**BioPharm International Published Article, 2021**

**Best Practices for Analytical Method Validation: Study Design, Analysis and Acceptance Criteria**

Thomas A. Little PhD, President, Bioassay Sciences

**Background**

A critical element in all drug development from early clinical trials to final form and dose of a drug is the qualification and validation of analytical methods and or bioassays. This paper identifies the relevant ICH/USP guidelines and suggests appropriate and efficient study designs, method of analysis and acceptance criteria. Much of the guidance used in this paper is based on ICH Q2 for validation of any analytical method, ICH Q9 quality risk management and USP <1033> for the validation of biological assays. All study designs suggested are the minimum number of determinations needed to qualify/validate a method. Not all analytical methods are identical and different methods may or may not require all of the elements in this paper; however, for the most part a quantitative or semi-quantitative method will require them.

**Study Designs**

It is useful to break down the qualification/validation of an analytical method into six study designs. Sample sizes are recommended to meet ICH, USP, FDA and other health authority guidance. A determination is one run of the assay starting from the unknown to the reportable result and includes sample preparation.

**Systems Suitability and Validity**

Systems suitability and validity criteria are important elements of all assays and biological assays. Systems suitability is typically based on a standard or on a well characterized positive control and associated limits. Validity criteria typically include a signal control, variation or repeatability control, detection and elimination of outliers, measures of curve fitting and confidence interval ranges for reportable values. Systems suitability and validity criteria are not part of the validation of an assay and must be defined prior to conducting the validation. Finalization of all systems suitability and validity criteria limits may be completed and finalized post validation. Only valid runs of an assay should be included in the validation report.

**Coefficient of Variation (CV) and Recovery**

CV (SD/Mean \* 100) or relative standard deviation (RSD) and recovery (Mean Measured Concentration/Theoretical Concentration)\*100 are not recommended measures of assay performance. They are commonly used but do not correctly reflect the influence of assay performance compared to critical quality attributes (CQA) or measures of productivity. They are not useful and may be misleading because they are scaled to the concentration and will likely indicate that high variation at high concentrations are acceptable and low variation at low concentrations are unacceptable. We do not process, manufacture nor test relative drug product/substance concentrations or impurities and thus CV is irrelevant when compared to the product critical quality attribute. USP <1033> recommends acceptance criteria should be determined relative to the tolerance of product specifications and not relative to a target dilution nor concentration. If CV or recovery is included in a validation report it should be used as report only and should not form part of the assay validation acceptance criteria.

**Study Design 1 Linearity 5x6**

Five concentrations of a reference standard by six independent determinations for each concentration covering 80% of the lower specification limits and 120% of the upper specification limit for a two-sided product specification limit. In order to evaluate accuracy a reference standard must be created or purchased. ICH Q2 does not recommend how to set the acceptance criteria; however USP <1033> states the following: *“When there is an existing product specification, acceptance criteria can be justified on the basis of the risk that measurements may fall outside of the product specification.”.* For a two-sided specification the tolerance is used, for a one-sided specification the margin may be used (Process Average – LSL) or (USL – Process Average) and the process average may be used in the denominator when there is no specification stated. 5.15 is a commonly used risk factor to include 99% of the analytical error.

Assay precision includes repeatability (intra-assay error), intermediate precision (intra and inter-assay error) and reproducibility (inter-lab). Reproducibility is the inter-lab or locational error and is not part of a normal method validation and is typically evaluated during tech transfer if needed.

|  |  |  |
| --- | --- | --- |
| **Reportable Result** | **Method of Analysis** | **Acceptance Criterion** |
| **Accuracy in units and 95% CI** | Mean Difference from Standard  ((Difference)/(USL-LSL))\*100 | <10% of Tolerance for an analytical method  <20% of tolerance for a bioassay |
| **Repeatability in units and 95% CI** | Variance components (within) POV or REML is the recommended method of calculating variance components.  ((SD/sqrt(n))\*5.15))/(USL-LSL)\*100 | <25% of Tolerance for an analytical method  <60% of Tolerance for an analytical method |
| **Intermediate Precision in units and 95% CI** | Variance components (within and between variation) POV or RMEL is the recommended method of analysis  ((SD/sqrt(n))\*5.15))/(USL-LSL)\*100 | <30% of Tolerance for an analytical method  <70% of Tolerance for a bioassay |
| **Linearity, lowest concentration and highest concentration where the response is linear.** | Studentized residuals of the linear fit of the measured concentration | Concentration where the 95% CI of the quadratic fit of the studentized residuals crosses +- 1.96 |
| **Range of the concentration where the assay is accurate, repeatable and liner** | Comparison of concentration and all acceptance criteria | Range where all of the above concentrations are acceptable |

The study design is a main effects only design of experiments with 30 runs. Factors to be included are concentration (5), analyst, day and instrument at two or more levels. Software programs like SAS/JMP are useful in designing the study. Table 1.0 is an example of the study design. Japanese health authority guidance also recommends the column to be included in the study design when appropriate.

Table 1.0 Linearity Study Design

|  |  |  |  |
| --- | --- | --- | --- |
| **Concentration** | **Day** | **Analyst** | **Instrument** |
| 40 | L1 | L1 | L2 |
| 50 | L1 | L2 | L1 |
| 30 | L1 | L1 | L2 |
| 60 | L1 | L2 | L2 |
| 70 | L1 | L2 | L2 |
| 50 | L1 | L1 | L2 |
| 40 | L1 | L2 | L1 |
| 40 | L1 | L1 | L1 |
| 30 | L1 | L1 | L1 |
| 30 | L1 | L2 | L1 |
| 70 | L1 | L1 | L1 |
| 70 | L1 | L1 | L2 |
| 60 | L1 | L1 | L1 |
| 50 | L1 | L2 | L2 |
| 60 | L1 | L2 | L2 |
| 70 | L2 | L1 | L1 |
| 50 | L2 | L1 | L2 |
| 50 | L2 | L2 | L1 |
| 60 | L2 | L1 | L1 |
| 50 | L2 | L1 | L1 |
| 40 | L2 | L2 | L2 |
| 30 | L2 | L1 | L2 |
| 30 | L2 | L2 | L2 |
| 60 | L2 | L2 | L1 |
| 30 | L2 | L2 | L1 |
| 40 | L2 | L1 | L2 |
| 40 | L2 | L2 | L1 |
| 70 | L2 | L2 | L2 |
| 60 | L2 | L1 | L2 |
| 70 | L2 | L2 | L1 |

**Study Design 2 Limit of Detection (LOD) and Limit of Quantification (LOQ)**

Two concentrations of a reference standard by six independent determinations for each concentration at or near the estimated LOD and LOQ limit.

|  |  |  |
| --- | --- | --- |
| **Reportable Result** | **Method of Analysis** | **Acceptance Criterion** |
| **Limit of Detection** | SD at concentration near LOD \* 3.3 (divided by the slope of the signal and theoretical concentration if the SD is not in units) | Report only |
| **Limit of Quantification** | SD at concentration near LOQ \* 10  (divided by the slope of the signal and theoretical concentration if the SD is not in units) | Report only |

**Study Design 3 Specificity Identification 2x3**

Two conditions a positive control and a negative control by three independent determinations.

|  |  |  |
| --- | --- | --- |
| **Reportable Result** | **Method of Analysis** | **Acceptance Criterion** |
| **Difference between the positive control (standard) and the negative control and the 95% CI** | Means ANOVA *t-*test | Must be statistically significant (*p*-value <0.05, two-sided test) |

**Study Design 4 Specificity Interference 3x3**

Risk assessment with all interfering compounds, impurities, excipients etc. Three concentrations of the reference standard for every interfering compound spike in representative concentrations of the interfering compound. Results are compared to the linear curve (unspiked) in study design one (5x6). USP <1033> states, “*For products or intermediates associated with complex matrices, specificity involves demonstrating lack of interference from matrix components or product-related components that can be expected to be present. This can be accessed via parallel dilution of the Standard sample with and without a spike addition of the potentially interfering compound*”.

|  |  |  |
| --- | --- | --- |
| **Reportable Result** | **Method of Analysis** | **Acceptance Criterion** |
| **Difference between the spiked and unspiked sample intercepts and 95% CI.** | Two-way ANOVA, concentration, spiked/unspiked and the concentration and spiked/unspiked interaction. ((Difference)/(USL-LSL))\*100 | Criterion is the same as accuracy,  <10% of tolerance. Evaluate both the main effect of the spiked/unspiked and the two-factor interaction if significant. |

**Study Design 5 Stability Indicating 1x5**

One standard concentration under stressed conditions (temperature or pH) at five time points.

|  |  |  |
| --- | --- | --- |
| **Reportable Result** | **Method of Analysis** | **Acceptance Criterion** |
| **Rate of degradation and the 95% CI from a linear or nonlinear fitted curve.** | Linear or nonlinear fitted regression curve and all curve parameters and 95% CIs | Must demonstrate a statistically significant (*p*-value <0.05, two-sided test) rate of degradation to be considered stability indicating. |

**Study Design 6 Robustness**

Robustness is a key element of method qualification/validation and is normally performed before the qualification/validation and is included in the validation report. Normally a single analyst, single instrument and a single representative concentration is used in the robustness study. A formal risk assessment (ICH Q9) of all reagents, reagent lots, sample preparation, sample handling, method parameters and sample hold times are evaluated for inclusion in the robustness study design. Software such as SAS/JMP are used to design the experiment with all of the main effects, two-factor interactions and quadratic model terms included in the study design from the risk assessment. Three independent determinations are performed for each experimental run. Mean and standard deviation of the n=3 determinations are calculated. The model is then fitted to the mean difference from the standard and standard deviation results. The nominal best set point or recipe for the method is defined and clear operational limits are set from the robustness study. Limits are set where the method will be accurate (<10% of tolerance) and repeatable (<25% of tolerance) and documented in the method standard operating procedure.

|  |  |  |
| --- | --- | --- |
| **Reportable Result** | **Method of Analysis** | **Acceptance Criterion** |
| **Robustness model equation for accuracy and repeatability. Set points for all method parameters evaluated. Operational limits for all method parameters evaluated.** | Multiple regression if all factors are continuous or ANCOVA mixed model if both continuous and categorical factors are in the study design. | Report only |

**Evaluating Acceptance Criteria**

When evaluating acceptability of an analytical method’s error and variation it should be viewed as a budget of error and how it influences out-of-specification error rates. Method errors are random and normally distributed for the most part. Table 2.0 show a typical recommended budget for the method and for the process. The variance of the errors are summative and can be viewed as a percentage of tolerance.

Table 2.0 Budget of Analytical Error

|  |  |
| --- | --- |
| Sources of Error | Recommended Budget (% of Tolerance) |
| Method Accuracy | 10 |
| Method Repeatability | 25 |
| Method Intermediate Precision | 30 |
| Process Mean delta from target | 10 |
| Process Variation | 25 |
| Design Margin | 20% or more depending on accuracy and the process mean delta from target |

To truly evaluate all of the moving parts of the process and the analytical method a simple calculator can integrate and evaluate all of the sources of error and demonstrate acceptance criteria and their influence on OOS rates for any CQA. This calculator factors in the process variation with the assay variation. If the process variation is higher the assay allowable error will need to be reduced to have an OOS rate less than 100 PPM. Also if the process variation is low the assay error may be larger and still meet the <100 PPM goal. Figure 1.0 shows an example of the elements of the above budget and integrates them all into an overall assessment of the probable OOS rate in percent and PPM. The calculator may be downloaded from the reference website. 

Figure 1.0 Assay Evaluation Calculator

References:

ICH Q2(R1) Validation of Analytical Procedures: Text and Methodology, 2005

ICH Q9 Quality Risk Management, 2006

USP <1033> Biological Assay Validation

www.bioassaysciences.com