



Animal and Plant  
Health Inspection  
Service

Veterinary Services

1400 Independence  
Ave, SW

Washington, DC  
20250

## VETERINARY SERVICES MEMORANDUM NO. 800.73

**TO:** Veterinary Services Leadership Team  
Directors, Center for Veterinary Biologics  
Biologics Licensees, Permittees, and Applicants

**FROM:** Jack A. Shere  
Deputy Administrator

**SUBJECT:** Diagnostic Test Kit Validation

### I. PURPOSE

This memorandum provides guidance to firms (licensees, permittees, and applicants) to support an application for a U.S. Veterinary Biological Product License or U.S. Veterinary Biological Product Permit for test kits intended to detect animal disease or immunological status, as authorized by title 9, *Code of Federal Regulations* (9 CFR), part 101.2.

This document outlines an approach for validating a diagnostic test, setting potency specifications, and confirming a product's dating period. Veterinary Services (VS) provides these guidelines to the firms conducting studies to submit to the Center for Veterinary Biologics (CVB) in support of licensure of a diagnostic test kit.

### II. REPLACEMENT

This document replaces VS Memorandum No. 800.73 dated January 6, 2015. VS removed some of the content in section VII (Outline of Production) of that memorandum and included the content in an update to [VS Memorandum No. 800.206](#).

### III. BACKGROUND

Diagnostic test kits detect the disease or immunological status of an animal, flock, or herd. Diagnostic kits must be validated to demonstrate that they are scientifically sound, reliable, reproducible, and fit for their intended use.

#### IV. ASSAY VALIDATION

Conduct assay validation studies to demonstrate the kit's diagnostic accuracy, analytical sensitivity, analytical specificity, and ruggedness. Also conduct an interlaboratory comparison study. The CVB strongly recommends submitting a protocol for comment before initiating a study. Upon completing the study, submit a report including all raw data. (See the [CVB Data Guide](#) on the National Center for Animal Health Portal Guidance website for instructions regarding electronic data submission.)

- A. Diagnostic Accuracy (Diagnostic Sensitivity and Diagnostic Specificity). Estimate diagnostic sensitivity and specificity of the experimental test kit by comparing to one or more reference tests in a sample set using the method described in [STATWI0002](#). This method allows for the possibility that the accuracy of the experimental kit may be superior to that of the reference test rather than assuming that all discrepancies between the two are the fault of the experimental kit.<sup>1</sup> Required inputs to the method include information about the reference test's accuracy and the prevalence within the sample set.

1. *Sample set*

- i. *Nature of the sample set.* Ideally, the sample set is composed of specimens from animals selected randomly from an intended target population, but in practice that is often not feasible. Instead, the sample set may be randomly selected from a group of available specimens.
- ii. *Sample set construction.* Create a sample set of specimens by randomly sampling a group of available specimens containing both positives and negatives. Describe in detail the selection of the sample set and provide an estimate of the presumed prevalence in the sample set. (The prevalence is the fraction that are truly disease positive. The presumed prevalence is the user's imperfect assessment of the prevalence.)

iii. *Examples of sample set selection methods* (not an exhaustive list):

*Example.* The firm randomly selects specimens from a repository that contains specimens designated as both positive and negative. Base the presumed prevalence on the fraction of presumed positive specimens within the repository.

*Example.* The firm randomly selects specimens from an assembled group that includes a specified ratio of presumed positives and presumed negatives. Base the presumed prevalence in the selected set on that specified ratio.

*Example.* The firm randomly samples several herds that each have a roughly known prevalence. If there are enough specimens, each herd gets its own

---

<sup>1</sup> If a reference test is incorrectly assumed to be perfect, any disagreement would serve to lower the apparent accuracy of the experimental test. By contrast, this method allows the experimental kit and the reference test to each have a probability of being correct for a given specimen.

sample set, which may improve the sensitivity/specificity calculations. If not, the firm prepares a single sample set, and derives its presumed prevalence from the individual ones.

*Exception.* Construct sample sets by sampling from a group of specimens that contains both positives and negatives. The only exception is when there is a test method that is considered undeniably infallible. For example, necropsy findings for canine heartworm are infallible for positive results but not negative results, the firm may construct a sample set of positives from the known positives. In this instance, construct an additional sample set by sampling from specimens that contain both positives and negatives.

2. *Reference tests.* Use one or more reference tests. Include information about the sensitivity and specificity of each proposed reference test in the protocol and final report. The firm should do this by providing their assessment of the reference test's anticipated sensitivity and specificity using the best information available to them, and stating the source of that information. Do not expect this assessment to be perfectly correct, since the reality is often not precisely known. It indicates what is actually known about the performance of the reference test and how to use that information in evaluating the experimental kit.
3. *Test concurrently.* Concurrently test each sample specimen with the experimental kit and the reference test or tests. Concurrent testing means both the experimental kit and the reference test or tests are run after selecting the sample set, as close in time as practical, ideally on the same day. If concurrent testing is not feasible, explain why in the protocol and describe how you will conduct all testing. Do not use the reference test to construct a sample set by including or excluding specimens based solely on the reference test's result. However, you may use specimens previously tested by the reference test for other purposes, including admission to a repository of specimens.
4. *Testing regimen.* Use the same testing regimen for all specimens in every sample set. For example, do not selectively conduct retests based on the results of the experimental kit.
5. *Visual and machine measurements.* For kits in which the response is determined by visual inspection, use the visual classification to estimate sensitivity and specificity. Take an instrument measurement, such as densitometry to assist in selecting panel members for serial release and establishing the relationship between the instrument measurement and the visual read. Make the visual classification without knowing the instrument measurement.
6. *Sample specimens.* Determine the diagnostic sensitivity and specificity separately for each proposed specimen type (e.g., blood, feces) and host species. There may be instances in which it is not necessary to determine sensitivity and specificity for each proposed specimen type when the specimen types are closely related (e.g., blood, serum, plasma); in that case, a reduced number of specimens may be sufficient to demonstrate similar performance. Do not use

experimentally infected animals to estimate diagnostic sensitivity, although such specimens may provide supplemental information. Specify in the protocol the host species, specimen types, and proposed number of specimens of each type. Include pertinent information about the acquisition of samples, such as geographic location and specimen treatment. Obtain sample specimens from U.S. sources or provide justification in the protocol for specimens obtained from foreign sources.

7. *Number of Serials.* While a single serial is acceptable, use more than one, as you will need the serials used in the diagnostic accuracy study to set serial release specifications and demonstrate appropriate panel members formulation. Test the same sample specimens on all serials. If that is not possible because specimens are limited in quantity, the firm should propose a suitable design in a protocol before starting the study.
  8. *Adding pooling claims.* Form pooled samples by combining individual specimens. To support a claim for the use of pooled samples in a test kit, conduct a study using pools of the desired size (n). First, test each individual sample specimen with the experimental test kit. Then, formulate pools by combining equal amounts of n specimens, one that tested positive and n-1 that tested negative. Every specimen in the study must be from a different animal, so that each animal contributes exactly one specimen to one pool.
  9. *Group samples.* Group samples are those where specimens are collected from a group rather than individuals. Group samples may be actively collected from all the group members, such as a bulk tank milk sample, or passively collected from an undetermined set or subset of the group members, such as oral fluids obtained from ropes. Group samples require special considerations, and the firm should seek concurrence from CVB before initiating any data collection.
  10. *Multiple serotype or serovar claims.* To support a label claim for the detection of multiple serotypes or serovars, the sample set should include all serotypes or serovars present in the target populations.
  11. *Presenting sensitivity/specificity estimates.* Sample sets constructed for the study may differ from an intended target population in notable ways. These include the nature of the specimens (e.g., stage of disease). Therefore, the estimates of diagnostic sensitivity and diagnostic specificity of the experimental kit in the study may not entirely reflect its performance when it is used in a target population for various purpose (e.g., screening or confirmatory diagnosis). For that reason, accompany estimates of sensitivity and specificity in the product literature with a description of the sample sets used.
- B. *Analytical Sensitivity/Limit of Detection.* Conduct a study to establish the relationship between analyte concentration and the percentage of samples classified as positive for a given concentration. If the concentration of analyte in the sample cannot be determined, use dilution in place of concentration. This information is helpful in selecting the appropriate formulation for the weak positive panel member

used in serial release testing. The analytical sensitivity and analytical specificity studies are often done before the diagnostic accuracy study.

- C. Analytical Specificity. Demonstrate the lack of cross reactivity of the kit by testing for analytes that are similar to, but different from, the intended analyte. For this purpose, similar means genetically related. Evaluate specificity in a separate study or by specimens included in the inter-laboratory comparison study.
- D. Determining the Cutoff Value. Some kits may produce a numeric test result used in conjunction with a cutoff value for determining the status (positive/negative) of the sample. For example, a specimen producing a signal-to-positive value less than the cutoff would be classified as negative, and a value equal to or greater than the cutoff would be classified as positive. Describe how the cutoff is determined.
- E. Ruggedness. Evaluate ruggedness by observing the effect of changes in incubation time, incubation temperature, or other critical test conditions on the final test results.
- F. Interlaboratory comparison. Conduct an interlaboratory comparison to evaluate the ability of the test kit to produce consistent test results when used by more than one laboratory.
  - 1. Prepare a test panel consisting of at least 20 members spanning the expected range of reactivity and containing no more than 5 negatives. Provide a detailed description of each panel member in the report and explain how the reactivity of all panel members was determined. It is preferable to acquire panel members from naturally infected animals rather than spiking negative diagnostic samples; however, CVB accepts spiked samples as positive panel members if an acceptable justification for their use is provided. You may use negative panel members that are positive for an analyte that the experimental kit is not intended to detect to demonstrate the lack of cross-reactivity.
  - 2. Send the panel to each of three participating laboratories. Each laboratory should test the panel members in each of two prelicense serials. Do not retest panel members with discrepant results between serials. Provide a detailed description in the protocol of the methods used to randomize panel members and to blind personnel to panel member identification.
  - 3. In addition to the panel, participating laboratories should evaluate the suitability of the test kit by testing specimens submitted to their laboratory. This is especially valuable for fresh specimens such as whole blood and fecal samples, particularly when there are processing steps needed for fresh specimens that are not necessary for the prepared test panel members.

## V. SERIAL RELEASE TESTING

Test every serial of a test kit for potency prior to release. The potency test of a diagnostic test kit is an indirect measurement of sensitivity and specificity using quantitative measurements from a panel of test preparations. A serial released to the market should have diagnostic

accuracy consistent with that of the serial or serials used in the pivotal study for establishing the product's diagnostic sensitivity and specificity. Design the potency test to accomplish this by preparing a test panel and setting potency specifications so that the potency test result reveals a problem if a serial does not have adequate diagnostic sensitivity and specificity. VS intends the guidelines provided here to maintain a connection between the diagnostic accuracy demonstrated in the pivotal study, the composition of the potency test panel members, and the potency specifications.

A. Serial Release Test Panel

1. *Panel composition.* Every product code must have a corresponding serial release panel. The panel consists of at least three members: A negative, a weak positive, and a strong positive. For kits intended to detect more than one analyte, additional panel members might be necessary. Supply aliquots of each panel member to the CVB laboratory for confirmatory testing.
2. *Specimen type.* All panel members should be composed of the same specimen type as the intended diagnostic specimens (e.g., blood, serum). For instance, if serum is the approved specimen type, all members including the negative should also be serum. Negative panels should not be sterile diluent, buffer, or any other medium. If the kit is approved for more than one specimen type, select one that is most feasible. If the approved specimen type is impractical for use in a serial release panel (e.g., fecal swabs), propose an alternative composition of the panel members.

B. Release Specifications. For kits that produce an increasing response with an increasing analyte concentration (e.g., S/P ratios increase with increasing concentration), set the specifications as follows. (Reverse as appropriate if the product produces a decreasing response with increasing concentration (e.g., real-time PCR Ct values decrease as concentration increases).)

1. *Negative panel member.* Set the release specification as an upper bound to assure that a new serial does not produce a signal too high that would lead to an increase in the number of false positive test results (i.e., decreased specificity). If an assay produces a numeric response, then the value of the potency specification is the cutoff distinguishing positive from negative samples. If the assay is read visually, then the value of the potency specification are the quantitative instrument measurement that distinguishes visually positive from visually negative.
2. *Weak positive panel member.* Set the release specification as a lower bound to assure that a new serial does not produce too low a signal that would lead to an increase in the number of false negative test results (i.e., decreased sensitivity). The weak positive panel member should be formulated to produce results low enough to detect a possible problem with newly manufactured serial producing too weak a signal.

- a. Set the potency specification so that it is within 3.5 standard deviations of the mean measurement. Alternative methods may be necessary for data that are not symmetric.
  - b. The lower bound for the weak positive panel member is usually the same value as the upper bound for the negative panel member. However, there may be some instances in which formulating the weak panel member to meet the predetermined specification is not feasible, such as when both visual classification and numeric instrument measurements are used for serial release testing. In such cases, the firm should consult with CVB to determine an appropriate means of determining the specification. In such cases, the specification is often based solely on repeated instrument measurements. This usually leads to different serial release specifications for the lower bound of the weak positive and the upper bound of the negative panel member.
3. *Strong positive panel member.* Set the release specification for the strong positive panel member as an upper bound to assure that a new serial does not produce too high a signal that would lead to an increase in false positives (i.e., decreased specificity). The strong panel member serves a similar function as the negative panel member but on the opposite end of the response spectrum.

Using data from testing the strong panel member with the serial or serials used for assay validation, set the specification as the mean plus 3 standard deviations. Alternative methods may be necessary for data that are not symmetric.

4. *Testing during development.* In studies determining an appropriate specification for the strong positive panel member or demonstrating an appropriate formulation of the weak positive panel member, conduct all testing according to the serial release procedure in the Outline of Production. For example, if each panel member is tested in triplicate wells on an ELISA plate, the testing to demonstrate formulation of the weak positive panel member and to determine the specification for the strong positive panel member should also be tested in triplicate wells on each plate.
5. *Serial release test.* For serial release, test each panel member five times according to the Outline of Production. If it is a plate-based kit, use one plate for each of the five measurements, but all panel members may be tested on the same plate, for a total of five plates. A satisfactory potency test result is one in which at least four of the five testing results meets the specification for each panel member.

## VI. CONFIRMATION OF DATING

Conduct a real-time stability study to confirm the initial dating period. VS assigns a dating period of 12 months to test kits before the firm completes a confirmation of dating study. VS will consider requests for longer dating periods if justified.

- A. Test According to Outline. Perform testing as specified in section V.C. of the Outline of Production.
- B. Number of Serials. Three consecutive serials should be used in the confirmation of dating study.
- C. Frequency of Testing. Test each serial at release and at the end of the proposed dating period using the serial release panel and criteria specified in the filed Outline of Production. CVB recommends, but does not require, interim testing. Use the same serial release panel for all tests performed on a given serial.
- D. Outline Statement. Before completing the stability study, include a statement in the Outline of Production that dating has not been confirmed. Once the study has been completed and approved, include the date of approval in the Outline of Production.

## VII. REPLACING PANEL MEMBERS

Prepare serial release panel members in sufficient quantity to conduct serial release testing for years. Replacement panel members will be needed when current panel members become depleted. The original panel members were directly connected to the pivotal study of diagnostic accuracy, but the replacement panel members must be indirectly connected to it because the original serials used in the study are no longer available.

- A. Kits Licensed Under this Guidance. Test the proposed and current panel member simultaneously on available released serials and demonstrate the responses are similar. The serial release specifications will not change for the replacement panel members.
- B. Kits Licensed Before this Guidance. To continue using the current serial release specifications, test the proposed and current panel members simultaneously on available released serials and demonstrate the responses are similar. If requesting changes to the established serial release specifications, submit a protocol before starting the study.

## VIII. CONTROLS TO ESTABLISH A VALID TEST

Some kits may include controls used to determine whether a particular run of the assay is valid. If controls are used for routine testing, they should be used in serial release testing as well. Submit data to support the validity criteria and a complete description of how the validity criteria were established.



**IX. SHIPPING TEST KITS**

To request an exemption to ship a test kit outside the recommended storage conditions specified on the label, submit data to CVB demonstrating that the test kit performs as expected for the temperatures outside the recommended storage conditions. To request an exemption to 9 CFR 114.11 storage conditions during shipment, submit a protocol to CVB before initiating the study. If granted, include the date CVB granted the exemption and the specific shipping conditions in section VI. B. of the Outline of Production.

**X. PROGRAM DISEASES**

The National Veterinary Services Laboratories (NVSL) evaluates diagnostic test kits intended for U.S. Federal and/or State eradication/control programs. CVB provides prelicense serials to NVSL for evaluation. The license and/or permit for such kits may restrict distribution to APHIS-approved laboratories. A U.S. Veterinary Biologics License or Permit does not guarantee the test kit will be used in an official eradication program. Significant changes in disease prevalence over time may affect the diagnostic implication of a test result and, hence, the role of the kit in disease eradication/control programs.