BioPharm INTERNATIONAL

July 2008

The Science & Business of Biopharmaceuticals

Coming to a Biotech Near You: Quality by Design

Part 2: Design Space in Development and Manufacturing

In a time of necessary speed and robustness, Quality by Design and Design Space are useful tools for companies to use

Regulatory support...economic pressures... the complexity of protein-based products. As we detailed in the previous installment in this series, (*BioPharm International*, May 2008) all of those trends are coming together to bring Quality by Design (QbD) to biotech companies. Because QbD envisions designing-in product and performance characteristics from the first rather than deriving them through testing after the fact, it opens the way to a risk-based approach to quality. The key to ensuring an acceptably low risk of failing to achieve the desired clinical attributes lies in determining the Design Space, defined in ICH Q8 as "the multi-dimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality."[1]

But as we also pointed out previously, one of the biggest obstacles to adopting QbD is the industry's inconsistent understanding and application of the tools and methods that are essential for accurately determining Design Space. In this article, we show how those tools are typically used in describing the relationships between the manufacturing process inputs and the critical quality attributes (CQAs). Used correctly, these tools help make it possible to map the multidimensional Design Space in which quality is ensured, and to carve out an efficient operating space in the center of that Design Space where process robustness resides.

Jason J. Kamm is managing consultant and **Conrad J. Heilman, Jr., PhD,** is senior vice president, both at Tunnell Consulting, King of Prussia, PA, 610.337.0820, kamm@tunnellconsulting.com heilman@tunnellconsulting.com.

THE OBJECTIVE: PROCESS UNDERSTANDING

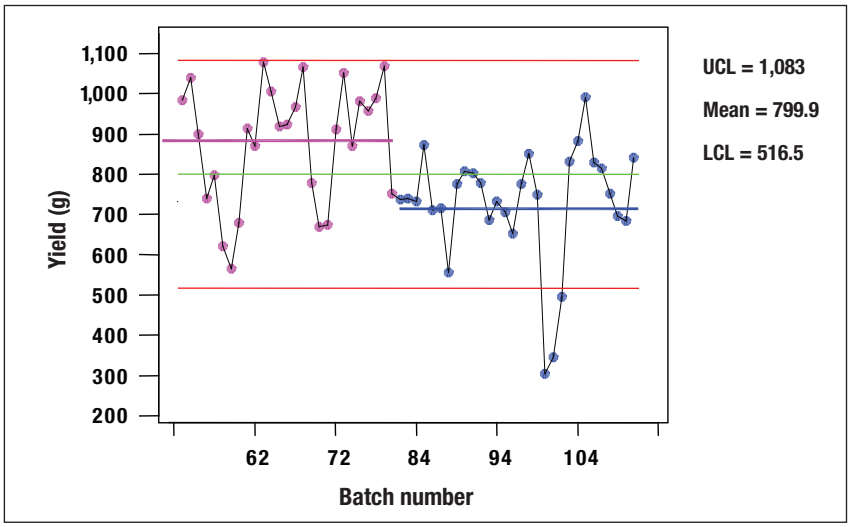
The CQAs are the desired outputs of the manufacturing process. In mapping out Design Space, the goal is to understand the relative impact on CQAs of input variables—process steps, process parameters, and raw materials. For example, in a large-scale mammalian cell culture bioreactor, there are any number of potential CQAs, which could include buffer components, media components, supplements (e.g., nutrients), dissolved gases, etc. One surrogate measure for these CQAs in a biologic operation is yield, and the case study detailed below focuses on this surrogate measure. The Design Space encompasses numerous permutations of the input variables in relation to each other that still produce the desired outputs. In other words, the goal is to achieve a profound and comprehensive understanding of the process and then carefully monitor and control those critical elements.

In a traditional approach to process understanding, every parameter could potentially receive the same degree of scrutiny. However, given all the possible permutations of process steps, process parameters, and raw material components, such an approach is highly inefficient and sometimes impossible to achieve. Although manufacturers using one-factor-at-a-time analysis are able to produce an in-specification product by locking down individual process parameters, such success is usually short-lived. Differences in raw material batches and drift in other parameters can soon bring new problems. By contrast, QbD's risk-based approach seeks to determine the critical parameters and their combinations and control for them in a flexible manner in the Design Space.

APPLYING THE STATISTICAL TOOLS TO HISTORICAL DATA

The first step in achieving process understanding is to make sure that you have gathered in one place all of the historical data about the development of the product. Ideally, a database and all development reports to date will already exist. But if the data are lacking or spotty, it's necessary to compile

Figure 1. A control chart that plots batch-to-batch variability in yield.



the data before trying to map the Design Space. Once all of the historical data is on hand, you can then apply and sequence the appropriate tools.

It's important to acknowledge that the amount of data available in a development setting is often limited. Therefore, the ability to conduct statistical analysis is affected and a modified approach to analysis may be required. However, as a product moves into manufacturing and production runs mount, the amount of data available for statistical analysis increases and improves one's understanding of the process and one's ability to detect CQAs and understand how they interact. In the example below, we applied QbD tools to a manufacturing process to solve a problem with bulk substance yield at a company we will refer to as PharmaCo. The product being produced was a monoclonal antibody produced in a standard mammalian cell expression system in a bioreactor.

Run charts and control charts. First, the historical data on yield were plotted on a run chart in order to graphically represent the batch-to-batch variability and shift in output over time. The chart depicted a sequence of batches, at different scales, with the standardized yield in grams on the vertical (y) axis and the time series, represented by the batch numbers, on the horizontal (x) axis (as in Figure 1, a control chart that has the run chart as its basis). In the underlying run chart, we see a few indications of atypical behavior. First, there is a shift in the average yield from about 900 grams to about 700 grams. Something, unknown to this point, has happened in the process to cause this shift. We also see instances of outliers—or points that breached the lower control limit (LCL). These points also point to some special cause.

Control charts add the control limits of the process to support a wider range of analytical techniques. The control limits, as in Figure 1, are statistically derived based on the data. An analysis of the control chart showed that over time a significant shift in the process occurred that resulted in a drop in yield of 18%. Moreover, the chart shows a high number of out-of-control points, including three extreme outliers that occurred three consecutive times, which were traced to a single raw material lot. In other words, the process is not in "statistical control." Further, patterns that the chart uncovers—in this case, problems with a raw material lot, among other things—may offer some clues to the source of the out-of-control variations in the process.

PharmaCo's manufacturing personnel acknowledged the three extreme outliers in the data points but were unable to provide an explanation of them because those three batches had been produced in precisely the same fashion as the three preceding batches. This is not an uncommon situation, but it is telling that there are elements of the process that are not well understood. In the language of QbD, there is a lack of process understanding.

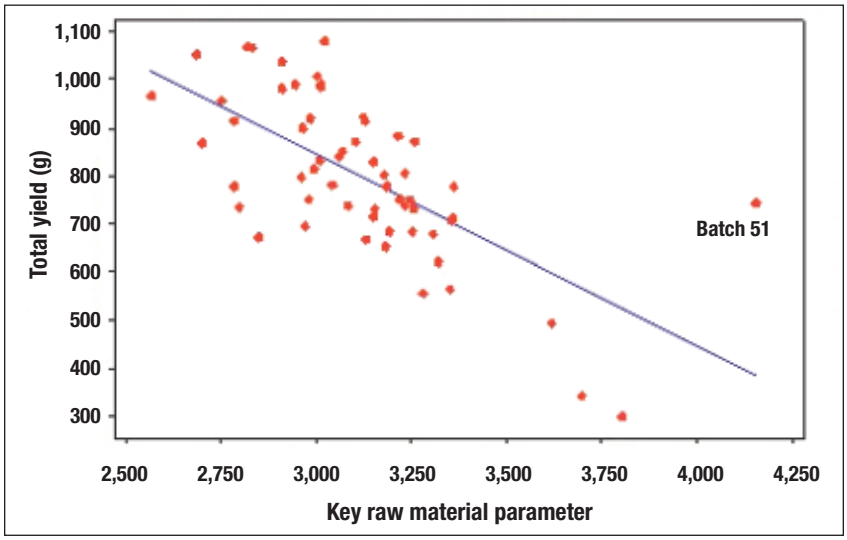
Bivariate analysis. To help understand the sources of the variation indicated by the control chart, the available data were subjected to bivariate analysis in which pairs of variables extracted from the historical data were examined and depicted on a scatter plot. In PharmaCo's case, bivariate analysis might plot yield against raw material lots or against process parameters to find correlations. Figure 2 plots the yield of the PharmaCo fermentation process against one of the key raw material parameters and shows that this process parameter correlates strongly with yield. At this point, however, this correlation does not imply cause and effect; further study is warranted. In fact, bivariate analysis conducted over a number of pairs of variables typically finds only weak correlations, not a strong correlation between two variables that might point to a magic bullet to solve the problem.

Of course, it would be possible to create a bivariate analysis for every conceivable pair of parameters, but it would be extremely inefficient to do so. Also, this method cannot uncover interaction effects, which could be important. Thus, bivariate analysis can be used as a prioritization step to reduce the number of variables for further study.

Multivariate analysis. Having identified several potentially critical variables through bivariate analysis, we applied multivariate analysis to look simultaneously at many variables that drive yields.

Multivariate analysis can be done through many commercially avail-

Figure 2. A scatter plot of a bivariate analysis comparing a key raw material parameter to yield.



able software packages. The goal is to create statistically valid models that describe the behavior of output variables (Ys) from input variables (Xs). A word of caution: these methods require some advanced statistical training to produce valid results.

In this case, multivariate analysis of historical data revealed three variables that accounted for 76% of the variability in yield. Further, the analysis showed that all three variables were related to raw materials.

Figure 3 shows the results of the multivariate analysis. A perfect model, which would account for 100% of the variability in yield, would show all of the data falling on the solid diagonal line.

BEYOND PURELY HISTORICAL DATA: DESIGN OF EXPERIMENTS

All of the previous steps focus on historical data. The next step—Design of Experiments (DoE)—uses the results of that historical analysis as a foundation on which to begin building experiments that will further refine understanding of the process. However, the three important variables uncovered by the multivariate analysis are not the end of the story. Before embarking on a DoE study, you should make sure that you haven't overlooked other variables that might be important.

For example, in biopharmaceutical manufacturing, variables that are typically important include mixing time, temperature, and the size of the bioreactor. Any such variable that you know to be important in a particular case should be included in the DoE study. In addition, you should solicit the input of personnel who are thoroughly familiar with the process and the product. For example, it might emerge from these conversations that certain variables known to be important were kept at the same setting for all batches. As a result, those potential variables might not have shown up as important in the previous analysis. Nevertheless, they should be included in the statistically designed experiments.

Once you've determined all of the relevant variables to be included, you then design experiments that are scaled correctly, balanced, orthogonal, and correctly sized. You must also have a statistically valid sampling plan to make sure you're making the right number of measurements in each batch. Then you can carefully execute the experimental trials.

The DoE study must include any parameters that were found to be important as a result of the bivariate and multivariate analyses. It should also include other variables that are found to be important in similar processes. Depending on the number of variables to be studied and the amount of variability one might expect to see in the response variables (CQAs), the DoE can be designed and sized.

For PharmaCo, this DoE might include batches representing different levels of raw material properties: RM1, RM2, and RM3. Since we are referring to only three variables, these could be arranged in a central composite design consisting of as few as 16 trials.

Using widely available statistical software, you can then perform multiple regression analysis on the results of the DoE to arrive at estimates of the effects of variables on the CQAs, their interactions, and their statistical significance. The model can then be refined manually based on such diagnostics as risk probability and others. The final regression model for yield should then have an excellent fit to the data—indicating that the model explains nearly all of the variation in the desired outcome. You can then produce a multidimensional picture of the Design Space (Figure 4).

This "contour profiler matrix plot" was created using the optimum set-

Figure 3. Multivariate analysis of three variables.

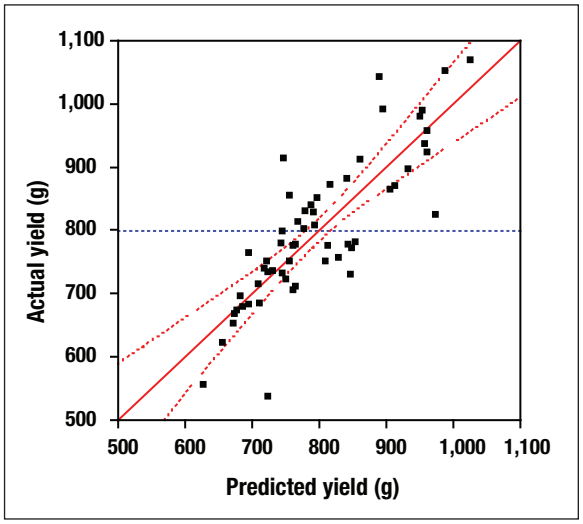
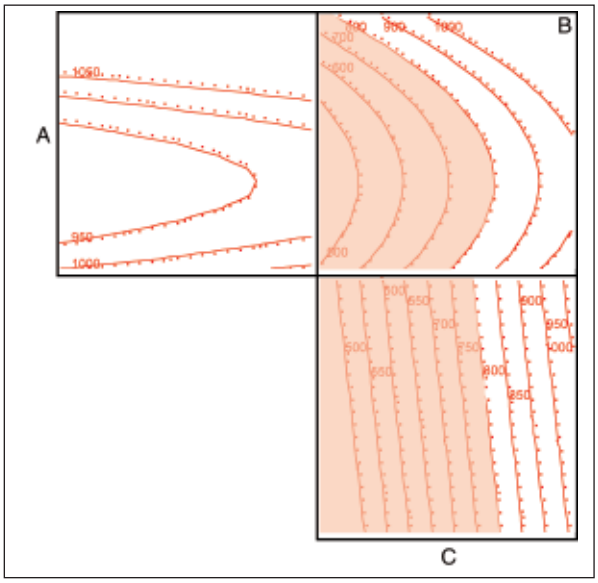


Figure 4. A contour profiler matrix plot.



tings for each of the three significant variables in the PharmaCo example. This tool brings together three-dimensional response surface plots, each of which was originally created in the modeling software. The X and Y axes are made up of the DoE variables, and the Z axis (the contour curves) represent yield (the response variable). In the white regions, yield is in specification, and in the shaded regions it falls outside specification.

Each of the cells is a three-dimensional plot in which the Y variable is to the left of the cell, the X variable is below the cell, and the Z variable is the measured response, yield. For example, the top left cell has raw material attribute 1 as the Y axis and raw material attribute 2 as the X axis; the surface depicted in the plot represents yield.

The goal is to find an available Design Space in which the variables can produce in-specification yield. That is: what values for the variables offer the optimum path to achieving the desired yield? The "contour matrix profiler" shows the areas (in white) in which you can operate successfully and those (shaded) in which you cannot. Those white areas are the Design Space for PharmaCo's fermentation process. In that Design Space, one can identify an area in the center as the operating space, in which the process or outcome will be in control. By choosing this "sweet spot" for yield, and again for the other CQAs as the operating space, they can be sure that the process won't drift into inoperable regions.

THE ADVANTAGES OF UNDERSTANDING DESIGN SPACE

Successfully defining the Design Space means you have achieved a full understanding of the various permutations of input variables and process parameters that ensure an in-specification product. As a result, you gain far more flexibility in changing process parameters and other variables. For example, when raw material batches vary, PharmaCo can make the necessary adjustments in the manufacturing process to compensate for the effects of differing properties and be confident that the resulting product will achieve the desired results.

Despite the significant operational and business benefits that the FDA's QbD initiative offers, only a small number of organizations know how to take full advantage of it. Many continue to use one-variable-at-a-time analysis, even though critical pro-cesses may depend on the complex interactions of several variables. The result: poorly understood processes, an inability to demonstrate adequate control, costly delays in development, and many processes that are neither robust nor reliable. But with carefully structured data analysis through statistical tools, biotech organizations can achieve the robust and reliable processes required to streamline scale-up, technology transfer, and validation, and produce high-quality biologic products. At a time when the pharmaceutical industry is under intense pressure to make safe products and reduce costs, these tools and approaches offer tremendous value.

As illustrated in the PharmaCo example, it's never too late to apply these statistical tools to a process. In fact, it's a straightforward process and can be accomplished in a relatively short period of time. It offers significant enhancement to process understanding, improves scientific rigor, and typically results in significantly enhanced qualitative and quantitative performance as well as cost savings. ◆

REFERENCES

1. International Conference on Harmonization. Q8, Pharmaceutical development. Geneva, Switzerland; 2005.

© Reprinted from BioPharm International, July 2008 AN ADVANSTAR PUBLICATION Printed in U.S.A.

Discover the Value

Founded in 1962, Tunnell Consulting is a globally-positioned management consulting and implementation firm whose expertise improves the operating performance of each client we serve. With a strong focus in the areas of pharmaceuticals, biotech, biologics and medical devices, and supported by our deep industry knowledge and extensive scientific credentials, we integrate strategic, technical, process and organizational skills to design sustainable solutions for our clients.

Tunnell Consulting

Headquarters
900 East Eighth Avenue, Suite 106 • King of Prussia, PA 19406 • ph 610.337.0820 • www.tunnellconsulting.com

King of Prussia, PA Washington, DC San Diego, CA San Juan, PR