Authors: Patrick Bird, PMB BioTek Consulting, LLC, and Quynh-Nhi Le, NEOGEN®

A Guide to Hemp Method Validations

Paper Three — Five Part Series

Executive Summary

The first two papers in this series have provided the information to allow end users to understand and select a validated method for their testing needs. This paper will focus on aspects for two different types of validations, those for chemistry and microbiology methods.

Different Targets Require Different Validations

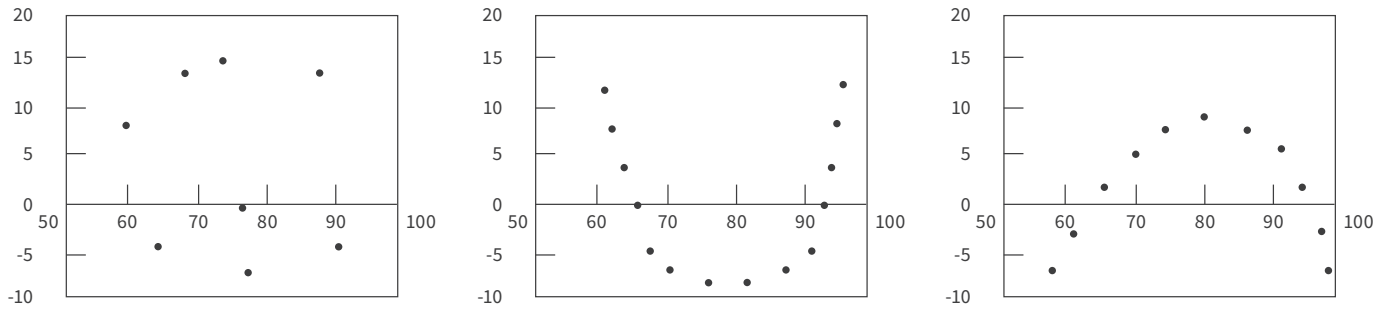
To optimize sales, cultivators must ensure their products are safe for consumption from harmful contaminants (pesticides, heavy metals, microbials) and possess enticing profiles (potency, terpene profile) that will attract and retain consumers. To capture this information, a variety of different types of tests, both chemical and microbiological, are performed. These tests can be simplified screening assays or complex methods that require training and expertise. Regardless of the complexity, each method should be put through a rigorous validation study. While the main requirements of validation studies for chemistry and microbiology methods are consistent[1], the study designs for each can be starkly different.

Chemistry

Chemistry methods (including methods for allergen detection and quantification) safeguard that harmful contaminants are at or below acceptable limits set by regulatory agencies, verify nutritional properties are accurate on labels, and help ensure that the product's organoleptic properties meet consumer expectations. Four main components of chemistry validations should be performed.

1. Calibration/linearity — These studies are designed for quantitative or semiquantitative methods and will evaluate the dose response relationship between the target analytes and the method's response. Testing is performed over the expected concentration range of the method and will include blanks (controls), low, and high levels.
   - Calibration curves are developed using pure target analytes, although a fortified matrix can be used. Testing is conducted at several concentrations of interest using multiple replicates (2–5). Best fit lines and correlation coefficients are determined. Correlation coefficients with values ≥ 0.98 indicate an acceptable dose response relationship.
   - Data is interpreted for non-random patterns (figure 1) using a regression model (linear, weighted linear, quadratic). If non-random patterns are observed, a different regression model should be employed. Data producing residuals at <20% typically indicate that the method's response to the target can be considered linear.

Figure 1: Regression Analysis Demonstrating Random and Non-random Patterns



I. Random Pattern | II. Non-random Pattern | II. Non-random Pattern

NEOGEN®
NEOGEN.com

2. Selectivity — This study is designed to demonstrate a method's ability to:
   a. Detect the full range of target analytes within the method's claim
   b. Detect the target analytes of the method without interference from matrices or components with similar properties
   c. Not detect molecules similar in chemical structure or activity to the target analyte

   - A list of potential interferants is selected based on their presence in the matrices of interest, as well as their similarity to the targets of the assay. Testing is performed on the non-target analytes and matrix to determine if cross-reaction with the assay occurs. Testing is repeated with the addition of a target analyte at a known concentration to determine if there is positive interference (enhancement of the analytical response) or negative interference (reduction of the analytical response).
   - A determination that a non-target analyte causes interference (either positive or negative) or cross-reactions does not indicate the method cannot achieve validation status. However, this information should be provided clearly in the instructions for use so that end users can take this information into account when performing the method.

3. Matrix studies — The matrix study is designed to evaluate the method's performance for the target analytes in the claimed matrix.
   - Matrix studies can be performed using reference materials, reference standards, and as a comparison with a reference method.
     - Reference materials (RM) or certified reference materials (CRM) are sufficiently homogeneous and stable materials with one or more specified properties established to fit their intended use in a measurement process[2]. RMs and CRMs are provided by third-party vendors or national metrology organizations.
     - Reference standards are a traceable standard of an analyte with documented identity and purity.
     - Reference methods are pre-existing official regulatory or other nationally/regionally recognized analytical methods against which the candidate method can be compared.
   - If the method claims multiple analyte detection simultaneously, testing must be performed with all analytes in at least one matrix.
   - Testing should be performed with blind coded and randomized test portions, reducing testing bias from the technician performing the analysis.
   - In a quantitative study, each matrix will be evaluated at 4–6 concentration levels with multiple replicates at each level. The study results will be analyzed for multiple statistical parameters, which may include recovery, bias, LOD, LOQ, repeatability, and intermediate reproducibility.
   - In a qualitative study, each matrix will be evaluated at 4–6 concentration levels with 20–30 replicates at each level. Depending on the study design, statistical approaches can determine bias between the candidate and reference method, the 90/95% probability of detection (probit analysis), or another appropriate statistical approach.
     - Probit analysis is a method used to analyze the relationship between the concentration, or dose, and the binary (presence/absence) response. In method validation, probit analysis is often used when evaluating a method's ability to detect antibiotic residues in dairy matrices. A 90/95% probit analysis is the concentration that will detect 90% of samples analyzed and is evaluated with a 95% confidence interval.
4. Robustness — This study is designed to evaluate a method's ability to withstand minor variations to specific testing parameters that could occur when end users perform the method.
   - Variations in testing parameters will focus on critical components of a method related to storage, extraction, reagent concentrations or volumes, and test portion to solvent/reagent ratios.
   - These parameters are varied and tested using an experimental design (e.g., classic, full factorial, fractional factorial, Plackett-Burman).
   - A classic design will evaluate a single parameter at a time to determine the effect of that parameter on the performance of the assay. Testing is performed by modifying the parameter at conditions above, below, and at the nominal testing conditions.
   - A factorial design (either full or fractional) will identify several parameters to evaluate at conditions above, below, and at the nominal testing conditions. For methods with minimal steps, a full factorial design can be employed. For more complex methods with multiple steps, a fractional factorial design can be utilized. The use of factorial design studies maximizes variations in a smaller study design, making it easier for the method developer to evaluate multiple parameters. (See tables 1 and 2)

Table 1: Example of Parameters for a Factorial Study Design

Test Parameter	Low Value	Nominal Value	High Value
Mass/Solvent Ratio (g/mL)	8/102	10/100	12/98
Reagent Volume (mL)	4.5	5	5.5
Extraction Time (min)	8	10	12
Extract Volume (µL)	0.4	0.5	0.6

Table 2: Example of Testing Conditions for a Factorial Study Design

Parameter	Combination							
	1	2	3	4	5	6	7	8
A or a	A	a	a	A	a	A	A	a
B or b	B	b	B	b	B	b	B	b
C or c	C	C	c	c	C	C	c	c
D or d	D	D	D	D	d	d	d	d

- When performing robustness testing, multiple replicates are prepared from a matrix fortified with the target analyte and evaluated.
  - Data is analyzed by Analysis of Variance (ANOVA), multi-factor regression, or other appropriate statistical analysis. The objective of the analysis is to determine if single or multiple parameters will impact the performance of the assay and generate a different result when varied.

Microbiology

Microbiology methods are designed to ensure that the product evaluated is safe for consumption by being free from pathogenic bacteria or that the quality of the product is not impacted by spoilage or contamination organisms. Testing is often performed to meet regulatory requirements.

1. Matrix Study — These studies compare the results of a candidate method to a known reference method. When no reference methods exist, cultural confirmation of the candidate method using standardized guidance (for example, SMPRs) is conducted to evaluate the method's performance.
   - Matrix testing can be performed with naturally occurring microorganisms already present on the product or through artificial contamination.
     - Artificial contamination is performed using viable cultures. The use of extracted bacterial components (DNA, protein) is not acceptable. Viable cultures allow the method to be evaluated under the conditions with which an end user would perform the method.
   - Prior to analyzing these cultures, the organisms should be stressed, where appropriate, to mimic conditions of real-world settings. For example, dried hemp material should be tested with a dried, or lyophilized, culture. The stressing of the culture is intended to mimic the condition with which end users would encounter the organism.
   - Stressing conditions allow for equilibration, or adaptation of the organism to its environment. This equilibration allows for a more realistic evaluation of the method's ability to detect or enumerate the target.
   - All matrix testing should be performed at multiple concentrations of the target analyte.
     - Quantitative methods — Testing is performed over a contamination range anticipated to be found in the field, with a low level near the LOQ of the method and two additional levels approximately one log higher. For matrices that are artificially contaminated, a blank level should be included. Standardized validation guidelines require a minimum of five test replicates to be evaluated for each contamination level.

Table 3: Experimental Design for a Quantitative Microbiology Study

Matrix	Target Analyte	Stabilizing Conditions	Inoculation Level	Replicates per Method	Culture Confirmation
Hemp Flower	Total Yeasts and Molds	Not Applicable; Natural Contamination	Low	5	SMPR 2020.002
			Medium	5	
			High	5	

- Qualitative methods — Testing is performed at multiple contamination levels, emphasizing the contamination level targeting either the LOD50 or LOD100 of the method. This emphasis is achieved by increasing the test replicates to 20–30 for this contamination level. When evaluating the method at the LOD50 level, testing should result in fractional positive results, a 25–75% positivity rate. Testing performed at the LOD100 level will result in all or nearly all positive results. This requirement helps to demonstrate the performance of the method at the tested LOD. Control and high levels are also included but tested at a lower frequency (five test replicates).

Table 4: Experimental Design for a Qualitative Microbiology Study

Matrix	Inoculation Organism (Condition	Stabilizing Conditions	Inoculation Level	Replicates per Method	Culture Confirmation
Hemp Flower	*Salmonella* Typhimurium (Lyophilized)	2 Weeks at 20–25°C	Non-inoculated	5	SMPR 2020.002
			Low	20	
			High	5	

- After performing the analysis, the bias, or difference in results between the candidate method and the reference method (or confirmed result) is determined.
  - Quantitative methods are evaluated for their repeatability or how closely the results between replicates at the same concentration align.
  - Qualitative methods are evaluated to determine if false positive (incorrect positive detection) or false negative (missing detection) results occurred.

2. Inclusivity and exclusivity study — The inclusivity and exclusivity study evaluates the ability of a method to detect target strains and discriminate them from closely related non-target strains or other background organisms traditionally encountered in the matrixes evaluated.
   - Evaluations are performed following the entire protocol of the candidate method. This includes all enrichment broths, extraction protocols, and detection technologies prescribed in the method.
   - Testing is performed on pure cultures, although the inclusion of matrix into the study is acceptable. Strains must be evaluated in a blind coded randomized study so that the analyst performing the test is not biased in their performance.
   - Target strains are evaluated at 10–100 times the LOD_{100} of the method, while non-target strains are evaluated after optimal growth conditions.
3. Robustness study — Microbiology robustness studies are performed similarly to chemistry robustness studies using classic or factorial study designs that will evaluate small changes to the testing protocol.
   - Critical parameters often evaluated are transfer volumes, time frames, and temperatures throughout the testing protocols. These factors are evaluated as they require calibration of specific instruments (pipettors, timers, heat blocks/incubators) where variation will occur between testing facilities.
   - In addition to these factors, other parameters, such as the brand of media, multiple analysts, and multiple instruments (incubators, thermocyclers, etc.), can be incorporated into the design, although traditionally, this is evaluated during method verification[3].

- Robustness testing can be performed with pure culture, although the inclusion of matrix is preferred. Multiple replicates are prepared with the target analyte and evaluated using each parameter to be studied.
- Data is analyzed in a similar approach to that of the matrix studies. Each combination of variables tested can be compared to the nominal combination to determine if a bias exists.
- Robustness testing will help identify if a single step or multiple steps are critical within the process. If small variations occur at these steps, method developers can provide more robust guidance in instructions to ensure the method performs appropriately.

Conclusions

Method validations for chemistry and microbiology require laboratory studies to determine whether the performance characteristics of the method meet established requirements for its intended use. Chemistry and microbiology methods can further be broken down into qualitative (presence/absence) or quantitative (numerical output), which can look similar but differ in study design, the number of replicates evaluated, and statistical analysis. The chemistry method validation includes linearity, selectivity, matrix studies, and robustness to determine whether the method is appropriate for detecting harmful contaminants in specific matrices. The microbiology validation includes matrix studies, inclusivity/exclusivity, and robustness to determine whether the method is appropriate for detecting pathogenic or spoilage bacteria in specific matrices. By understanding each validation type, the end user is confident that the method's performance is fit for purpose and meets regulatory guidelines when applicable.

1 See Paper 2, What does it mean to have a validated method, for more details
2 Definition courtesy of the National Institute of Standards and Technology (NIST)
3 See Paper 5, Validation versus Verification