Rank	10	7	10	10	7		
Attribute Parameter	Genotoxic Impurity	Tablet Potency	Drug Release Rate	Shelf Life	Content Uniformity	Score	Strategy
Sampling Method	9	5	5	5	9	288	MSA
Excipient Attribute	10	5	5	9	1	282	DoE
Drug Particle Size	1	9	9	1	9	236	Model
Roll Force	1	9	5	1	9	196	DoE
Screen Size	1	5	1	1	5	100	DoE
Blend Speed	1	5	1	1	5	100	Model
Blend Size	1	5	1	1	5	100	Model
Roll Gap Width	1	5	1	1	5	100	DoE
Compression Force	1	1	5	1	1	84	DoE

Step 3: Prioritize Team Actions

All attributes and parameters should be evaluated in terms of their roles in the process, and on their impact on the final product or in-process material. They should also be reevaluated as new information becomes available.

Following the relationship building of the Ishikawa diagram, the team will select attributes that best define the process. It is typical to perform a risk assessment to prioritize actions taken by the team in developing the process. As per the Pareto principle, it is important to identify the significant parameters for further study. Not all parameters will have an impact and prior knowledge of this improves the impact of the planned studies. Prioritization establishes a risk-based approach to development. Table 3.1 provides the results of a risk assessment.

- The attributes from the Ishikawa diagram are listed across the top and the team of knowledgeable experts rates their importance from 1 to 10, with 10 being the most important in impacting the final product quality.
- The parameters are listed down the left side and the hypothesized or known strength of the relationship between the attribute and parameter is supplied in each box.
- The score is determined for each parameter by multiplying the attribute score and the parameter strength and summing across the attributes. For example, the score for the excipient attribute is $10 \times 10 + 5 \times 7 + 5 \times 10 + 9 \times 10 + 1 \times 7 = 282$.
- The score is sorted from high to low and the strategy to study each parameter is determined.
- The team's actions and the work performed in developing the product are prioritized based on importance as indicated by the total score.

Step 4: Take Action to Understand and Solidify Functional Relationship

Throughout the product life cycle, various studies can be initiated to discover, observe, correlate, or confirm information about the product and process. Knowledge exists in many forms including fundamental knowledge, data-driven models from experimentation, data-driven models on related compounds or equipment, and experimental studies meant to establish or confirm relationships. Studies to gain knowledge are planned for areas where information does not exist. These studies should be planned and conducted according to sound scientific principles and appropriately documented.

Ultimately, the result of the functional understanding is coined as a knowledge space. The region defined as the design space is a subset of the knowledge space. Operation within the design space will ensure product quality. Note that this is clearly not the traditional statistical definition, as a design space in statistics refers to the study range. The design space is meant to be defined in a multifactor fashion and is optional from a regulatory perspective. Another region traditionally defined to represent a region of goodness is the proven acceptable range (PAR). This range has traditionally, although not exclusively, been set in a univariate manner. Rather than compare and contrast a design space with a PAR, suffice it to say each can be called a “region of goodness.” This term is used in future discussion in this chapter to cover both regions. For more on the topic of design space see the papers by Peterson (2004, 2010), Vukovinsky et al. (2010a, b, c), and Stockdale and Cheng (2009).

An initial set point is established within the region of goodness to define the nominal operating condition. Around that point, a normal operating range (NOR) is defined that considers expected operational variability. In the ICH literature, the NOR is permitted to vary within the region of goodness. For example, incoming raw material variability might necessitate a change in set point and NOR or additional process understanding at scale could be used to establish a new set point within the region of goodness.

Step 5: Confirm

Once an NOR is determined, the selected operating conditions or ranges are confirmed. In many cases the NOR has been determined based on experimental design and predictive modeling, but it hasn't been run at either development or full scale. The paper champion needs to be realized and the knowledge confirmed. The initial confirmation might be at development scale (Garcia et al. 2012). The ultimate confirmation, process qualification, is usually conducted at the manufacturing facility where the product will be produced.

Step 6: Document Control Plan

Justification of the controls should be sufficiently documented and internally reviewed to verify and preserve their value for use or adaptation later in the process life cycle. Process knowledge and understanding is the basis for establishing an approach to process control for each unit operation. Strategies for process control can be designed to reduce input variation, adjust for input variation during manufacturing, or combine both approaches. Manufacturing controls mitigate variability to assure quality of the product. Controls can consist of material analysis

and equipment monitoring at significant processing points (21 CFR 211§ 211.110 (c)). Decisions regarding the type and extent of process controls can be aided by earlier risk assessments, then enhanced and improved as process experience is gained. The degree of control over attributes and parameters should be commensurate with their risk to the process. In other words, a higher degree of control is appropriate for attributes and parameters that pose a higher risk. The planned commercial production and control records, which contain the operational limits and overall strategy for process control, should be carried forward to the next stage for confirmation.

Step 7: Iterate as Needed

Typically, all development decisions are not made in one shot. This is an iterative process that continues as new information becomes available.

3.4 PV-1 Statistical Tools

Knowledge is defined as facts, information, and skills acquired through experience or education. It is the theoretical or practical understanding of a subject. Knowledge does not need to be recreated ab initio for every product being developed, but should be created where necessary. That is, in the design process, teams leverage relevant existing data along with fundamental knowledge to make initial decisions, perform risk assessments that identify gaps, and take actions to gain more knowledge. Figure 3.5 summarizes statistical tools that are important in PV-1. These tools include data-based decision making, data collection and experimental design,

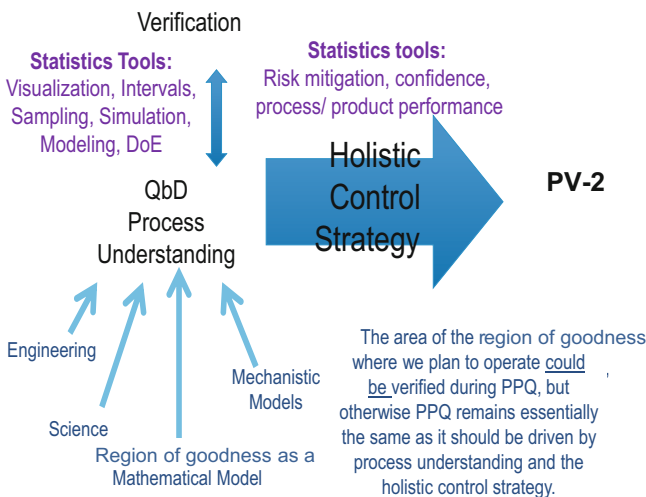


Fig. 3.5 Application of statistical tools in PV-1

descriptive data analysis, and complex modeling. These tools are used to gather, summarize, or quantify knowledge in PV-1.

Some of the PV-1 statistical tools are

1. **Visualization:** It is said that a picture is worth a thousand words. The benefit of effective and simple display of information cannot be overstated and the ability to take a set of data, summarize the information, and visually display this information is both an art and a science. A primary goal of data visualization is to communicate information clearly and efficiently in order to induce the viewer to think about the substance being displayed without distorting or misrepresenting the information. There are many graphical tools available in spreadsheet and statistical software programs. It is necessary to learn these tools in order to present a meaningful data analysis. Section 2.4 provides more discussion on this topic.
2. **Simple Descriptive Statistics:** Descriptive statistics is the discipline of quantitatively describing a set of data. This usually includes a description of the central tendency of the data (mean, geometric mean, median, or mode) and a measure of the dispersion or variability in the data (range, standard deviation, or variance). The data summary can be displayed visually as a boxplot by itself or with other groups of similar data as a comparison. Section 2.4 provides more discussion on a boxplot.
3. **Statistical Intervals (Confidence, Prediction, and Tolerance):** Statistical intervals are the most useful tools for quantifying uncertainty. Section 2.5 discusses these tools in detail.
4. **Sampling Plans:** In pharmaceutical development and manufacturing, sampling is used in many applications. Included are sampling processes used for making batch release decisions, demonstrating homogeneity of drug substance and drug product, accepting batches of raw material, and selection of units for environmental monitoring. Examples of sample plans are discussed throughout this book.
5. **Monte Carlo Simulation:** Simulation is most useful for studying future events that can be predicted from historical data and theorized or established models. The impact of considered changes can be simulated to obtain an understanding of future outcomes under various possible scenarios. Simulation applications are provided throughout this book.
6. **Measurement System Analysis (MSA):** These analyses are referred to as repeatability and reproducibility (R&R) in some industries. They involve the design and analysis of experimental data to understand, quantify, and reduce the variability in the measurement system (analytical method). Variability in the measurement system is normally reduced to categories of bias, linearity, stability, repeatability, and reproducibility. These types of data analysis are critical for the development of useful analytical methods, and are discussed in Chap. 6.
7. **Hypothesis Testing:** Hypothesis testing is a formal statistical process of comparison and inference. Such tests are often required by regulatory agencies in many evaluations. This topic is discussed in Sect. 2.10.

8. **Models and Modeling:** Prior to running experiments, information based on either first-principles or data-driven models should be exercised to help inform relationships.
9. **Data-driven Modeling:** Data-driven models are developed through fitting models to data. In PV-1, there is often data related to the process or compound being developed. Sometimes, as is the case with material property data, chemical structure data, and processing data, small to large data sets exist and data-driven models are developed to best express relationships. In the case of a material property data base, relationships between material properties and product attributes would be examined and data-driven models developed to predict product properties based on the material attributes. These data-driven models permit a decrease in experimentation, or at least provide a starting point for further experimentation. Common modeling techniques include simple linear regression, partial least squares, regression trees, and machine learning algorithms.
10. **First-principle or Fundamental Models:** First-principle, engineering, physics, or fundamental models explain relationships between parameters, material attributes, or manufacturing factors and product attributes. These models seek to predict product attributes directly from established laws of science.
11. **Design of Experiments (DoE):** DoE is a highly used tool in investigating unknown relationships within the framework of PV-1. DoE provides a systematic approach to study prioritized factors and establish a relationship with quality or in-process attributes. More information on DoE is provided in Sect. 3.5.

3.5 Design of Experiments

To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of.

R. A. Fisher (1890–1962)

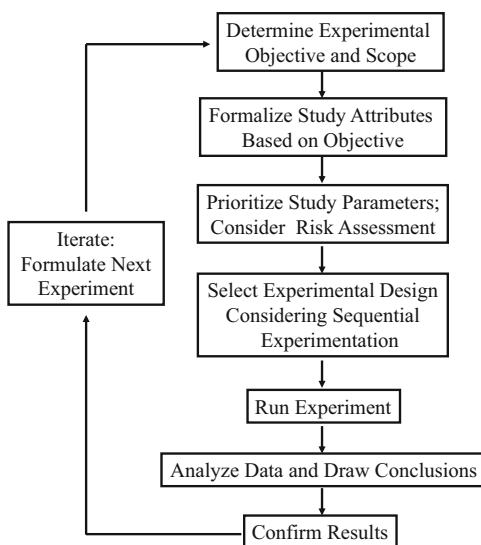
DoE has become a bedrock of the framework of PV-1. Why has this become such an integral part of the process? The strength of DoE is in the application of a systematic approach to data-based decision making along with the selection of a study design. Because of the complexity of most processes, several factors are usually studied in a series of experiments. Historically, students learn to vary one-factor-at-a-time (OFAT) and this practice is applied on the job in research, development, and manufacturing. A reason provided in support of this approach is that if more than one factor is changed, the experimenter will not be able to determine which factor was responsible for the change in the response. In reality, the proper selection of experimental runs combined with the proper analysis removes this source of concern. In addition, there are two major deficiencies with an OFAT study. The first is that there are often interactions between the parameters under study. An interaction means that the effect a process parameter has on the

response may depend on the levels of another process parameter. Statistical experimental designs that permit the estimation of interactions will allow for their study, whereas OFAT studies do not. The other deficiency of OFAT is that data previously collected to study other factors is set aside and new data are collected. The structure of the statistical experimental design allows all the data from the entire study to be used to draw conclusions on each factor. This results in savings of both time and money over the OFAT process. In fact, the statistical design approach provides a proper design structure that when combined with the analysis method maximizes the amount of information for the minimum number of runs (i.e., the knowledge development process is highly efficient)

Many textbooks and papers have been written on this subject, and the reader is encouraged to have some of these books in a personal library. Three books are Box et al. (2005), Montgomery (2012), and Morris (2011). Since so much is available on the topic, there is no intention to provide comprehensive technical details in this book. Rather, the focus in this chapter is on the high level application of DoE within a PV-1/QbD environment.

Underlying all processes are mathematical and statistical models, the behavior of which is interrogated via experimentation. Designing an efficient process with an effective process control strategy is dependent on the process knowledge and understanding obtained. DoE studies can help develop process knowledge by revealing relationships between process parameters and the resulting quality attributes measured on process outputs. Efficiently determining an approximate equation representing the underlying physical equation is best accomplished by DoE and an effective experimental strategy. The experimental design process is only one element of QbD, and it closely follows the QbD process as shown in Fig. 3.6.

Fig. 3.6 Experimental design flowchart



As shown in Fig. 3.6, the experimental process consists of the following seven steps:

1. Determine the experimental objective and scope: Before initiating any series of experiments, define the purpose of the study. Is the goal to improve yield, increase selectivity or the reaction, achieve a particular dissolution profile while minimizing content variability, or optimize a potency method? It is very important to be clear about the purpose of the experiment and the decisions to be based on the study. Under the QbD paradigm, agreement on the goal and alignment on the experimental objective is especially important. Everyone on the team needs to be progressing toward the same goal.
2. Formalize study attributes based on the objective: Agree on the attributes (responses or outputs) to study in the experiment. Attributes should be aligned with the experimental objective. In addition, an important consideration at this stage is inclusion of attributes that are not primary to the objective. Selected attributes should include those directly related to the experimental objective, and those not of primary interest, but with an ability to impact the study later in the process. In addition, selection of a measurement system, how responses will be measured, and the required level of precision should all be considered before running the experiment.
3. Prioritize study parameters which may affect the responses: Define the factors (e.g., process parameters, material attributes, and starting materials) that are hypothesized to have an impact on the quality attribute responses. Scientifically analyze the issue at hand and the process under study in order to determine all the possible factors influencing the situation. This could be achieved by examining literature reports or other prior knowledge, employing fundamental or mechanistic understanding, or through preliminary laboratory experiments. Judicious selection of factors is important in keeping the number of experiments manageable. Consider the appropriateness of a single large design, or a series of reduced sequential designs. A risk assessment can be used to prioritize variables for DoE studies. There are two ways to usefully categorize process performance parameters. One is traditionally statistical, and the other is of special consideration within the pharmaceutical industry. In each case, there are special issues with each parameter type that should be addressed prior to designing and running the experiment.
 - a. *Control and noise factors*: Selecting factors to control is not always an easy decision. In fact, even if a factor is not one of the process parameters to be controlled during manufacturing (so-called noise factors), it might be beneficial to control the noise factor during experimentation. For example, humidity at the site might be an uncontrollable factor during manufacturing. However, if an effect on response attributes can be demonstrated during experimentation which might drive improvement of the manufacturing process to mitigate the effect of humidity. Thus, it is important to consider all factors that can impact values of the quality attributes in designing your experiment.

- b. *Scale dependent vs. scale independent or scalable*: Parameters that are scale independent or scalable can be studied at a smaller scale than full scale commercial manufacturing and results are applicable for full scale. With scale dependent parameters, there is a dependency of the results based on the scale of the equipment. A strategy is needed to assess the DoE results at scale. Examples of scale independent factors include pressure, temperature, and Gerties roller compactors. Examples of scale dependent factors include mixing rpm and high sheer granulation.

In addition to defining process parameters, it is necessary to define the experimental domain by assigning the upper and lower limit ranges to all continuous variables. For discrete variables, one must define categories. Probably the most difficult component of designing an experiment is selection of the levels or ranges for each parameter. Consider an experiment with the process performance parameter “revolutions per minute (rpm)”. How does one decide to set the low and high study levels to (75, 125), as opposed to (50, 100) or (50, 150), or (75, 150)? Selection of such ranges depends on the experimental objective and the overall development strategy. In general, the range limit span should be as wide as is practical, but neither too large nor too narrow. If limits are too narrow, there is a risk of not seeing the parameter effect. If the range is too wide, the parameter will be characterized on a macro level, but may not provide information on the micro level and hide effects of other parameters. It is often helpful to examine existing experimental data, fundamental knowledge, and similar compounds or processes of interest.

4. Select experimental design and consider sequential experimentation: Experimental design is based on the following principles:
 - a. *Randomization*: Statistical methods require that observations be independently distributed random variables, and randomization helps make this assumption valid. Randomizing helps “average out” uncontrolled noise variables (lurking or extraneous variables). There are situations where the experiment is not run in a completely randomized fashion due to practical situations. However, this should be by design and the data analyzed in a manner consistent with the design.
 - b. *Blocking*: Blocking removes unwanted variability and allows focus on the factors of interest. Pairing is a special type of blocking. As an example, consider a comparison of the bias for two analytical methods. It is expected that there will be variation among the test samples measured in the experiment. For that reason, a paired design requires that each method be used to measure each test sample. In this manner, variation among test samples will not manifest in differences between the measured values from the two analytical methods.
 - c. *Replication*: Replication allows the estimate of experimental error to be obtained. This estimate is the basic unit of measurement for determining whether observed differences in the data are *statistically* different.

Replication represents the between run variability. True replication means that the process is completely restarted for each replicate run. Often times, the person running the experiment will merely take repeated measures from a single experimental setup. This is not a true replication, but rather repetition. The manner in which one employs these three principles is impacted by the realities of the scale of the design, and time and other resource constraints. It is easy to get confused by the many statistical designs that are described in the literature.

Design selection is largely dependent on the objective, goals, process knowledge, and the stage of the experimentation. Sections 3.5.1 and 3.5.2 provide high level descriptions of some common experimental designs.

5. Run the experiment: The details in this step are often ignored. It is extremely important that the one who carries out the experiment understands the underlying details of the experiment. In general, run the experiments in a random order to distribute unknown sources of variability and minimize the effect of systematic errors on the observations. It may be tempting to re-order for convenience, but running an experiment in such a non-random pattern can create problems with the data analysis. Some designs can account for planned restrictions on randomization (e.g., split-plot or hard-to-change factors). However, any such restrictions should be built into the experimental design.

It is important for the person running the experiment to understand the process. The person should understand the difference between experimental factors and factors which should stay fixed during the entire DoE. It may be tempting for those who truly understand the process to make slight tweaks to try and “save” a run. Such adjustments are not allowed, and risk destroying the study conclusions. Review the experimental protocol and determine if there is something that should be changed before running the experiment. Record actual levels of the process parameters and note if they deviate from the planned levels. Record other unexpected events.

6. Analyze the data and draw conclusions: If care was taken in setting up the experimental plan and the experiments were executed as expected, then the data analysis will be relatively straightforward. It might be necessary to meet with a subject matter expert should any of the data appear as outliers, or if the analysis results have influential observations or the models don't appear to be feasible.
7. Confirm results and document: Verify as necessary the decisions made based on the experimental data and the model. It is very important to perform confirmation runs as necessary to verify the best predicted condition. Often times, the best model predicted condition or region has not been run in the experiment, so physical confirmation is highly recommended. Plan the next sequential design, if appropriate.

Results of the experimental effort must be documented. Recommended documentation should include a brief description of the background information that led to the experimentation, the objective, study process parameters (including names,

levels, and units), response attributes (including measurement method), details of replication and randomization, the design matrix, a summary of the data, the statistical methods and software used in the analysis, and summary results.

3.5.1 *Full and Fractional Factorial Experiments*

The factorial family provides powerful and flexible designs for collecting information on main effects and interactions in a minimum number of runs. They are highly flexible and can be used in the screening and interaction phases and as a base for optimization. The ability to add onto these designs facilitates sequential experimentation and enhanced refinement of knowledge.

To execute, factors for experimentation are selected and a fixed number of “levels” (usually high-low) are defined for each parameter. A full factorial design considers all possible combinations of the levels of each input factor. This design permits estimation of the main effects and interactions. In general, assume there are ℓ_1 levels for the first factor, ℓ_2 levels for the second factor, and ℓ_k levels for the k th and last factor. The complete arrangement of $\ell_1 \times \ell_2 \times \dots \times \ell_k$ experimental combinations is called a full factorial design (e.g., a $2 \times 2 \times 3$ full factorial design yields 12 experimental runs). A full factorial design including five factors varying each factor across two levels is written as 2^5 , and has 32 experimental runs.

The 2^5 full factorial design permits estimation of the five main effects, 10 - two-way interactions, 10 three-way interactions, five four-way interactions, and one five-way interaction. The remaining experimental run is used to estimate an overall mean.

It is usually not required to estimate all multifactor interactions, and so a specific fraction of the full factorial is selected. This necessarily results in a reduction of the number of experimental runs. This so-called fractional factorial design is a mathematically correct subset of the full factorial that permits estimation of main effects and some subset of interaction effects. Some loss in experimental information (i.e., resolution) generally results by fractionating, but knowledge of the desired information can be used a priori to select an appropriate fraction. For example, a half fraction of the 2^5 full factorial design includes $2^{5-1} = 16$ runs. This design permits estimation of the main effects and all of the 10 two-factor interactions. Figure 3.7 displays a 2-level full factorial with 3 factors (A, B, C) on the left, and a half fraction of that design on the right.

Understanding the structure of a factorial experiment is important as a base to understanding all designed experiments. An experiment to study the effect of factors A and B on attributes of interest would consist of the four unique runs: (A low, B low), (A low, B high), (A high, B low), and (A high, B high). This is denoted as a 2^2 full factorial experiment, and is shown in Table 3.2. A full factorial experiment to study three factors at two levels, 2^3 , has eight unique runs as shown in Table 3.3.

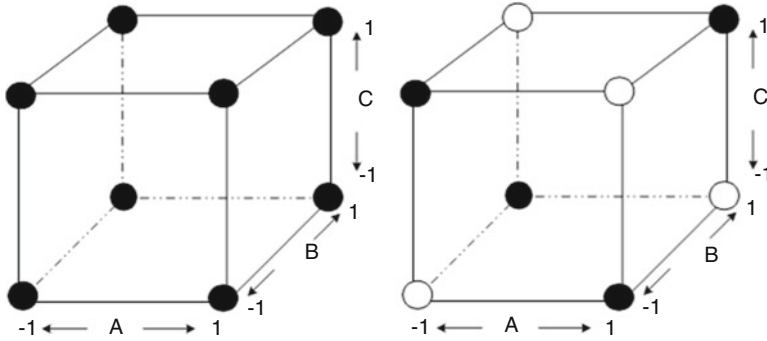


Fig. 3.7 Two level three full factorial (*left*) and half fraction factorial (*right*)

Table 3.2 2² full factorial experiment

Standard order	A	B
1	Low	Low
2	Low	High
3	High	Low
4	High	High

Table 3.3 2³ full factorial experiment

Standard order	A	B	C
1	Low	Low	Low
2	Low	Low	High
3	Low	High	Low
4	Low	High	High
5	High	Low	Low
6	High	Low	High
7	High	High	Low
8	High	High	High

As noted earlier, it is generally accepted that even complicated relationships between parameters and attributes can have a large proportion of the relationship explained by linear effects. Less can be explained by interactions, and even less from nonlinear or quadratic effects. Hence, 2-level factorial experiments are all that is required in many cases.

An example is now provided to demonstrate the power of the two-level factorial structure. In the development of a wet granulation process, it is desired to study the impact of impeller speed, binder level, and binder addition rate on the average particle size (D50). Since it is desired to look at all possible combinations of the three process parameters, the selected design is the 2³ full factorial experiment shown in Table 3.4. The experiment was run in a random order and three replicated center point conditions were run in addition to the eight factorial

Table 3.4 Wet granulation design with particle size data

Standard order	Impeller speed (A)	Binder level (B)	Binder addition rate (C)	D50
1	Low	Low	Low	156.5
2	Low	Low	High	146.3
3	Low	High	Low	198.8
4	Low	High	High	209.5
5	High	Low	Low	158.4
6	High	Low	High	161.6
7	High	High	Low	136.4
8	High	High	High	142.7

Table 3.5 2^3 design illustrating all estimable effects

Standard order	Impeller speed (A)	Binder level (B)	Binder addition rate (C)	$A \times B$	$A \times C$	$B \times C$	$A \times B \times C$	D50
1	-1	-1	-1	1	1	1	-1	156.5
2	-1	-1	1	1	-1	-1	1	146.3
3	-1	1	-1	-1	1	-1	1	198.8
4	-1	1	1	-1	-1	1	-1	209.5
5	1	-1	-1	-1	-1	1	1	158.4
6	1	-1	1	-1	1	-1	-1	161.6
7	1	1	-1	1	-1	-1	-1	136.4
8	1	1	1	1	1	1	1	142.7

runs (not shown in the table). The results are in Table 3.4 in standard order without the center points to simplify the analysis and more effectively demonstrate the power of the analysis.

In the analysis of this particular experiment, it is possible to obtain the linear effects of the parameters (A, B, C), the two-way interactions ($A \times B$, $A \times C$, $B \times C$), and the three-way interaction ($A \times B \times C$). Now, replace the word “low” in Table 3.4 with a “-1” and the word “high” with a “1”, as shown in Table 3.5. Notice that the sum of multiplied values in the same row of any two columns is equal to zero. In matrix algebra nomenclature, such columns are said to be linearly independent. An experimental design in which all columns are linearly independent is said to be an orthogonal design. An orthogonal design permits estimation of all of the effects individually without interference from any other effects. Notice that each column in Table 3.5 is unique. All eight values of D50 will be used to estimate all seven effects.

Table 3.6 displays this same information in an alternate form. White space in Table 3.6 indicates the correct -1 or 1 positioning of the data for each effect.

Data analysis will normally be conducted using a computer program. For this example, a simple analysis representation which will match a computer analysis is shown in Table 3.7. Each D50 value is placed in the column of its row

Table 3.6 Alternative representation of Table 3.5

Random Order Trial Number	Standard Order Trial Number	Response (Observed Value)	Impeller (A)		Binder Level (B)		Binder AddRate (C)		AB		AC		BC		ABC	
			-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
	1	156.5														
	2	146.3														
	3	198.8														
	4	209.5														
	5	158.4														
	6	161.6														
	7	136.4														
	8	142.7														
Total																
Number of Values		8	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Average																
Effect																

corresponding to the level of the performance parameter for which it was collected. For example, in the first row, A is at level -1 , and so D50 in the first row is placed in the -1 column of A. From this display, patterns may become apparent, and certainly, data from standard order trial numbers 3 and 4 appear greater than the rest of the data.

To perform the analysis, add each column of values and place the sum in the row labeled “Total.” For example, the sum for Impeller (A) at -1 is $156.5 + 146.3 + 198.8 + 209.5 = 711.1$. Next average each column by dividing the column total by the total number of observations included in the total (4 in this example). For Impeller (A) at -1 , the average is $711.1/4 = 177.78$. Comparing the average between the low and high level of each factor, it is observed that some of the differences are large (e.g., 177.78 versus 149.78 for Impeller (A)) and some of the differences are small (e.g., 162.53 versus 165.03 for Binder Addition Rate (C)). The difference between the $+1$ average and the -1 average is summarized into a factor effect shown in the last row of the table. By this method, the effect of A is found by $149.78 - 177.78 = -28$. From this row, it can be seen that $A \times B$, A, and B are the largest effects.

Table 3.8 presents results of a regression model as described in Sect. 2.12 that is fit to include the three large effects (A,B, $A \times B$) using the data in Table 3.5.

Note the intercept term is the overall average of all eight values of D50. The regression estimate (slope) for each parameter is equal to the effect value in Table 3.7 divided by two. Recall that the effect is the difference from low (-1) to high ($+1$), whereas the slope is the difference for one unit change (e.g., from -1 to 0 or from 0 to $+1$). Since the overall average of the attribute D50 represents a

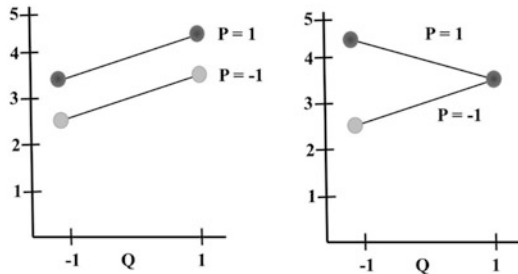
Table 3.7 Data analysis, for example

Random Order Trial Number	Standard Order Trial Number	Response (Observed Value)	Impeller (A)		Binder Level (B)		Binder AddRate (C)		AB		AC		BC		ABC	
			-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
	1	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5	156.5
	2	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3	146.3
	3	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8	198.8
	4	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5	209.5
	5	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4	158.4
	6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6	161.6
	7	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4	136.4
	8	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7	142.7
Total		1310.2	711.1	599.1	622.8	667.4	650.1	660.1	728.3	581.9	650.6	643.1	667.1	664	646.2	
Number of Values		8	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Average		163.775	177.78	149.78	155.7	171.85	162.53	165.03	182.08	145.48	162.65	164.9	160.78	166.78	166	161.55
Effect			-28		16.15		2.5		-36.6		2.25		6		-4.45	

Table 3.8 Estimates of regression slopes

Term	Estimate	Prob > t
Intercept	163.775	<0.0001
Impeller speed (A)	-14	0.0024
Binder level (B)	8.075	0.0169
Binder level (B)*Impeller speed (A)	-18.3	0.0009

Fig. 3.8 Examples of interaction plots



model baseline, the estimates describe the amount of change from baseline as a given factor moves from -1 (low) to +1 (high).

As discussed in Sect. 2.12, existence of an interaction $A \times B$ means that the effect of A on the response attribute depends on the selected level for B. This means information from two levels of each parameter is required to decompose the shape of the interaction. In the case of the example, the existence of the $A \times B$ interaction points to its strength but not to the functional nature. The functional nature can be described using an interaction plot.

Two interaction plots are provided in Fig. 3.8 for factors P and Q. The vertical axis represents the response attribute, and the horizontal axis shows the two levels of Q. There is one line on the plot for each level of P. The circles represent the average of the response attribute at the given combination of P and Q.

No interaction exists between P and Q in the plot on the left. This is because the lines are parallel, and the amount of change in the response attribute as Q changes from -1 to +1 is constant for both values of P. The y-intercept is different for the two lines, but the rate of change (slope) is identical. Thus, the change in the response attribute as a function of Q is not dependent on the setting for P. Such effects are said to be additive rather than interactive. Similarly, the change in the response as a function of P is not dependent on the setting for Q. (This can be demonstrated by placing P on the horizontal axis, and drawing a line for each level of Q.)

The interaction plot on the right of Fig. 3.8 indicates a strong interaction between P and Q. Notice that as Q moves from -1 to +1 when $P = 1$, there is a decrease in the average response. However, if $P = -1$, movement of Q from -1 to +1 results in an increase in the average of the response. This is an important discovery when interactions of this magnitude exist, and provides important information to be considered in process development.

Table 3.9 Calculations of interaction example

Run	Response, No Interaction	P (-1)	P (1)	Q (-1)	Q (1)	P*Q Low, No Interaction	P*Q High, No Interaction	Response, Yes Interaction	P (-1)	P (1)	Q (-1)	Q (1)	P*Q Low, Yes Interaction	P*Q High, Yes Interaction
1	2.5	2.5		2.5			2.5	2.5			2.5			2.5
2	3.5	3.5			3.5	3.5		3.5	3.5			3.5	3.5	
3	3.5		3.5	3.5			3.5	4.5	4.5			4.5	4.5	
4	4.5		4.5		4.5		4.5	3.5	3.5			3.5		3.5
Total	14	6	8	6	8	7	7	14	6	8	7	7	8	6
Number of Values	4	2	2	2	2	2	2	4	2	2	2	2	2	2
Average	3.5	3	4	3	4	3.5	3.5	3.5	3	4	3.5	3.5	4	3
Effect		1		1		0			1		0		1	

The analysis of the data from these two graphs is provided in Table 3.9. The no interaction graph on the left of Fig. 3.8 shows visually that as Q changes from low to high, there is a 1 unit change in the response. As P changes from low to high, there is also a 1 unit change in the response and there is no dependency between P and Q. Note the sum that defines the effect of the interaction is 0. This must be true when there is no interaction, and the lines when graphed will be parallel. On the other hand, the interaction graph on the right shows there is not consistent behavior in the effect of Q changing between P low and P high. Table 3.9 calculates the effects given this situation. In this case, the effect of P changing from low to high is 1, the effect of Q changing from low to high is 0, and the effect of the interaction is 1. The weight of importance in correctly understanding the situation has shifted from an individual parameter effect to the interaction effect.

3.5.2 Other Experimental Designs

Other experimental designs are now briefly discussed.

1. *Plackett–Burman Designs (PBD)*: The PBD design is used in screening where one has a large set of candidate factors and it is necessary to select a small set of the most important factors. Unlike the factorial design structure, the PBD design is constructed in multiples of four rather than powers of two. For example, a PBD design with 12 runs may be used for an experiment containing up to 11 factors. These very economical screening designs are most normally used when only main effects are of interest and are most useful if you can safely assume that interactions are not significant. Another useful application is in ruggedness testing or confirmation within a region of goodness where there should not be an effect on the attributes of interest. Alias structures can be very messy in some situations and it is advised that someone with experience in experimental design be consulted in selecting an appropriate design. Because these designs are used for screening, follow-on designs are usually conducted with the process parameters identified as significant. PBDs are difficult to augment except under specific circumstances when combined with a computer optimal design.

2. *Central Composite Designs (CCD)*: The CCD is used in optimization or to map a region of interest in more detail. These are response surface designs to which a full quadratic model can be fit. The CCD is part of the factorial family of designs and contains a factorial or fractional factorial design that is augmented with both center and axial points. As the name implies, axial points appear on the axis outside of the cube defined by the full factorial corner points. If the distance from the center of the design to a factorial point is defined to be ± 1 unit for each factor, then the distance from the center of the design to an axial point is $\pm\alpha$ with $|\alpha|$ is greater than or equal to 1. The precise value of the distance depends on the properties desired for the design and on the number of factors included in the design. The axial points require that each design factor can be changed across either three or five levels. Similarly, the number of center points depends on preferred design properties.
3. *Box–Behnken Designs (BBD)*: The BBD is a three-level design used for fitting response surfaces. BBDs are experimental designs used to fit a model which includes main effects, two-factor interactions, and quadratic effects. They are formed by combining 2^k factorials with incomplete block designs. In the experiment, each factor is placed at one of three equally spaced values, usually coded $-1, 0, +1$. The design itself is structured as a series of two level (full or fractional) factorial designs ($-1, +1$) in usually 2–3 factors while the other factors are kept at the center (0) values. In this design, several center points are run. The structure of the BBD provides a convenience of not running at extremes, should the extreme be a concern. However, the predictive ability is not generally as good as the CCD. Like a CCD, these designs can be augmented. The augmentation for BBD permits estimation of cubic and quartic effects. In the case of 3–4 factors, the BBD will require a fewer number of experiments than the CCD.
4. *Split-Plot Fractional Factorial Designs*: Split-plot designs are required when there are constraints on randomization of the experimental runs. For example, the temperature of an incubator cannot be randomly changed across units placed in the same incubator. Factors such as temperature in this example are referred to as hard to change factors. Hard to change factors appear often in CMC applications, and proper analysis of the data requires proper recognition of these factors in the experimental design.
5. *Mixture Designs*: Mixture designs include factors that are compounds or ingredients in a mixture. The objective of these experiments is to determine the optimal proportion of each ingredient in order to accomplish some objectives.
6. *Computer Optimal Designs*: Computer optimal designs allow alternatives that are not considered in the more classical designs. In particular, they allow definition of an experimental range that is not defined by a cube or sphere. They also allow selection of specific models that might include a pre-selected subset of interactions. They also allow the opportunity to select designs with small sample sizes relative to the number of parameters to be estimated. These designs are popular as they can reduce the required number of experiments and are helpful in augmenting experimental runs to a previously designed study.

They can also be helpful in tricky situations, such as when there are an uneven number of levels of the experimental factors, when certain combinations of the factors cannot be run, or when multiple level discrete factors combine with continuous and mixture factors. Care should be given in employing this design as the design is only optimal if the pre-specified model is current. This requires understanding of the underlying mathematics of statistical experimental design and practical knowledge of the process under study. Two popular criteria include both D-optimal and I-optimal designs. The D-optimality criterion minimizes the joint confidence region of the regression coefficients, and I-optimality minimizes the average prediction variance over the design space.

There are many other designs that are useful in special applications, and new designs to be developed. Some of these other designs include saturated designs (designs where the number of parameters is equal to the number of data points), definitive screening designs, and hybrid designs. Information on these designs can be found in the statistical literature.

3.5.3 Experimental Strategy

Determination of an experimental strategy is both an art and a science. If research studies are sequential in nature or cover multiple unit operations, it may be advantageous to break up a study into parts. Strategy depends on prior knowledge, available time, and material and equipment availability. Regardless of the particular intricacies of a situation, it is best to make decisions as expeditiously and efficiently as possible. To do so, a hierarchical effect principle is employed. Many processes involve complicated relationships between process parameters and attributes. In general, the large portion of the relationship can be explained by the linear effect, less by interactions between parameters, and less again by a nonlinear or quadratic effect. It takes two experiments (low, high) to estimate a linear trend, four experiments to estimate an interaction between two parameters, and three to five experiments (depending on the nonlinearity) to estimate curvature. This generality is consistent with a strategy to first understand linear relationships and interactions, and then examine curvature as needed.

Consider two such examples:

- HPLC process parameters are known to be linear in their effect on certain attributes. The signature for a new piece of HPLC equipment may be unknown, but the underlying trend in the parameter to attribute relationships could be well known based on prior fundamental and experimental knowledge. A couple of familiarization runs plus a screening design combined with prior knowledge might be all that is needed to establish the functional relationship between the process parameter and attribute for the compound being developed.
- In developing understanding around an active ingredient's synthetic route with minimal prior knowledge, one might require all stages of experimentation. As a

first familiarization step, a small number of experiments at the extremes could be run to gain knowledge on the compound and equipment. A screening experiment could then be run to identify the significant few from the trivial many parameters. Once the important 3–5 parameters are determined, a factorial or central composite design is run to estimate interactions and quantify nonlinearities.


In each of these examples, a scientist works to understand the particular strategy and integrate all prior knowledge and tools in order to set up the most efficient experimental strategy. Table 3.10 presents the four stages of strategy:

1. Familiarization,
2. Screening,
3. Interaction, and
4. Optimization.

Each stage is now described in more detail.

1. *Familiarization*: As the name implies, the basic purpose of this phase is to better understand the problem at hand. The experimenter should keep in mind that engaging in a full DoE without a basic understanding of the system practically assures a study of limited value. If the system is well known, this step can be skipped. There are no set guidelines or specific requirements for executing familiarization runs as part of an experimental design (with perhaps the

Table 3.10 Overview of experimental strategy by level of understanding

	Little System Knowledge		Detailed Understanding 	
	Familiarization Phase	Screening Phase	Interaction Phase	Optimization Phase
# Parameters (Factors)	2 - 15	5 - 15	2 - 8	2 - 5
# Experiments	2 – few	11 – 19	7 – 35	11 - 31
Questions to be Answered	If there is little knowledge about the study environment, a couple of runs to establish ranges and investigate the system is valuable. Need to be aware of when to stop and move on to formal efforts.	What factors should be further studied? Is there a workable solution? Should the ranges be adjusted in the next study? Is there an area or direction of goodness? Are any responses nonlinear? What is the control strategy?	Are the center points repeatable and reproducible? Which factors interact? Is there a potential solution? Is there an area or direction of goodness? Are any responses nonlinear? What is the control strategy?	Are the center points repeatable and reproducible? Are nonlinear responses adequately modelled? Is there an area or direction of goodness? What is the control strategy?

exception of the initial runs in a sequential simplex). However, a familiarization phase is essential. The following outcomes would generally describe a successful completion of this stage:

- Any new equipment has been tested and enjoys a degree of reliability.
- Potential performance parameters have been identified with some degree of certainty.
- A range for the performance parameters has been defined that appears practical from a process point of view (i.e., not difficult to control and are scalable) and provide results that are not extraordinarily atypical.
- At least several replicate runs have been completed to estimate the system variability.

2. *Screening*: The main purpose of a screening design is to select a small number of performance parameters from a large set of potential parameters in a minimal number of experiments. Many times, one can identify several potential performance parameters after only a few experiments. At this very early stage, the relative impact of these parameters on the quality attributes may be based more on prior knowledge than on empirical experimentation. Since there is a severe penalty in terms of the number of experiments required to complete a full factorial design, the wise experimenter will embark on a full DoE with only those process parameters that are truly important in this stage.

Screening designs are obtained by using fractional factorial designs, Plackett–Burman designs, or computer optimal designs. One drawback to screening designs is they have a complicated confounding of interactions. However, any process parameters that affect the attributes to an extent greater than the experimental error will be identified. Although some modeling can be done with the data, the basic idea is that once a screening design is completed, the experimenter will eliminate the superfluous variables and embark on a more detailed study of the important process parameters using higher resolution factorial designs.

Screening designs can also be used for purposes that don't require additional experimentation. One such case is the confirmation of an area of robustness. In demonstrating robustness, key process parameters are studied across their recommended manufacturing ranges and the responses or quality attributes are measured. The effects of raw materials and environmental or human factors may be considered in the experiment as noise factors. Noise factors are controlled at the time of the DoE and then left uncontrolled in process operation. Robustness demonstration can be conducted in conjunction with the optimization phase. Variability reduction activities take place in manufacturing, but are best performed early in the development life cycle.

3. *Interaction*: The interaction phase permits study of interactions between the input process parameters. During this phase fewer input factors are studied as compared to the screening phase, because knowledge of interaction effects requires more testing than knowledge of main effects. There may also be a greater level of process understanding. The several process parameters being

studied are believed more likely significant than the many studied in the screening phase. If there are no more than 5–6 process parameters, and only 2-factor interactions are of interest, the screening and interaction phase may be conducted at the same time.

Possible designs considered in this phase include fractional to full factorial designs, depending on the level of interaction required in the study. Data from the screening design can be used in a “fold-over” study (Box et al.) to reduce the total number of runs. A computer optimal design may be used if the region is of unusual shape, if a known model exists, or if design modifications are required unexpectedly in the process of running the experiment (e.g., design repair).

4. *Optimization*: Optimization refers to examination of nonlinear effects, usually quadratic effects, about a smaller region of interest (e.g., the NOR). This typically occurs following the interaction and the screening phases. Designs used in this phase include central composite designs, optimal designs, and Box–Behnken designs. The most popular designs in this phase are central composite designs as in many cases information from experiments included in the screening and interaction phase can be reused and included in the study design and analysis. Note that the screening, interaction, and optimization phases do not need to be sequential and can be conducted simultaneously.

Strategic questions to answer that are crucial to proper execution of an experiment include the following:

1. How will the responses be measured? What measurement system will be used? What is the expected variability?
2. What performance parameter factors are hypothesized to have the largest effect on the quality attributes of interest?
3. Are there any known interactions between factors? Increased prior knowledge can help in decreasing the required experiments.
4. How will the rest of the parameters and material attributes be controlled or blocked during the experiment?
5. Are there noise factors which cannot be controlled? How can their effect be minimized? Should blocking be used to minimize the effect of the hard to control sources of variability?
6. Can the entire experiment be randomized or is this not practical? Should there be a partial randomization scheme?
7. How many replicates are needed for each attribute in an experimental run? It is often both acceptable and necessary to perform unreplicated experiments, but it is important to understand the considerations of these experiments. For example, consider tablet potency as a quality attribute for a study. How will potency be measured? Certainly not by a single replicate assay injection from a single tablet. More likely, it could be measured as the average of two replicate HPLC injections from a composite of five tablets. In general, knowledge of past estimates of variability for similar compounds or similar processes will help inform the replicate strategy. Ultimately, this subject is important to ensure

sufficient statistical power to detect differences which are meaningful to the experimental objective.

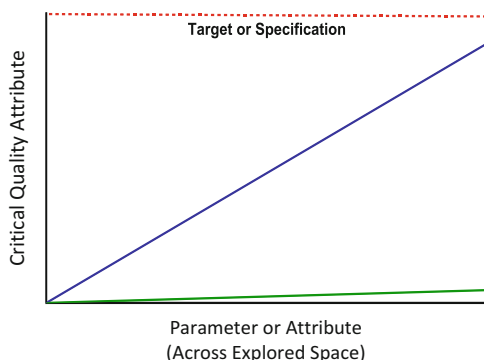
8. Will center points be used to estimate variability? For example, in a 16 run factorial design, it is usually of benefit to run at least three center points at the beginning, middle, and end of the experiment. These center points are used to assess variability across the experiment and also to judge nonlinearity or curvature in the experimental space.
9. Is it expected that center points will be in the center of the experimental design, or will those points be at a manufacturing set point that may be off-center in the experimental space? What is the effect to the properties of the experimental design if the points are off-center?

3.6 Nominating a Parameter as Critical

The assessment of critical quality attributes (CQAs) and the control of critical process parameters (CPPs) that affect these attributes is an important component of the overall control strategy for drug substance and drug product manufacturing. There are many different approaches for assessing process parameter criticality, and although the determination of criticality is not primarily a statistics function, statistics can play a part in helping to identify CPPs.

One particular challenge involves assessing when a relationship between a process parameter and a CQA represents a significant impact on that CQA. For example, Fig. 3.9 provides two statistically significant relationships between a CQA and a process parameter across the explored space. Both equations are statistically significant, however, it is clear that the blue equation has a practically more meaningful relationship than the green equation. The blue equation has a chance of producing product outside specification if operated within the range, whereas, the green equation does not. Assessing impact based solely on statistical significance (p-value) is not appropriate, because statistical significance does not take into

Fig. 3.9 Statistically significant equation comparison



account the strength of the relationship relative to the relevant quality requirements. Ignoring this fact can lead to the inclusion of relatively unimportant process parameters as critical elements of the control strategy. Including these unimportant process parameters as CPPs is undesirable as it effectively dilutes the focus on process parameters that are truly important for ensuring product quality. It can also place an unnecessary burden on manufacturing operations resulting in an increased cost.

An alternative two-step procedure is provided by Wang et al. (2016).

Step 1: Perform a process risk evaluation for each relevant CQA.

For each CQA, evaluate the data set responses across the investigated range without focusing on any single or particular parameter. A Z-score assessment is employed to determine how close the results are to the specification or targeted response for the attribute. The Z-score is calculated as

$$Z^* = \min\left(\frac{U - \bar{x}}{S}, \frac{\bar{x} - L}{S}\right)$$

(3.1)

\bar{x} = average of data across the explored space,
 S = standard deviation of data across the explored space,
 U = upper target or specification, and
 L = lower target or specification.

It is not necessary to have both an upper and lower limit to calculate Z^* . In the case of a one-sided specification, Z^* is simply the single value corresponding to the specification of interest.

Figure 3.10 provides an illustration for a one-sided Z^* . For this case, if “specification #1” is the upper specification limit, then the Z^* for this data is expected to be small, indicating that the data is at risk of being greater than the upper specification limit at some operating conditions in the explored space. Alternatively, for “specification #2,” the data is far from the specification limit indicating that there is no risk of being beyond the specification for a well-controlled process operating within the explored space.

In Fig. 3.10 cutoff values for Z^* of 2 and 6 were selected as decision points in the analysis. Say that potency is a CQA and the analysis of the DoE data found a significant relationship between potency and milling speed, roll force, and compression force. The prediction equation is

$$\text{Potency} = 98 + 2.5 \times \text{Millspeed} - 0.5 \times \text{Roll force} + 1.5 \times \text{Compression force} \quad (3.2)$$

- If Z^* is less than 2, then all of the parameters in the significant model are CPPs. For the potency example, mill speed, roll force, and compression force are all significant.

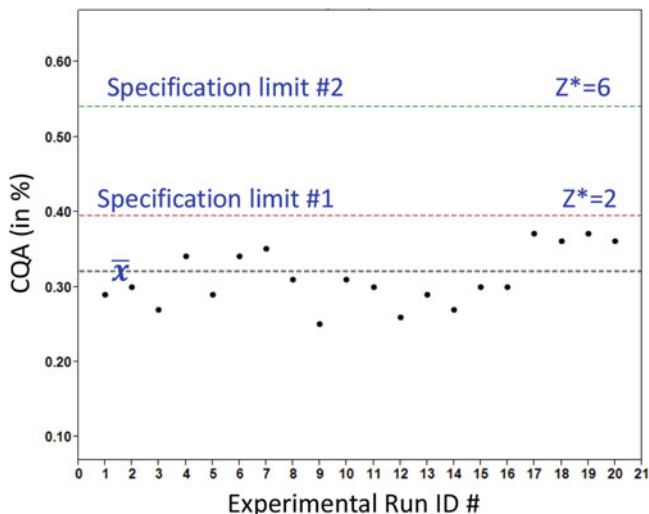


Fig. 3.10 Potential Z-score cutoff values for determining significance

- If Z^* is greater than 6, then none of the parameters in (3.2) are CPPs. For the potency example, the response is performing similar to the green line in Fig. 3.9. Hence, there is no risk across the explored region and no CPPs.
- If Z^* is between 2 and 6, then go to Step 2.

Step 2: Assess the criticality of individual parameters as necessary.

This step is performed if Z^* is between 2 and 6. Here, the fitted statistical model is utilized to quantify individual parameter effects against the proposed specification. This is termed the 20% rule for this application. For the potency CQA, if the specification is 95–105%, then the specification width is 10%. This specification width is multiplied by 20% to yield 2% for this example. As the mill speed coefficient in (3.2) is greater than 2%, the mill speed is a practically meaningful CPP. The other two parameters, roll force and compression force, are not CPPs.

3.7 Determining a Region of Goodness

A significant outcome of the DoE is determination of a region of goodness to operate the process. For example, two responses, impurity 1 and impurity 2, were studied in a two-factor (B, C) full factorial design with replicated center points. From the experimental data, models were developed and summarized in a contour plot (see Fig. 3.11).

It is desired to minimize impurity. The arrows in Fig. 3.11 show the direction of this minimization, or the so-called direction of goodness. For impurity 1, the

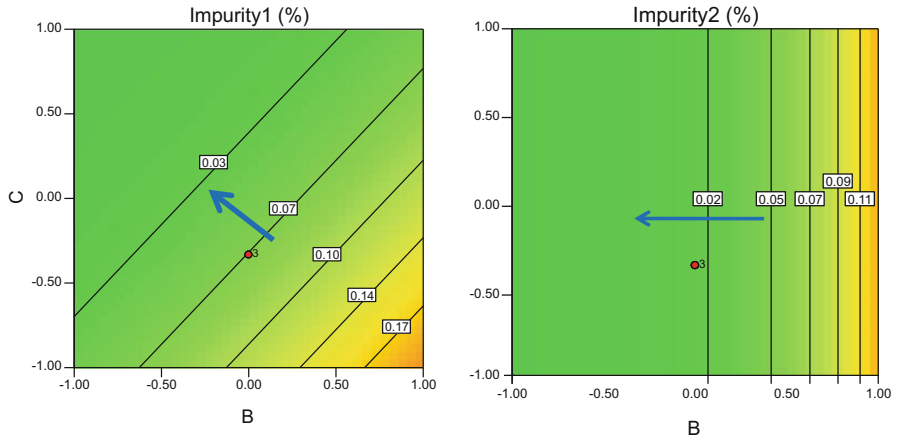


Fig. 3.11 Contour plots of impurity 1 and 2

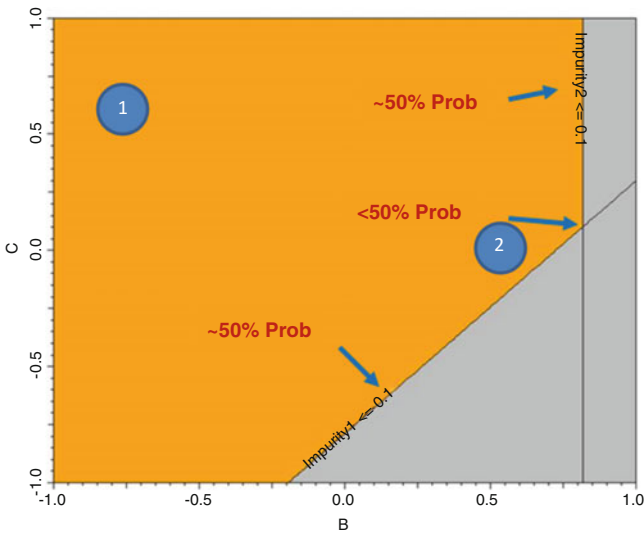


Fig. 3.12 Region of goodness defined by pass/fail mindset

combination of factor B at the low level with factor C at the high level is the best combination to minimize impurity 1. For impurity 2, factor B at the low level is the best, and factor C has no impact.

Assume the specification for each impurity is 0.10%. Examination of Fig. 3.11 shows the region and boundary where each impurity is less than 0.1%. It is common to summarize this information by providing pass (orange) and fail (gray) regions as shown in Fig. 3.12. The orange region represents an area where both 0.10%

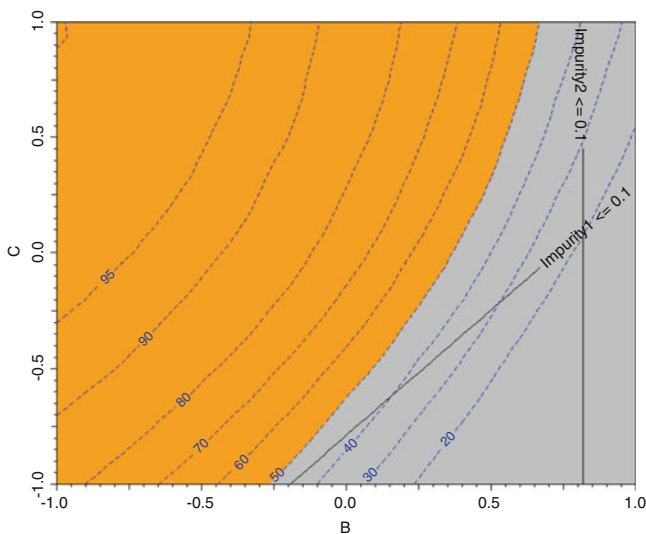


Fig. 3.13 Probability of passing considering the relationship between responses

specification limits are simultaneously met. However, across this region, there are a range of success probabilities. That is, based on Fig. 3.11, it is expected that there will be a higher probability to pass the specification of 0.1% in the upper left quadrant of the region than in the rest of the space. The predicted value at point #1 in Fig. 3.12 for impurity 1 and 2 is 0.01. The predicted value at point #2 for impurity 2 is close to 0.10. It makes sense that although the orange region will produce product that passes specifications, the probability of passing a specification of 0.10% must be greater at point #1 than point #2. In general, at the impurity limit of 0.10% it is expected that the probability of passing the specification is about 50% and in the region of overlapping requirements there is less than a 50% probability of passing. An improvement to examining the pass/fail plot in Fig. 3.12 is to assess the probability of passing and make decisions based on this probability.

Peterson et al. (2009) proposed an approach to calculate the probability of simultaneously passing all relevant specifications using seemingly unrelated regression (SUR) and has also, although unpublished, outlined a parametric bootstrap simulation approach to calculating this probability across the space of interest.

Using a bootstrap method, the probability to simultaneously pass multiple specifications is provided in Fig. 3.13. The levels on these contours now show the probability of passing while taking into account the predictive distribution, not simply the average prediction.

This more descriptive probability can be used to make better informed decisions about the process. That is not to say that the method is perfect or cannot be improved. A Bayesian approach to specify a range of parameter and variability estimates might help stabilize predictions in some cases.

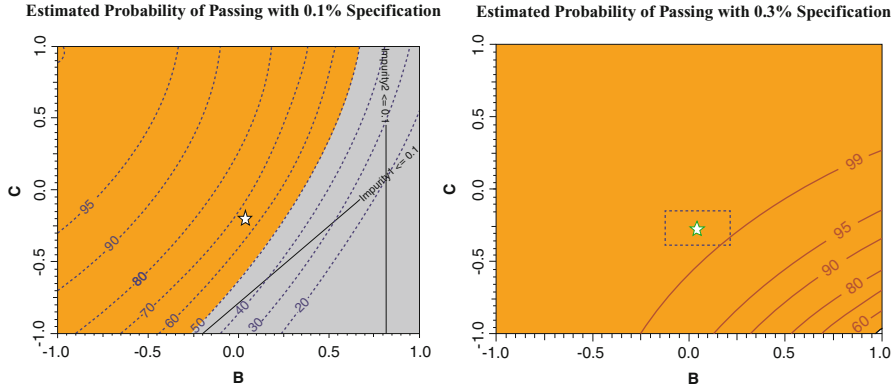


Fig. 3.14 Example using the probability of passing specification

Continuing the earlier example of the two impurities, assume the process was initially set to operate at the point indicated by a star in Fig. 3.14. Following the probability calculation, it is determined that there is a 70% chance of passing both specifications simultaneously. There could be several solutions that may improve or remediate this probability, and knowledge of the estimated probability is a step in proposing the process.

- It may be that the process is improved by downstream processing. So although there may be a cost associated with a 70% probability of passing at this stage, the probability will be improved in the future.
- It may be that the experiment was performed sub-scale. There is a known improvement to the probability when performed at scale.
- The set points of parameters B and C may need to be adjusted to improve the probability of passing.
- The initial specifications on the impurities of 0.1% may need to be increased. The effect of increasing the specification to 0.3% is provided in the right-hand side of Fig. 3.15.
- Finally, true process variability may be greater or less than the magnitude realized in the experimental data. The simulation can be performed again with the more appropriate error structure.

3.8 Process Capability and Process Robustness

The process capability index, abbreviated broadly as C_{pk} , is a widely used summary statistic describing the ability of a process to produce output within specification limits. The index plays a prominent role in PV-3, and is discussed in Chap. 5 of this book. An assessment of capability is also useful in PV-1. Obtaining a meaningful estimate of process capability early in a product’s life cycle is difficult because

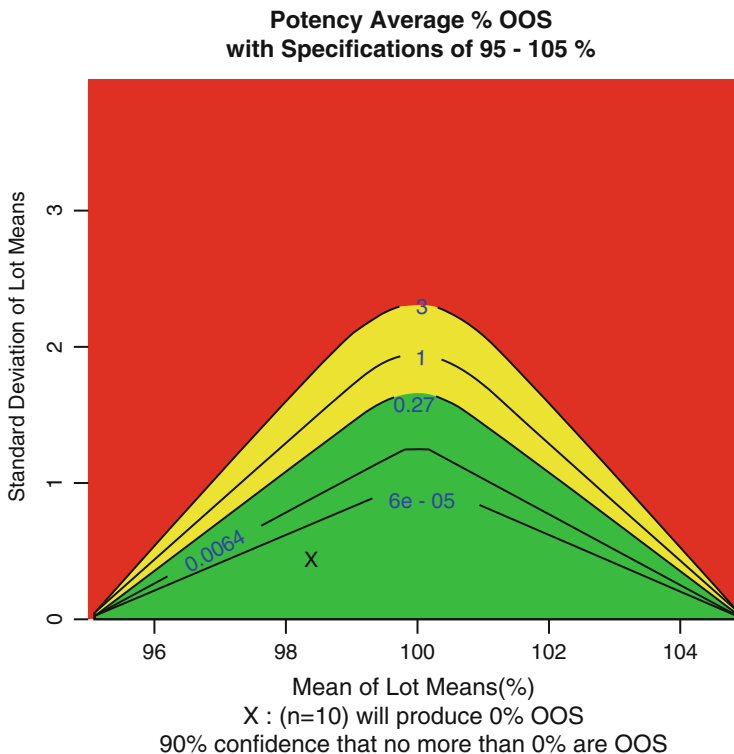


Fig. 3.15 Robustness contour

many lots are needed to provide a meaningful capability index. For these indices to have predictive meaning, the process must have demonstrated adequate statistical control prior to their calculations. This effort requires at least 25 lots.

Within the last decade, the concept and industrial practices of QbD have led to greater process understanding in R&D leading to increased knowledge of process capability that is not specifically captured by the small number of lots manufactured early in a product's life cycle. Although there may only be a couple produced lots, the scientific understanding, fundamental knowledge, and development experience is substantial and provides an opportunity to assess process capability. A proposal for a robustness calculation, meant to distinguish an early estimate of control and capability developed within a QbD framework from the rigorous assessment of control and capability implied by a capability statistic was proposed by Vukovinsky et al. (2017). This contour-based tool calculates the percent out-of-specification (% OOS), based on the mean, standard deviation, and specification of an attribute. The contours provide a clear visualization of the ability of the process to meet the specification, making it a useful tool for products in development as well as new and marketed products. Figure 3.15 provides an example contour plot for potency,

based on a sample size of only 10 lots resulting in a mean potency of about 98.4%, a standard deviation between lots of 0.5%, and a specification of 95–105%.

These %OOS contour regions use the following coloring scheme:

- Green: less than a 0.27% OOS rate (good performance).
- Yellow: greater than or equal to 0.27% OOS rate and less than 3% (further discussion required).
- Red: greater than 3% OOS rate (requires improvement).

The OOS% contour levels of 0.27, 0.006, and $6e-5$ displayed on the plot are approximately related to C_{pk} values of 1, 1.33, and 1.67, respectively. Associating the green contour with 0.27% OOS implies a minimum C_{pk} of one in transition to manufacturing.

Once a process robustness contour plot is constructed, the relative location of the present process within the colored contour is examined to assess the product performance. In Fig. 3.15 the “X” represents the location of the attribute of interest. The ultimate goal for the product should be emphasized more than the color zone containing the “X”. The relative location provides information concerning the sensitivity of the attribute to change in the sample mean and sample standard deviation and can guide the search for potential improvements in product performance or the need to modify data-driven specifications. As with any summary statistic, there is variability in the %OOS estimates. This variability is described in Fig. 3.15 footnote as an upper confidence estimate on the %OOS. This estimated upper bound is based on the data, and can be quite wide for a small sample size. The fundamental, scientific, and experimental understanding of the process gained through the design process along with the calculated bound should be considered in process decisions.

Once constructed, the contour plots should support an active discussion about the product performance amongst a cross-functional team. In general, data external to the summarized lot data, estimates of variability components from methods and processes, or knowledge from modeling efforts on similar products can be used to assess potential future process behavior and expectations. All of these discussions can use the robustness contour as a foundation.

3.9 Control Strategy Implementation

ICH Q8(R2) documents a “Minimal Approach” to Control Strategy which is contrasted with the “Enhanced, Quality by Design Approach.” Here, criticality of parameters is determined following scientific investigation through the QbD process.

The concept of criticality can be used to describe any material attribute, characteristic of a drug substance, component, raw material, drug product or device, process attribute, parameter, condition, or factor in the manufacture of a drug

product. The assignment of attributes or parameters as critical or non-critical is an important outcome of the development process and provides the foundation for the control strategy. Critical Process Parameters (CPPs), the relationship between Critical Quality Attributes (CQAs) and Critical Process Parameters (CPPs), and the ranges for CPPs (PAR and NOR) are documented as a control plan. The control strategy provides a plan to prevent operating in regions of limited process knowledge or those that are known to cause product failure.

Underlying the criticality assignment process is the concept that the primary assessment and designation of criticality should be made relative to the impact that quality attributes or process parameters have on the safety, efficacy, and quality of the product. The material in Sect. 3.6 provides one option to determine criticality. Once criticality is determined, a control strategy that focuses on the most appropriate control points and methods is developed.

ICH Q10 defines a control strategy as

a planned set of controls derived from current product and process understanding that assures process performance and product quality. The controls can include parameters and attributes related to drug substance and drug product materials and components, facility and equipment operating conditions, in process controls, finished product specifications and the associated methods and frequency of monitoring and control.

QbD also introduced the concept of a traditional versus a dynamic control strategy. In a traditional control strategy, any variability in process inputs (such as quality of the feed material or raw materials) results in variability in the quality of the product because the manufacturing controls are fixed. In a dynamic control strategy, the manufacturing controls can be altered (within the region of goodness) to remove or reduce the variability caused by process inputs.

A holistic control strategy mitigates any risk from a single unit operation. The control strategy includes the process definition, control limits of process parameters, and release limits, amongst other considerations. It is important in determining the manufacturing process that specifications be set appropriately (see Chap. 7).

A statistically related example illustrates the translation from an equation derived from a DoE to a control strategy. Here, dissolution (*Diss*) is found to be a function of API particle size (*API*), magnesium stearate surface area (*MgSt*), lubrication time (*LubT*), and compression force (*Crush F*).

$$\begin{aligned}
 Diss = 108.9 - 11.96 \times API - 7.556 \times 10^{-5} \times MgSt - 0.1849 \times LubT \\
 - 3.783 \times 10^{-2} \times CrushF - 2.557 \times 10^{-5} \times MgSt \times LubT.
 \end{aligned}
 \tag{3.3}$$

Assume these parameters are both statistically significant and their effect on dissolution is practically meaningful. Equation (3.3) describes the current understanding and can be used to define meaningful limits on the parameter specifications and process controls. Using this information, quality is built into the process by managing the process inputs. Although there may not be direct control of *Diss*, it might be controlled upstream by one of the variables on the right of Eq. (3.3).

- API: To control dissolution, it is important to maintain the D90 API particle size within a certain range. Here, the predicted equation is used to determine the range of 5–30 μm and the high shear wet milling equipment is set to achieve a value within this range.
- MgSt: The surface area of the magnesium stearate (lubricant) particles is controlled to ensure dissolution. This assurance is performed upon receipt of the MgSt from the supplier.
- LubT: Lubrication time is controlled between 1 and 8 min via automated equipment.
- CrushF: Tablet hardness is controlled by the crushing force at the time of compression to a targeted amount and within an acceptable range.

All decisions concerning the CQA are documented within a control plan.

3.10 Preparation for Stage PV-2

After the control strategy has been defined and the product and process ranges are established, product and process qualification (PV-2) is performed to demonstrate that the process will deliver a product of acceptable quality if operated within the region of goodness. This will also confirm whether the small and/or pilot-scale systems used to establish the region of goodness can accurately model the performance of the manufacturing scale process. PV-2 is really a confirmation of the understanding and control strategy. Following PV-2, the regulatory filing is compiled, which includes the acceptable ranges for all critical operating parameters that define the manufacturing process.

References

- ASTM E2363 (2014) Standard Terminology Relating to Process Analytical Technology in the Pharmaceutical Industry. ASTM International, West Conshohocken, PA
- Box GEP, Hunter JS, Hunter WG (2005) Statistics for experimenters: design, innovation, and discovery, 2nd edn. Wiley, NY
- Code of Federal Regulations (CFR), Title 21, Food and Drugs Administration (FDA), Part 210. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=210>
- Deming WE (2000) Out of the crisis, Reprint edition. The MIT Press
- FDA. CDER (2011) Process validation: general principles and practices, guidance for industry
- Garcia T, McCurdy V, Watson T, Ende M, Butterell P, Vukovinsky K, Cheuh A, Coffman J, Cooper S, Schuemmelfeder B (2012) Verification of design space developed at subscale. *J Pharm Innov* 7:13–18
- Glodek M, Liebowitz S, McCarthy R, McNally G, Oksanen C, Schultz T, Sundararajan M, Vorkapich R, Vukovinsky K, Watts C (2006) Process robustness – a PQRI white paper. *Pharm Eng* 26:6
- ICH (1999) Q6A Specifications: test procedures and acceptance criteria for new drug substances and new drug products: chemical substances

- ICH (2005) Q9 Quality risk management
- ICH (2008) Q10 Pharmaceutical quality system
- ICH (2009) Q8 (R2) Pharmaceutical development
- ICH (2012) Q11 Development and manufacture of drug substances (chemical entities and biotechnological/biological entities)
- Juran JM (1992) *Juran on quality by design: the new steps for planning quality into goods and services*. Free Press
- Montgomery DC (2012) *Design and analysis of experiments*, 8th edn. Wiley, NY
- Morris M (2011) *Design of experiments: an introduction based on linear models*. Chapman & Hall/CRC
- Peterson JJ (2004) A posterior predictive approach to multiple response surface optimization. *J Qual Technol* 36:139–153
- Peterson JJ (2010) What your ICH Q8 design space needs: a multivariate predictive distribution. *Pharm Manuf* 8:23–28
- Peterson JJ, Miró-Quesada G, del Castillo E (2009) A Bayesian reliability approach to multiple response optimization with seemingly unrelated regression models. *Qual Technol Quant Manage* 6(4):353–369
- Stockdale GW, Cheng A (2009) Finding design space and a reliable operating region using a multivariate Bayesian approach with experimental design. *Qual Technol Quant Manage* 6:391–408
- Vukovinsky KE, Altan S, Bergum J, Pfahler L, Senderak E, Sethuraman S (2010a) Statistical considerations in design space development, I of III. *Pharm Technol* 34(7)
- Vukovinsky KE, Altan S, Bergum J, Pfahler L, Senderak E, Sethuraman S (2010b) Statistical considerations in design space development, II of III. *Pharm Technol* 34(8)
- Vukovinsky KE, Altan S, Bergum J, Pfahler L, Senderak E, Sethuraman S (2010c) Statistical considerations in design space development, III of III. *Pharm Technol* 34(9)
- Vukovinsky KE, Li F, Hertz D (2017) Estimating process capability in development and for low volume manufacturing: process robustness assessment using %OOS contour plots. *Pharm Eng* Jan/Feb 2017
- Wang K, Ide N, Dirat O, Subashi A, Thomson N, Vukovinsky K, Watson T (2016) Statistical tools to aid in the assessment of critical process parameters. *Pharm Technol* 40(3):36–44