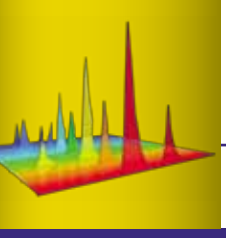


Understanding and Modeling Product and Process Variation

Variation understanding and modeling is a core component of modern drug development.

Understanding and accounting for product and process variation is one of the most fundamental aspects of product development. It is a mandatory activity required by regulatory agencies and a key consideration when developing new drug products and substances. Variation types and sources must be understood and accounted for when developing new products and processes (1). The goal of product and process development is to account for 100% of the variation in any product or process. In general, variation should be quantified and accounted for in three groups: the variation in each Y response, the variation transmitted by the transfer function of each X factor on every Y response, and the transmitted variation of each X factor at defined set points. Failure to understand and control sources and types of variation results in high out-of-specification events and agency concerns.

SOURCES OF VARIATION

Sources of variation include:

- Materials and excipients
- Equipment
- Process methods and/or sequence
- Process set points and allowable ranges
- Environmental conditions
- Measurement methods, calibration, standards, and probes
- Analysts and operators
- Software/control systems

TYPES OF VARIATION

Types of variation include:

- Within/between batch variation
- Within/between fill variation
- Positional or location variation
- Instant-in-time and period-of-time variation

FDA's *Guidance for Industry Process Validation: General Principles and Practices* states (2), "A successful validation program depends upon information and knowledge from product and process development. This knowledge and understanding is the basis for establishing an approach to control of the manufacturing process that results in products with the desired quality attributes. Manufacturers should:

- Understand the sources of variation
- Detect the presence and degree of variation
- Understand the impact of variation on the process and ultimately on product attributes
- Control the variation in a manner commensurate with the risk it represents to the process and product

Each manufacturer should judge whether it has gained sufficient understanding to provide a high degree of assurance in its manufacturing process to justify commercial distribution of the product. Focusing exclusively on qualification efforts without also understanding the manufacturing process and associated variations may not lead to adequate assurance of quality."

VISUALIZATION OF VARIATION

If the variation is positional such as concentration on a plate or shelf position in an oven, a visual representation typically is best. All variation should be examined both visually and analytically (quantitatively). The primary limitation of variation visualization is not quantitative and needs to be augmented with other quantitative measures to achieve a complete understanding. **Figure 1** is a 96-well plate example with a controlled reference that is used to examine coating consistency. Contour plots, 3D surface plots, and bubble plots are common graphical representations.

Thomas A. Little, PhD, is president of Thomas A. Little Consulting, drlittle@dr-tom.com.

Image: PASIEKA/Science Photo Library/Getty. Figures are courtesy of authors.

Figure 1: Contour plots, 3D surface plots, and bubble plots.

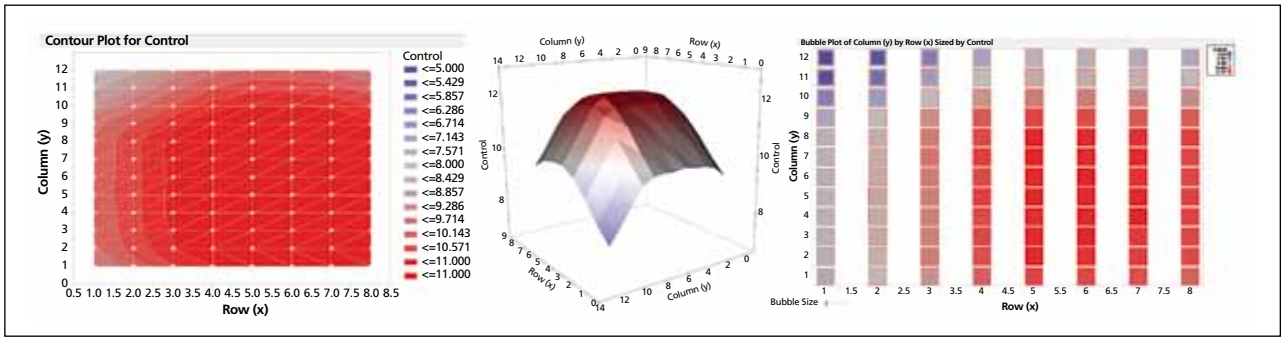
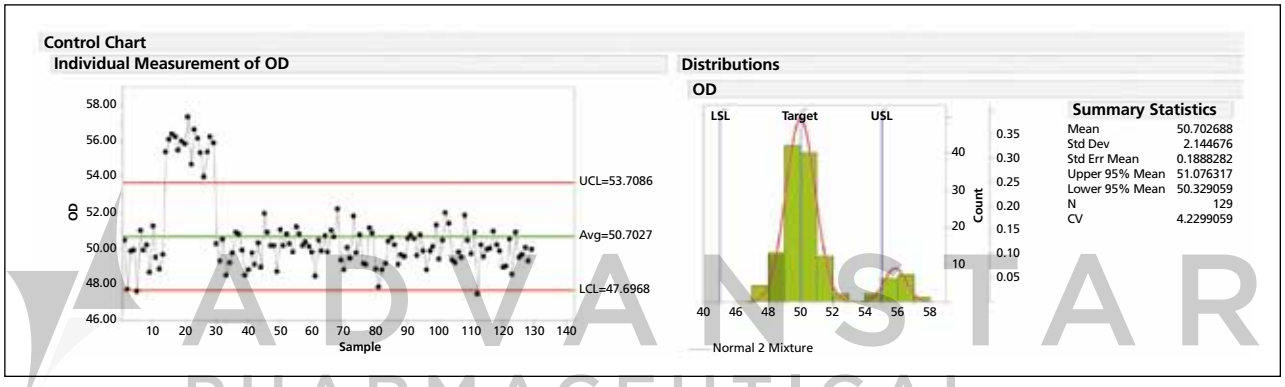


Figure 2: Control chart and distribution analysis.



UNIVARIATE

Single parameters, either factors or responses, should be visualized and quantitated using control charts for time-series data and histograms for continuous variables. Distribution curve fitting should be used to better represent the population. Core statistics such as the mean, standard deviation, and coefficient of variation should be computed for the sample data, and confidence intervals (CI) should be computed for both the mean and the standard deviation. To ensure the statistics are representative of the population, a valid sampling method and sufficient sample size must be used prior to generating the statistics and CIs. A histogram alone is incomplete and only communicates one-half of the information; the control chart communicates the balance (see **Figure 2**).

ANOVA

Both one-way analysis of variance (ANOVA) (3) and bivariate regression analysis provide a more complete analysis of the influence of a single factor on a single response. ANOVA breaks down the variation into within-group and between-group variation based on sums of squares (SS) (see **Figure 3**). Sum of squares divided by total degrees of freedom (n-1) is the variance for each group. RSquared is the proportion of variation explained by the X factor. Root mean squared error (RMSE) is the variation unaccounted for by the ANOVA model. ANOVA model is often used for determination of intermediate precision and within and between batch variation, as follows:

Stdev=Sqrt (variance within group + variance between group)

When product and processes are multiple factor, one-way ANOVA and bivariate analysis are inappropriate and falsely indicate the variation.

BIVARIATE

Bivariate regression analysis is similar to the ANOVA (3); however, as the X factor is continuous it adds one more dimension to the analysis (see **Figure 4**). RSquare is the amount of variation explained by the linear model, RMSE is the residual error. To fully understand the transmitted variation, simulation can now enter in with all the three components of variation: RMSE, model, and the variation in X. The bivariate fit can be used to estimate the mean at any setting of X; the simulation can be used to estimate individual units at any setting of X. From a drug development perspective, the latter is

Figure 3: ANOVA and variation breakdown.

Oneway Anova

Summary of Fit

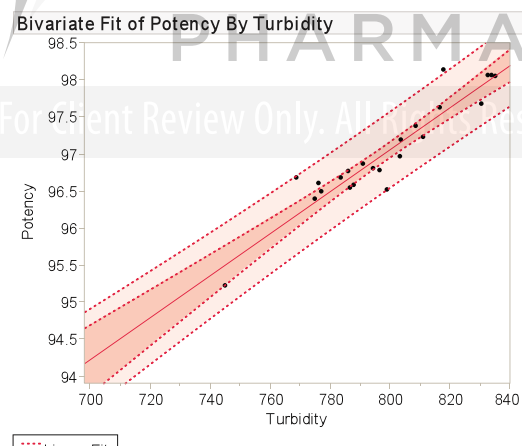
Rsquare	0.69634
Adj Rsquare	0.635608
Root Mean Square Error	1.200388
Mean of Response	99.33085
Observations (or Sum Wgts)	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Analyst	4	66.085695	16.5214	11.4658	<.0001*
Error	20	28.818609	1.4409		
C. Total	24	94.904304			

Source	SS	Total DF	Variance	Stdev	% of Variation
Analyst	66.086	24	2.754	1.659	069.63%
Error	28.820	24	1.201	1.096	030.37%
C. Total	94.904	24	3.954	1.989	100.00%

Figure 4: Bivariate variation modeling.



Linear Fit

Potency = 74.400025 + 0.0283328*Turbidity

Summary of Fit

RSquare	0.884953
RSquare Adj	0.879475
Root Mean Square Error	0.241659
Mean of Response	97.01647
Observations (or Sum Wgts)	23

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	9.433515	9.43352	161.5348
Error	21	1.226385	0.05840	Prob > F
C. Total	22	10.659900		<.0001*

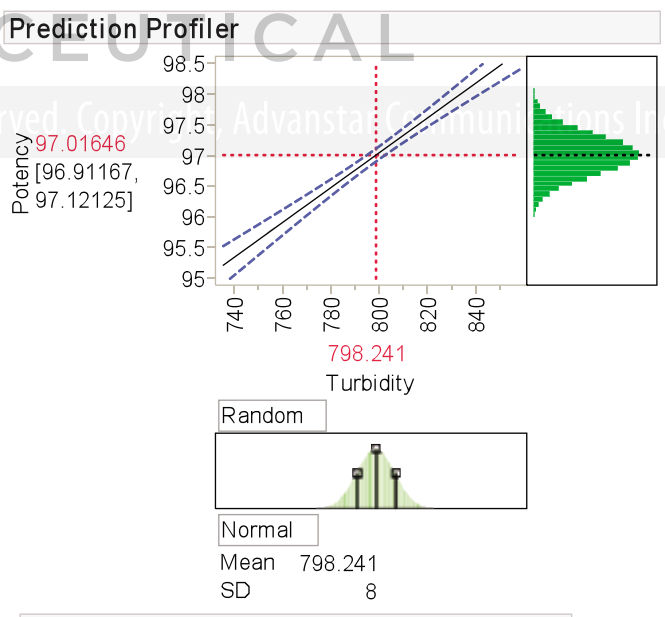
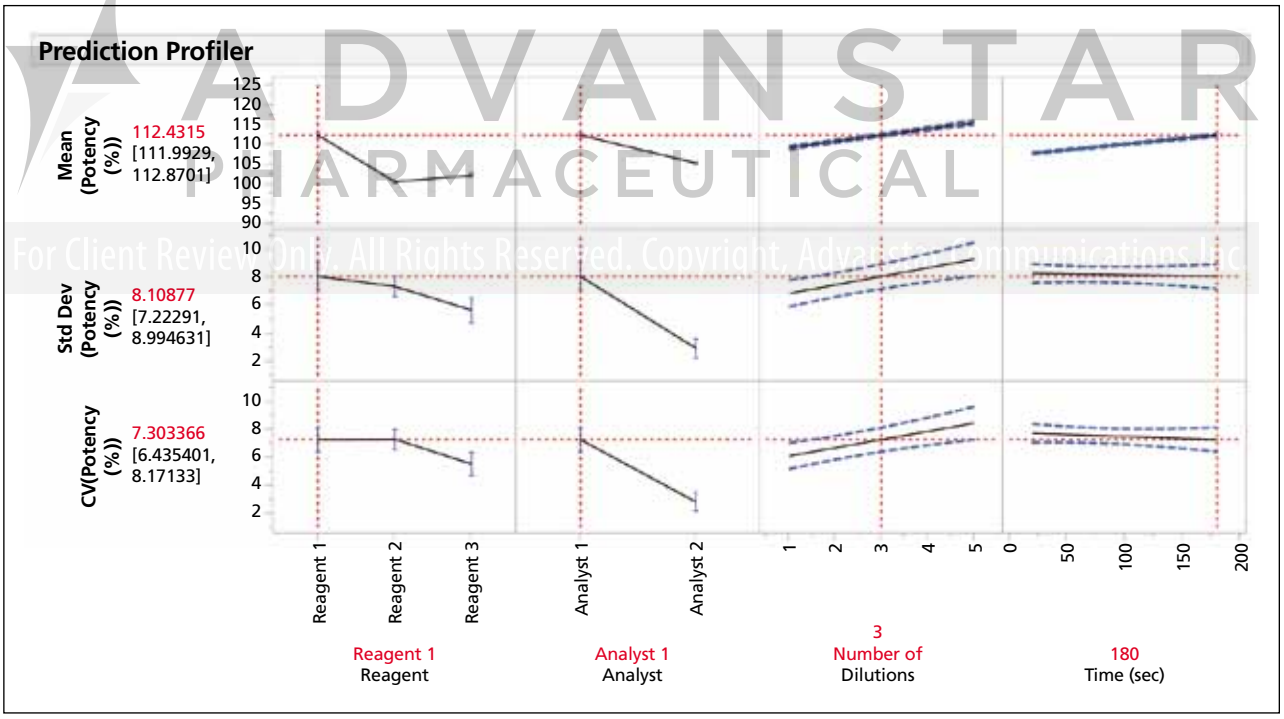


Figure 5: Components of variation analysis methods.

Method	Between Group Variation (Mean)	Within Group Variation (Stdev)	Sample Size per Group	Factor Types Allowed	Approach to Population Statistics
Expected mean Square	Full Breakdown	Does not breakdown the within group variation	Equal sample size per group	Categorial only	Predictive
Restricted maximum likelihood (USP 1033, Reference 4)	Full Breakdown	Does not breakdown the within group variation	Allows for unequal sample size per group	Categorial only	Predictive
Partition of Variation (POV, 2013, Reference 5)	Full Breakdown	Full Breakdown	Allows for unequal sample size per group	Categorical and Continuous	Descriptive

$$\sqrt{S^2_{\text{Process}} + S^2_{\text{Materials}} + S^2_{\text{Measurement Method}} + S^2_{\text{Stability}}}$$

Figure 6: Profiler with mean, Std Dev, and CV.



much more important as one does not release the mean performance, one releases the unit and or batch.

COMPONENTS OF VARIATION

Variance components analysis allows for a more complete breakdown of the variation and provides both a visual and analytical examination of the variation. There are several different methods for components of variation (COV) analysis and they have their limitations. COV is one of the most important analytical tools to parse and account for multiple factor variation. Generally, the partition of variation is the most complete and flexible method of performing variance components analysis (see **Figure 5**) (4, 5).

VARIATION MODELING

IN DESIGNED EXPERIMENTS

Factors may influence both the mean and the standard deviation. Factor selection prior to experimentation needs to be risk based (6). If repeated measures or replicates are used in the study design, the mean, standard deviation and coefficient of variation (CV) can be modeled to determine how factors can be adjusted to hit the target at the lowest transmitted variation. Summary tables are generated from the replicated measures and the mean, standard deviation, and CV are typically used for process modeling. A model is then fit to the statistics rather than to the individual values.

During analysis, care needs to be exercised to properly account for the sample size/sample frequency in the model. Once the sample size is accounted for correctly in the model definition the statistical tests and confidence intervals will also be correctly reflected in the analysis. Design space may be generated for both the mean response as well as the standard deviation (see **Figure 6**).

TRANSMITTED VARIATION IN SETTING RANGES AND LIMITS

Understanding how variation in X transmits onto Y is essential in drug development. Generally when setting normal operating ranges (NORs) for all X parameters, they are explored at 3K sigma or 99.75% of the normal range. If they are deemed safe and low risk, wider proven acceptable ranges (PARs) are used to determine how far the X parameter can be opened up and still assure the product/substance will meet all acceptance criteria. Often 4, 5, or even 6 K sigma limits are used to determine their influence in the design space and potential failure rates. Defect and failure rates in parts per million and capacilitiy indices such as C_{pk} are also measures of capability that are used in evaluation of PAR ranges. Desirability, functions can also be used to determine safe operating ranges rather than K sigma if using profilers.

CONTROLLING AND REDUCING VARIATION

There are two primary methods for controlling product, process, and material variation: test and release and measurement and adjust within an approved design space following a documented control procedure. Feed forward and feedback control loops are excellent ways of maintaining critical quality attributes (CQAs) at their intended target. Control loops can either be in-process or between-batch runs depending on how they were developed and submitted to the agencies.

Controls should be equality considered for all three sources of variation: process, analytical methods, and materials. It will not be possible to adjust a process when the analytical method is not stable and in control. Proper sampling plans with appropriate sampling methods and sample sizes are used to control the product and process.

CONCLUSION

Variation understanding and modeling is a core component of modern drug development. It is a required element of Stage I validation and increases product and process knowledge and reduces risk. The goal is to understand 100% of the drivers influencing product variation, acceptance criteria, and CQAs. Once the understanding of how factors influence the product variation is completed, a clear strategy on how to control the types and sources of variation is essential to gaining the benefits from variation understanding. Control plans with appropriate adjustment and acceptance testing protocols will deliver the benefits of improved understanding of the variation in key analytical methods, materials, product attributes, and process parameters.

REFERENCES

1. ICH, *Q8(R2) Pharmaceutical Development*, 2009.
2. FDA, *Guidance for Industry, Process Validation: General Principles and Practices* (CDER, Rockville, Md, 2011).
3. NIST, Section 3.2.3.1, *Engineering Statistics Handbook* (NIST, 2012).
4. USP, <1033> Biological Assay Validation (USP, 2010).
5. Thomas A. Little Consulting, Tools and Templates, POV, www.qualitybydesignconsulting.com/toolsandtemplates/index.html, 2013.
6. ICH, *Q9 Quality Risk Management*, 2006. ◆