Developing Representative Sampling Plans for Development, Problem Solving and Validation

Thomas A. Little Ph.D.
10/02/2014

President Thomas A. Little Consulting
12401 Wildflower Lane
Highland, UT 84003
1-925-285-1847
drlittle@dr-tom.com

In all product and process development and in all validation activities, developing a representative sample and the correct and defendable use of statistical method is a critical activity. In all sampling activities a scientifically sound sampling plan should be developed, justified, and implemented. Those activities include clinical trials, formulation, stability (long term and accelerated), process characterization, analytical method development, method validation, process validation and release testing.

Modern drug development, ICH and health authority's guidelines such as (ICH) Q8, Q9, Q10 and Q11 all assume statistically valid sampling procedures in test and product and process development.

Specifically Q8, Q9, and Q10 Questions and Answers (R4), states:

> *"Q6: Do traditional sampling approaches apply to real time release (RTR) testing?*
>
> *A6: No, traditional sampling plans for in-process and end-product testing involve a discrete sample size that represents the minimal sampling expectations. Generally, the use of RTR testing will include more extensive on-line/in-line measurement.* ***A scientifically sound sampling approach should be developed, justified, and implemented****.*

(Approved April 2009) processing, in-process materials, and drug product quality can provide an opportunity to shift controls upstream and minimize the need for end product testing."

The purpose of this paper to organize a logical framework and identify the appropriate tools that aid in the development of a statistically valid sampling protocol. Due to the complexity of the kinds of issues and the types of problems to be solved each sampling procedure will be somewhat unique; however, there are common issues and questions that each sampling protocol will need to consider and address.

The following are generally key steps for using statistical methods for problem solving and product development:

1. Define the business case
2. Define the problem
3. Define all objectives, goals and study questions
4. Define all factors and responses of interest and associated analytical methods
5. Define the population
6. Define the sample that will represent the population, study questions and the defined problem
    a. Sampling Method
    b. Sample Size
7. Collect data per sampling plan
8. Summarize data with key statistics for the sample and confidence intervals for the defined population
9. Draw conclusions and inference
10. Verify conclusions, solve the problem and achieve the business case

These steps are covered in detail below.

1. Define the business case

The business case explains the development objective, how each activity relates to quality objectives, QTTPs, CQAs, timelines or product development requirements and why the activity needs to be completed. The business case gives context to development and validation activities and helps CMC teams to understand how each activity fits into overall business and quality imperatives. Failure to understand the business case for development and validation activities causes discontinuity in development activities and the logic of what, why and how much becomes hard to justify and or file to the appropriate health agencies.

2. Define the problem

Problem definitions define what we don't know and or what is wrong with our current performance. Problem definitions are essential for problem sampling plan development and rationalization. Limitations and or scope associated with each problem definition is critical in defining the associated population of units associated with the problem. Also associated with the problem definition, there should be clarity on what is the sampling unit (batch, vial, drum etc.)

3. Define all objectives, goals and study questions

From the problem definition, objectives, goals and study questions must be defined. Goals come in four forms: 1) maximize, 2) minimize, 3) match target and 4) none. Limits and acceptance criteria should also be defined. The sampling plan will be designed to ensure we answer all study questions, achieve the goals and are sufficiently precise and accurate relative to the limits and tolerances of our acceptance criteria.

4. Determine all factors and responses and analytical methods

Based on the study questions and goals what must be measured to meet the business case and solve the problem? All factors that influence the problem and study questions must be defined. All responses and analytical methods associated with the problem need to be clarified. In many cases the factors and/or responses that are currently measured in batch records or in on-line data systems do not address the problem we are trying to solve or the process we are trying to characterize. Care should be exercised to assure there are no missing measurements or we will not meet the business case or address the problem correctly.

5. Define the population

Based on the study questions, unit definition and problem statement, what is the population of units that need to be understood? Population definition is critical prior to sampling plan formalization. For R&D the population is often a function of formulation and or configuration. Another term that is often associated with the population is "volume". How will the product perform in volume and at scale versus the limited product testing that is performed in small scale studies?

6. Define the sample that will represent the population, study questions and the defined problem

A sampling plan needs to be formally defined and scientifically justified to assure it is representative of the population and it meets the all defined study objectives and acceptance criteria.

WHO guidelines for representative sampling states:

"***Representative sample***, *Sample obtained according to a sampling procedure designed to ensure that the different parts of a batch or the different properties of a non-uniform material are proportionately represented."*

ICH Process Validation, regarding sampling plans states:

"The sampling plan, including sampling points, number of samples, and the frequency of sampling for each unit operation and attribute. The number of samples should be adequate to provide sufficient statistical confidence of quality both within a batch and between batches. The confidence level selected can be based on risk analysis as it relates to the particular attribute under examination. Sampling during this stage should be more extensive than is typical during routine production."

All statistically justifiable sampling plans require two primary considerations, sampling method and sample size.

6a. Define the sampling method

The sampling method is defined to clarify "***how***" samples are taken, "***where***" they are taken and "***how often***". How many samples are taken is sample size. To determine the sampling method care must be exercised to assure the samples are taken from the primary sources of variation. Partition of Variation (POV) or Components of Variation (COV) analysis is used to determine the proportion of variation within/between batch for example. In the below example 54% of the variation occurs within the lot, so at least 54% of the measurements need to be allocated to the within lot sampling to be representative. The batch-to-batch variation is only 3% so we don't need many batches.

Partition of Variation (POV) Crossed, Dependent Analysis

2 Level POV Analysis: Sublot -Fill/Orientation, Lot Number

Component	Pop Variance	% of Total	Sqrt(Var Comp)	F Ratio	Prob>F	Sig.
	.	.	.	.	.	
Between Total	0.000021	4.55	0.004601	0.5900	0.9622	
Between Sublot-Fill/Orientation	0.000002	0.43	0.001422	0.2498	0.9721	
Between Lot Number	0.000016	3.41	0.003981	1.1415	0.3247	
Between Sublot-Fill/Orientation*Lot Numbe	0.000003	0.71	0.001814	0.2370	0.9964	
	.	.	.			
Within Total	0.000444	95.45	0.021080			
Within Sublot-Fill/Orientation	0.000107	22.94	0.010334			
Within Lot Number	0.000253	54.27	0.015895			
Within Sublot-Fill/Orientation*Lot Number	0.000042	9.01	0.006476			
Common	0.000043	9.24	0.006559			
	.	.	.			
Total	0.000466	100.00	0.021577			

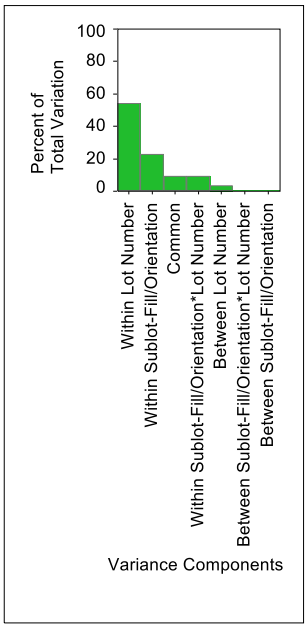

Figure 1.0 Partition of Variation Analysis for Sampling Method Determination

6b. Define the sample size

Sampling method must be defined first and then sample size. There are many ISO standards for variables and attributes sampling for lot acceptance and there are also the NIST standards for determining the sample size. SAS/JMP and other statistical packages have their own sample size and power calculators. Every sampling protocol involves risk, the risk should be known rather than unknown.

To use a sample size calculator you need to know the confidence interval you want (1-alpha), the power of the test, how reliably you want to detect a change (1-beta) and the practical change you want to observe (delta). For example: alpha = 0.05, the power =0.95 and the delta in pH that we want to detect is 0.2 and the standard deviation of pH at the point of evaluation is 0.125 from historical measurements. What is the sample size? 8 samples will do it. Sample size does not address sampling method. Sampling method tells you how to take the sample, sample size tells you how many. Power curves are used when the delta is unknown.

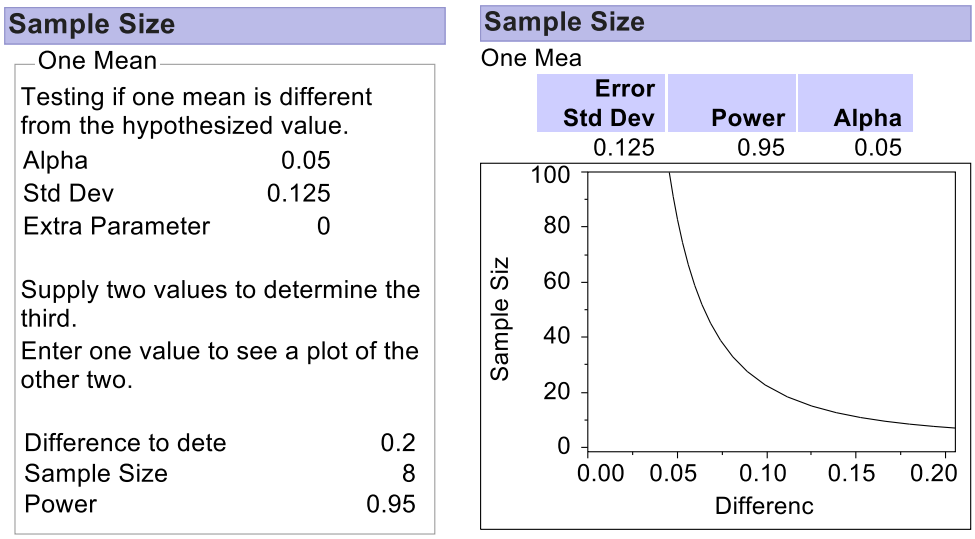

Figure 2.0 Sample Size and Power Curve

7. Collect data per the sampling plan

Follow the sampling plan and collect data and all associated data tags (time, date, analyst etc.) during sampling.

8. Summarize data into statistics for the sample and confidence intervals for the population

Statistics and graphs are used to summarize and aggregate the data. The statistics describe the sample and the confidence intervals are used to describe the population. Confidence intervals control for risk (95% for example), variation in the data and sample size. Confidence intervals should be in every graph and in every table.

9. Draw conclusions and inference

Based on the data and all associated confidence intervals, models and graphs, draw conclusions relative to acceptance criteria, limits, CQAs and the business case.

10. Verify conclusions, solve the problem, meet quality requirements and achieve the business case

Verify the conclusions and prediction made from the sample are subsequently observed in the population. Demonstrate that solutions are generally applicable though on-going longitudinal monitoring and continuous validation protocols. Verification indicates our earlier estimates taken from the representative sample are correct. Failure to verify our conclusions and inferences generally indicates additional uncontrolled factors are at play and need to be understood and or controlled before we will achieve the business case.

Summary

The ability to define a scientifically justified and statistically sound sampling procedure is a fundamental skill in modern systematic drug development. It impacts every aspect of development and validation. A structured approach using the key considerations outlined in this paper will aid in assuring it has a defendable technical basis for sampling method and sample size selection and controls and addresses risk relative to a clearly defined business cases, CQAs, problem statements and study questions.

References:

Annex 4, WHO guidelines for sampling of pharmaceutical products and related materials, 2005

ICH Q8, Q9, and Q10 Questions and Answers (R4), 2011

ISO 10725 Acceptance sampling plans and procedures for the inspection of bulk materials. 2000

ISO 3951:1989. Sampling procedures for inspection by variables. Specification for single sampling plans indexed by acceptable quality level (AQL) for lot-by-lot. 1993.

ISO 2859-1. Sampling procedures for inspection by attributes. Sampling schemes indexed by acceptance quality limit for lot-by-lot inspection. 1999.

ICH Process Validation, General Principles and Practices, 2011