

UC Davis

UC Davis Previously Published Works

Title

Suggested guidelines for validation of real-time PCR assays in veterinary diagnostic laboratories

Permalink

<https://escholarship.org/uc/item/7m75b6gr>

Journal

Journal of Veterinary Diagnostic Investigation, 32(6)

ISSN

1040-6387

Authors

Toohey-Kurth, Kathy
Reising, Monica M
Tallmadge, Rebecca L
[et al.](#)

Publication Date





2020-11-01

DOI

10.1177/1040638720960829

Peer reviewed

Suggested guidelines for validation of real-time PCR assays in veterinary diagnostic laboratories

Kathy Toohey-Kurth,^{1*}  Monica M. Reising,* Rebecca L. Tallmadge, 
 Laura B. Goodman,  Jianfa Bai, Steven R. Bolin, Janice C. Pedersen,
 Mangkey A. Bounpheng, Roman M. Pogranichniy, Jane Christopher-Hennings,
 Mary Lea Killian, Donna M. Mulrooney, Roger Maes, Shri Singh,
 Beate M. Crossley 

Abstract. This consensus document presents the suggested guidelines developed by the Laboratory Technology Committee (LTC) of the American Association of Veterinary Laboratory Diagnosticians (AAVLD) for development, validation, and modification (methods comparability) of real-time PCR (rtPCR) assays. These suggested guidelines are presented with reference to the World Organisation for Animal Health (OIE) guidelines for validation of nucleic acid detection assays used in veterinary diagnostic laboratories. Additionally, our proposed practices are compared to the guidelines from the Foods Program Regulatory Subdivision of the U.S. Food and Drug Administration (FDA) and from the American Society for Veterinary Clinical Pathology (ASVCP). The LTC suggestions are closely aligned with those from the OIE and comply with version 2021-01 of the AAVLD Requirements for an Accredited Veterinary Medical Diagnostic Laboratory, although some LTC recommendations are more stringent and extend beyond the AAVLD requirements. LTC suggested guidelines are substantially different than the guidelines recently published by the U.S. FDA for validation and modification of regulated tests used for detection of pathogens in pet food and animal-derived products, such as dairy. Veterinary diagnostic laboratories that perform assays from the FDA Bacteriological Analytical Method (BAM) manual must be aware of the different standard.

Key words: accreditation; methods comparability; real-time PCR; validation; verification.

Introduction

Real-time PCR (rtPCR) assays have become the workhorse molecular assays in veterinary diagnostic laboratories (VDLs) since ~2000. PCR assays targeting commonly encountered veterinary pathogens are commercially available (standard test methods) but often laboratories choose to develop their own nonstandard test method (laboratory-developed tests, LDTs). Nonstandard rtPCR assays targeted for use in the laboratory may be completely novel, conversions of conventional PCR assays to a rtPCR format, or extensions of a standard test method outside its intended scope (e.g., new specimen matrix). Whatever the scenario, all nonstandard test methods must be validated for use prior to implementation in an accredited laboratory. In addition, a standard test method, or a previously validated nonstandard test method, may need to be modified in order to accommodate sequence changes within the pathogen, incorporation of a new specimen type or species, or to take advantage of new reagents or instrumentation. Regardless of whether the assay is new or modified, it must be demonstrated to be fit-for-purpose through a predefined process.

Consensus guidelines exist for laboratory assays utilized in human medicine and are regulated by the FDA, which oversees commercial assays^{18,45} and has proposed guidelines for LDTs.¹⁸ Validation pathways for human medicine exist

California Animal Health and Food Safety Laboratory, University of California–Davis, San Bernardino (Toohey-Kurth) and Davis (Crossley) branches, CA; Center for Veterinary Biologics, Ames, IA (Reising); Population Medicine & Diagnostic Sciences, Cornell University, Ithaca, NY (Goodman, Tallmadge); Kansas State Veterinary Diagnostic Laboratory, Kansas State University, Manhattan, KS (Bai, Pogranichniy); Veterinary Diagnostic Laboratory, Michigan State University, Lansing, MI (Bolin, Maes); National Veterinary Services Laboratories, Ames, IA (Pedersen, Killian); Texas A&M University, College Station, TX (Bounpheng); Veterinary & Biomedical Sciences Department, South Dakota State University, Brookings, SD (Christopher-Hennings); Oregon Veterinary Diagnostic Laboratory, Oregon State University, Corvallis, OR (Mulrooney); Breathitt Veterinary Center, Murray State University, Hopkinsville, KY (Singh).

* These authors contributed equally to this work.

¹Corresponding author: Kathy Toohey-Kurth, California Animal Health and Food Safety Laboratory, University of California Davis, San Bernardino 105 W Central Avenue, San Bernardino, CA 92408.
klkurth@ucdavis.edu

for both performance⁸ and for publication, such as the MIQE and STARD lists.^{6,7,11,13} Detection of pathogens in veterinary clinical samples is less regulated by federal governments, with oversight provided for some commercial kits through the Center for Veterinary Biologics, U.S. Department of Agriculture (CVB-USDA), but not for LDTs. Thus, it is incumbent upon professional organizations within the veterinary diagnostic community to self-regulate the validation of nonstandard test methods, including LDTs, as well as modifications of validated tests. We provide here an overview of guidance that will help to ensure the quality of rtPCR assays performed in VDLs.

The World Organisation for Animal Health (OIE) has enlisted a consortium of international experts that author the *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*,⁵² which is a guidance document for VDLs. Several chapters in the OIE manual describe the critical elements of de novo assay development and validation,^{53–56} as well as critical elements necessary for modification of a validated assay.⁵⁷ The American Association of Veterinary Laboratory Diagnosticians (AAVLD) is an organization of 50 U.S. states and 9 Canadian provinces whose diagnostic laboratories are dedicated to providing quality laboratory services for food and companion animals. AAVLD provides an accreditation standard that aligns with ISO/IEC 17025:2017²² and uses validation guidelines provided in the OIE manual.⁵³ In addition to AAVLD, ISO 17025, and OIE, 2 other organizations produce standards or guidelines that potentially impact assay validation within a veterinary laboratory. These include guidelines from the American Society for Veterinary Clinical Pathology (ASVCP)^{16,17,43} and the FDA branch that governs U.S. food safety and oversees the actions of the U.S. Food Emergency Response Network (FERN) and the Veterinary Laboratory Investigation and Response Network (Vet-LIRN).⁴⁶

However, there exists a need for more specific guidance covering assay development, modification, and validation of rtPCR test methods. Available guidelines are often considered too expensive to follow,^{8,11,13} not sufficiently specific (OIE),^{53–57} or are not relevant for most of the molecular work performed in AAVLD-accredited VDLs.^{16,17,43,46} The AAVLD Laboratory Technology Committee (LTC) was convened to generate a set of suggested guidelines intended to consolidate the guidance from OIE and to facilitate alignment with the AAVLD Requirements for an Accredited Veterinary Medical Diagnostic Laboratory (<https://www.aavld.org/accreditation-requirements-page>). To this end, the LTC recommendations are described and compared to OIE guidance and the AAVLD accreditation requirements. The LTC provides herein a set of guidelines that allow validation of a de novo assay and assessment of minor changes in previously validated assays, with the goal of providing a cost-effective means of evaluating the impact of minor modifications. Also described are the requirements that the FDA food safety branch additionally imposes on VDLs.

De novo development of rtPCR

Fit for purpose

The OIE manual provides guidance for nucleic acid assay development and subsequent validation in 5 pertinent chapters.^{53–57} Here our goal is to review and mesh the guidance from these chapters into one document and offer practical suggestions from the LTC. The scope and intended use of an assay should be determined prior to the start of the assay design process. At the end of the validation process, the developer must assure that the studies support a statement of validity detailing fitness for the intended purpose.

Terms and definitions

To facilitate the discussion of assay development, a glossary of terms is provided (Table 1) comparing definitions provided by OIE⁵² and AAVLD.² Real-time PCR assay development utilizes the performance criteria of analytical sensitivity and specificity to initially assess the efficacy of the candidate assay on a smaller scale, prior to devoting resources to a full validation. *Analytical sensitivity* is referred to as limit of detection (LOD) and is obtained by “titrating a known strong positive sample to extinction.”⁵³ LOD determined by serial dilutions and plotted with log-transformed concentrations gives an approximate linear relationship and yields additional characteristics of dynamic range, linearity, and slope. The slope, in turn, provides PCR amplification efficiency (AE) using the formula: % Eff = $100(10^{(-1/\text{slope})} - 1)$.^{8,54} *Analytical specificity* allows the developer to test susceptibility to specimen matrix inhibitors (selectivity), ability to detect all relevant strains (inclusivity), and the ability to distinguish targeted organisms from genetic near-neighbors and specimen-specific microflora.

Development

The assay developer must understand the biology of the pathogen(s) in the host(s) and must match the pathogen target sequence characteristics to the intended use of the assay. Target sequence identification and primer/probe design are the most important steps in developing a useful assay, whether it is conversion of a conventional PCR or development of a new assay. The developer must determine whether the assay will target single or multiple pathogens. Multi-target panels require the use of identical annealing and extension times and usually the same components within a master mix; aside from primer concentration, there is little room for optimization. Once the target gene is selected, additional in silico evaluation of the primer and probe sequences is performed using commercial or publicly accessible software.^{15,41,44} It is critical to perform a search of nucleotide databanks (<https://blast.ncbi.nlm.nih.gov>) to determine if target sequences have homology to other agents in addition to the expected target. Design and selection of multiple primer/probe pairs for each target are advised so that the best choice can be made following

Table 1. Glossary of terms and definitions used in assay validation and verification.

Term	OIE glossary ⁵² or reference
Amplification efficiency (AE)*	Empirically derived doubling rate compared to theoretical doubling of 100%. ¹¹
Accuracy†	Nearness of a test value to the expected value for a reference standard reagent.
Analyte/target*	Component of a sample detected or measured by the test method.
Comparability* [‡]	Preferred term to describe a modified test with performance characteristics comparable to the validated test within statistically defined limits. ⁵⁷
Cycle threshold (Ct)*	Cycle number at which the fluorescent signal for an analyte exceeds the background.
Cutoff*	Ct value used to distinguish between negative and positive results on a continuous scale.
Genome copies*	Number of target genomes or genome mimics (plasmid, amplicon, armored RNA, etc.). Using genome copies as synonymous with genome copy equivalents. ⁴⁰
Limit of detection (LOD)*	Smallest detectable amount of analyte in a matrix that would produce a positive result.
Linearity (R^2)*	Capacity of the method to obtain test results proportional to the target concentration.
Precision*	Degree of dispersion within a series of measurements of an identical item.
Range, operating or dynamic*	Span of analyte concentrations over which the method provides suitable precision and accuracy. ¹¹
Repeatability‡	Level of agreement between replicates within and between runs in the same laboratory.
Reproducibility‡	Ability of a test to provide consistent results when the same samples are tested in different laboratories.
Sensitivity, analytical*	Smallest detectable amount of analyte that can be measured with a defined certainty.
Specificity, analytical*	Degree to which the assay distinguishes between the target and other components.
Sensitivity, diagnostic (DSe)†	Proportion of known infected reference animals that test positive.
Specificity, diagnostic (DSp)†	Proportion of known uninfected reference animals that test negative.
Specimen/sample†	Specimen is the material submitted for testing (feces, blood, etc.). Sample is derived from the specimen. (AAVLD states that a sample is used for testing.)
Validation†	Process that determines the fitness of an assay for intended purpose.
Verification§	Objective evidence that accuracy and precision of a validated assay are comparable when used in another laboratory. ⁵⁷

* Not defined in AAVLD glossary.

† Same definition in AAVLD glossary.

‡ AAVLD glossary defines repeatability as occurring on the same day, whereas reproducibility is different days or operators in the same laboratory.

§ Different in AAVLD glossary: internal process to determine if a validated assay performs as expected.

‡ Method comparison is now used primarily for equivalency testing.

empirical testing. Basic design guidance^{10,15,20,31,33,39,42,54} and, more recently, refined primer/probe selection rules are additionally available.¹²

Feasibility

In the feasibility study, a limited panel of 6–8 carefully selected samples are tested^{53,54} in order to limit costs while evaluating which of the candidate primers and probes perform the best. Here the LTC recommendations depart slightly from OIE guidelines, suggesting that the developer initially use a reference strain in buffer as opposed to samples in matrix, until basic efficacy of the primers and probes are established. Determination of the LOD and the operating range requires diluting a very concentrated target suspension through 5–9 ten-fold dilutions in triplicate, which could rapidly deplete a positive sample. The primer/probe set with the highest efficiency, desired LOD, linearity evaluation of R^2 close to 1, and spanning the desired range is selected for further evaluation. Then the dose-response curve is repeated in the most likely matrix rather than in buffer to assure that inhibition (selectivity) is considered, similar to OIE guidance.⁵³ If a partial construct of a gene was used as the target, it is

advisable to test native target at this point to ensure that nucleic acid secondary structure does not impede primer and probe binding.^{12,31} Additionally, performing a melt curve analysis using an intercalating dye such as SYBR green on the selected primers is important to rule out primer–dimer formation empirically.^{10,20,31}

As a next step, the LTC recommends testing 3 or 4 samples of major genetic strains (weak and moderate positives) and 3 or 4 samples containing the most likely commensal competing strains or genetic near neighbors to determine preliminary inclusivity and exclusivity, respectively. It is additionally highly recommended to view PCR as a system and to identify an internal control and extraction method at the outset of the feasibility process.^{10,12,19,24,58} Concurrently, a positive amplification control, optimally using a traceable reference strain,^{2,56} should be identified, and sufficient aliquots of this and the negative (no template) control generated to complete feasibility and validation studies.^{53,54,56} Because rtPCR is either quantitative or semiquantitative, it may benefit a laboratory to calibrate their assay using commercial quantified strains. The OIE suggests use of international or national reference standards to calibrate an assay.⁵⁶ The ATCC provides certified reference strains and nucleic acid

Table 2. Determination of LOD based on 2-fold dilution series for *Mycoplasma cynos* assay.⁴⁰

Genome copy number	80.0	40.0	20.0	10.0†	5.0‡	2.5	1.3	0.6
1	31.3	32.6	32.7	33.9	35.1	37.0	0.0	0.0
2	31.3	32.1	33.7	33.9	34.2	0.0	37.4	0.0
3	31.4	32.9	33.0	34.5	0.0	37.4	38.0	37.4
4	31.3	32.4	33.5	34.6	34.7	37.6	36.0	37.8
5	31.6	32.4	33.6	34.9	36.8	37.6	37.5	0.0
6	32.4	32.4	33.1	33.9	34.4	36.5	0.0	0.0
7	31.4	32.8	32.6	34.4	35.5	36.5	38.5	0.0
8	31.8	32.2	33.1	34.6	35.1	37.2	37.4	0.0
9	31.4	32.5	33.4	34.4	34.4	0.0	0.0	0.0
10	31.4	32.4	33.7	33.5	34.6	35.3	37.5	37.5
11	31.4	32.3	33.4	35.1	34.0	36.6	37.4	0.0
12	31.6	33.1	33.8	33.5	35.8	35.8	37.4	0.0
Average Ct*	31.5	32.5	33.5	34.3	35.0	36.8	37.4	37.6
Proportion positive	1	1	1	1	1	0.92	0.83	0.75

* Average cycle threshold (Ct) for the test wells showing a Ct.

† Limit of detection (LOD) of 10 genome copies and cutoff of 34.3 Ct, determined by 2-fold dilution series.

‡ LOD of 4.6 genome copies determined by probit analysis and cutoff of 35.0 Ct, defined by an average Ct of 5.0 genome copies.

standards (<https://www.atcc.org>). The ATCC site also provides links to commercial firms as well as to other national and international sources of standards. Additional recommended controls for routine performance of PCR are discussed in accompanying articles found in this focus issue.^{44,58}

De novo validation

Once an assay is developed, validation commences and is carried through 4 stages of the validation pathway as outlined in the OIE guidelines.⁵³ In stage 1, analytical sensitivity, specificity, and preliminary repeatability are determined. Stage 2 involves evaluation of the assay using diagnostic samples. Stage 3 is comprised of more extensive repeatability and reproducibility evaluation. Stage 4 involves establishing practical considerations for implementation of the assay.

Stage 1, analytical characteristics

Several methods can be used to determine the LOD, including 10-fold dilution series, 2-fold dilution series,^{53,57} or probit analysis.^{8,24,55} The simplest and most economical method to determine LOD for a rtPCR assay is by utilizing the dilution series that was generated during the analytical sensitivity determination. The final 10-fold dilution for which all 3 replicates are positive provides a conservative LOD estimate provided it was determined in the sample matrix intended for use. Using a hypothetical example, 3 of 3 independent 10-fold dilution series consisting of known target genome copies could detect 10 copies of the target, but only 1 of 3 of the dilution series could detect the target at 1 genome copy. With this example, the LOD would be 10 genome copies.

In general, the number of replicates for LOD determination of a de novo assay validation ranges from 3 to 12 replicates, with the higher number required in some statistical analyses.⁸ The replicates may be 10-fold dilutions or, with a greater expenditure of resources, 2-fold dilutions. An example of a LOD determined by a 2-fold dilution series is provided by the developers of a *Mycoplasma cynos* assay described in this focus issue (Table 2).⁴⁰ Twelve replicates for each of 8 two-fold dilutions gave an LOD of 10 genome copies (i.e., the last dilution in which all replicates were positive; Table 2). These same data can be used for probit or logit analysis by applying generalized linear models to the results of the dilution series experiment.⁸ Based on the concentration-response curve generated (Fig. 1), the concentration in which 95% of targets are positive is 4.6 genome copies (anti-log₂ of 2.2). It is important to note that, in this example, the concentration corresponding to the estimated LOD from the probit analysis was not tested directly.

Cutoff. The cutoff value provides classification of results as either positive or negative and may also include an indeterminate range. The target nucleic acid dilution series detailed above generates a standard curve that allows for accurate quantification of target genome copies. In VDLs, generally a single positive amplification control (PAC) concentration is routinely included for a rtPCR assay, thus limiting the precision of quantification, and making most of the assays used semiquantitative rather than absolute. Thus, the mean cycle threshold (Ct) value of the lowest concentration reliably detected in the dilution series may become the cutoff metric used to interpret diagnostic assay results on a daily basis. For a LOD determined by probit analysis, the cutoff is defined as the mean Ct of the next most concentrated dilution tested to

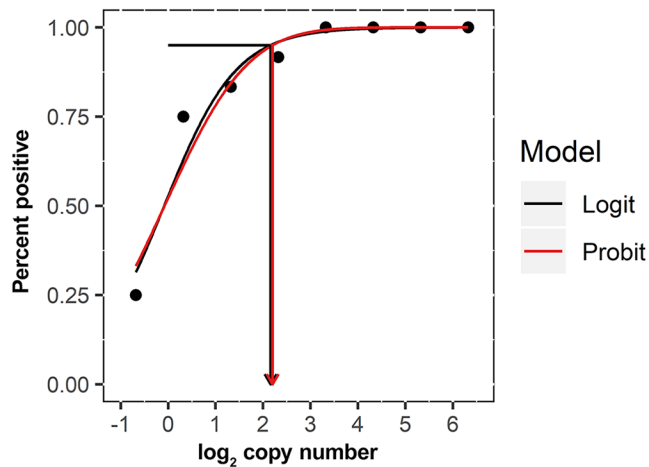


Figure 1. Probit and logit analysis (R Core Team) based on the concentration-response curve generated from the proportion positive results (Table 2). The concentration at which 95% of targets are positive for logit is 4.5 genome copies (95% CI: 3.0, 6.8) and for probit is 4.6 genome copies (anti-log₂ of 2.2; 95% CI: 3.1, 6.9).

the LOD defined by probit. In the example in Figure 1, the average Ct value is 35.0 at a genome copy of 5, which is the tested value at the concentration immediately above the probit value.

Analytical specificity. In the feasibility stage, a preliminary assessment of specificity was performed to aid in vetting the primer selection. Once validation commences, a more thorough evaluation of inclusivity, selectivity, and exclusivity is warranted. Traceable isolates of major strains are used for LOD determinations to assess inclusivity and to be compliant with the AAVLD accreditation requirements. Multiple matrices are evaluated to determine selectivity, and traceable genetic near neighbors of strains of commensals are tested to evaluate exclusivity. Examples of studies with thorough analytical specificity studies are referenced.^{34,40,50}

Repeatability. Repeatability is a measure of agreement among replicates of a sample within runs and between runs in the same laboratory. There are many sources of variation that will affect the outcome or result of an assay within a laboratory, variation among operators being one of these factors. The guidance provided by the LTC recommends that one operator perform this comparison, which differs slightly from the OIE, which recommends that testing should be done by more than one operator.⁵³ To extend the LTC recommendation, additional operators within a laboratory perform repeatability in order to show competence during their training and, through this practice, assay repeatability measures become congruent with the OIE guidelines. The LTC recommends testing 3 concentrations (high, medium, and low), which is consistent with other guidelines including OIE,⁵³ and to test 5 replicates of these concentrations in 6 separate runs.³²

The LTC studies recommended to obtain stage 1 analytical sensitivity, analytical specificity, and repeatability data with acceptance criteria are summarized and compared to OIE guidelines (Table 3). In general, recommendations are similar, although some suggestions from the LTC provide additional guidance. For example, LOD determination involves testing across the operating range of the assay, usually requiring 5–9 log₁₀ dilutions.⁸ Detailed examples of analytical and diagnostic validation are referenced.^{40,50,51}

Stage 2, diagnostic characteristics

In stage 2, diagnostic sensitivity (DSe) and diagnostic specificity (DSp) are determined. Recommendations for the number of samples needed for a full validation are found in the OIE guidance.⁵³ The analytical cutoff should be established from the LOD (determined by 2-fold, 10-fold, or probit/logit methods) then applied to a reference diagnostic sample set in order to verify the cutoff Ct value.

The source and quality of samples used in the validation process is a key element to reliably evaluating diagnostic performance characteristics. The acquisition of samples should be described and fully documented to include source, storage, collection, transfer, transport media, preservative used for storage, and relevant animal information (breed, age, sex, disease status, clinical history).⁵³ Random selection of samples will guarantee good representation and help avoid bias,²⁵ however, it is critical to include representatives of all host species and all sample matrices.^{53,55} Controversy exists as to whether any of the samples used during the diagnostic performance evaluation can be spiked samples (inoculated with known quantities of a pathogen). This is discussed for assays pertinent to wildlife species in a companion article in this focus issue.²³

In the absence of knowledge about the true disease status, statistical methods can be used to estimate DSe and DSp.^{21,30,47} Validating assays for wildlife often require alternative statistical methods, as reviewed in this focus issue.²³

The LTC recommends studies to obtain stage 2 DSe, DSp, and repeatability data with acceptance criteria, as summarized and compared to OIE recommendations in Table 3. Note that in performance of the stage 2 studies, operators should be blinded to the true status of samples. Examples of studies performed using the LTC stage 2 validation guidelines are referenced.^{40,51}

Stage 3, reproducibility

Stage 3 is characterized by establishment of reproducibility, which is the assay characteristic that measures consistency of test results derived from assays performed in different laboratories using aliquots of the same samples.⁵³ This is a critical step for LDTs used outside of the originating laboratory. In order to establish reproducibility, a panel of well-characterized samples representing a range of concentrations and

Table 3. Recommended studies for validation of a newly developed real-time PCR (Stage 1, analytical characteristics; Stage 2, diagnostic characteristics; Stage 3, reproducibility).

Performance characteristic	OIE ^{53,54}	AAVLD LTC
Analytical sensitivity	LOD dilution series of 1:10 or lower. Range (no specified range), linearity, AE. Test in matrix.	Same* test in buffer first, then matrix. Suggest a range of 5–9 ten-fold dilutions. ⁸
Analytical specificity	Inclusivity: test reference strains of major serotypes. Exclusivity: test all near genetic neighbors and matrix commensals. Selectivity: test for detection in the presence of inhibitors.	Same. ⁵⁰
Repeatability	Minimum of 3 reference strains in concentrations spanning the assay range. ⁵³	Same; suggest 5 replicates over 6 d. ^{†19,32}
Diagnostic sensitivity; diagnostic specificity	Determine number of samples based on prevalence of disease and expected DSe and DSp. [‡]	Select well-characterized samples tested on a reference method with infection status of animal known. Numbers tested are based on prevalence. ^{‡53} Internal or interlaboratory panel is acceptable. [§]
Reproducibility	Panel for interlaboratory comparison necessary if the assay is deployed to another laboratory.	Same.

AE = amplification efficiency; LOD = limit of detection.

* Laboratory Technology Committee (LTC) acceptance criteria: at least 90% AE, and $R^2 \geq 0.99$ over full range for linearity, range of 5–9 \log_{10} .

† Calculate standard deviation or coefficient of variation. References are cited for evaluating repeatability.

‡ Calculate DSe and DSp. Acceptable DSe and DSp depends on the purpose but generally will be very high with a well-developed rtPCR assay.

§ Refers to an internal or interlaboratory panel of well-characterized samples previously compared to an accepted reference assay.

all possible matrices is collected and tested using the reference method. These samples should be aliquoted and stored to maintain stability and homogeneity. Assays developed for internal use may be tested using a reference panel and in an alternative laboratory in order to fully align with the OIE pathway. The designated reference panel can also serve as a reference panel for determining accuracy following any subsequent assay modifications. One-time comparison with a reference panel, documentation of “internal or interlaboratory comparison to an accepted methodology and protocol,” together with “ongoing documentation of internal or interlaboratory performance using known reference standards for species and/or diagnostic samples of interest” is one way to fulfill test method validation per version 2021-01 of the AAVLD accreditation requirements.² The LTC recommendations for reproducibility are compared to OIE guidance in Table 3.

Stage 4, implementation

Implementation includes the practical steps of bringing an assay on board: cost accounting, management approval, finalizing the standard operating procedure (SOP), control chart establishment, incorporation into laboratory information management systems, etc. Interpretation criteria should be established prior to launching the assay. Test results must be considered in context of “fitness for purpose.” Factors to

consider beyond the laboratory aspects of test validation include: presence or absence of clinical disease; whether the agent is transboundary, rare, or endemic; possibility of environmental contamination; and the potential of detecting vaccine components. For example, the rtPCR Ct value may be interpreted differently if the animal was known to be recently vaccinated with live⁴⁹ or inactivated vaccines.³⁸

Once the test is launched, assays must be continually monitored by the use of daily in-house reference controls. All guidelines concur on the necessity of having a positive amplification control and a “no template” amplification control for each target; internal controls are strongly recommended. All controls are reviewed in companion articles in this focus issue.^{44,58} Establishment of criteria to monitor assay performance is needed to complete the OIE validation pathway and to meet the AAVLD accreditation requirement.² In addition to daily monitoring of the assay, primer and probe sequences should be monitored routinely to determine whether emerging/novel nucleotide sequence mutations in the target region alter the performance of the assay.¹⁵ As a final note, tests offered to a client must be completely validated, and the test method should be communicated with the client on the test report. If the client requests use of a non-validated test method (e.g., use on a new species or specimen matrix) and the laboratory conducts such testing, there must be a disclaimer statement added to the test report in order to be compliant with AAVLD accreditation requirements.

Box 1. Minor modifications of a PCR assay.

Master mix or extraction reagent replacements
 Change in probe labeling
 PCR or extraction platform changes
 Robotic handling

Method comparability

A validated assay in routine use may need to be modified. In keeping with the OIE naming convention,⁵⁷ the LTC refers to this as method comparability rather than comparison. Modifications include reagent replacement, platform changes, or extraction modifications.^{32,57} The LTC considers changes to the primer or probe sequence to be major changes, although OIE guidelines suggests that changes within the same amplified region are minor. The OIE guidance considers species or matrix additions to be major modifications to an assay. Box 1 shows the consensus between OIE and LTC designations of minor modifications. Under optimal circumstances, a direct comparison between the reference procedure and the modified procedure, hereafter referred to as the candidate method, is desirable. The procedures described are intended to make a direct comparison, but without the stringency of demonstrating statistical equivalence, which would require defining margins that are related to irrelevant scientific differences.^{3,37} The recommended testing should reveal any egregious changes to the performance characteristics of the modified assay. The LTC recommended method comparability studies for acceptance of a modified rtPCR assay are compared to OIE guidance in Table 4.

Method comparability analytical characteristics. An analytical sensitivity study is recommended using up to 3 reference strains,³² which is a reduction in resources from the requirement to test all major strains in a de novo assay. The 3 reference strains selected should span the sequence diversity.³² An analytical specificity study is not necessary but can be performed optionally unless a change of the sample matrix or species is evaluated, which would necessitate specificity testing.

The repeatability of the assay is assessed by conducting a study to estimate the intra- and inter-assay variance using one reference strain. Testing 3 concentrations, as recommended for a de novo assay, is one of the most resource-intensive steps in the validation process, with little value realized from testing strong and mid-range concentrations. Thus, our recommendation is to test a low concentration that is within the linear range of the assay but near the LOD. The test material is prepared in advance in bulk such that 1 aliquot is tested 5 times on each of 6 independent testing occasions^{32,48} by 1 operator. To analyze the findings statistically, a linear mixed-effect model²⁹ can be used to estimate the inter-assay variance (between testing occasion) and the intra-assay variance (within testing occasion variance).^{9,48}

To analyze the results with more readily accessible statistics, standard deviation or coefficient of variation can be used. Examples of acceptable standard deviation and coefficient of variation are referenced in Table 4.

Method comparability accuracy. Sample size determination is not stated for comparability; however, a schematic suggests 8 negatives and 24 positives in duplicate is sufficient;⁵⁷ a panel of 30 positive and 30 negative samples for assay evaluation is recommended by the LTC and is referenced in the OIE guidance.⁵⁶ Ideally, these samples are field diagnostic samples that are tested using the reference assay method (gold standard) because these can then become the reference panel needed to be congruent with the OIE guidance. The intention is to test enough samples to span the operating range of the assay. The operator should be blinded to sample status. In a side-by-side comparison, aliquots from the same sample should be tested concurrently (in the same test session) using the reference and candidate method. To assess the findings, the data can be plotted using a square correlation plot^{27,28} with modification³² or a Bland–Altman plot.^{4,5,26} The paired mean difference can be estimated when appropriate. For rtPCR-based assays, this would be reduced to the number of samples in which a Ct was observed using both methods. The correlation and concordance correlation coefficients along with confidence intervals can also be estimated using the samples that produce a response for both methods.²⁸ Finally, kappa can be estimated, and a confidence interval provided for the assay comparison.¹

A primary impetus for formation of the LTC was to generate guidelines that limit the expenditure of resources needed to re-validate an assay. To this end, side-by-side testing may be unnecessary because of either the nature of the modification or use of a reference panel. For example, cross-contamination analysis is necessary when robotic equipment is evaluated to replace manual procedures but would be difficult to set up in a parallel situation performed manually. In other cases, a well-characterized panel of samples already tested by the reference method could be tested with the candidate method without repeating the reference testing concurrently, as long as care is taken to preserve the integrity of the original sample. Values near the cutoff may be lost in the reference sample due to repeated freeze–thaw cycles, so it is important that aliquots of the reference panel undergo the same freeze–thaw and storage conditions.⁵⁶ Any discrepancy between the candidate and reference assay would require repeating all samples within a similar quantitative range side-by-side. This reference panel may be a set of internal well-characterized samples, or a panel generated by interlaboratory collaboration, both of which are an option for validation according to the AAVLD standard.² It is expected that the candidate assay yields the same or better Ct value than the reference assay for the panel samples.

Given the limited number of samples used, a method comparability experiment is not intended to estimate the DSE

Table 4. Recommended method comparability studies for acceptance of a modified real-time PCR assay.

Performance characteristic	OIE ^{53,54,56,57}	AAVLD LTC
Analytical sensitivity	LOD dilution series of 1:10 or lower. Range (no specified range), linearity, AE. Test in matrix.	Same*; OK to limit to 3 reference strains representing strain diversity. ³²
Analytical specificity	Test for specificity by including samples that are positive for potentially cross-reacting agents.	Most comparability studies will not require specificity testing. Different species or matrix will require specificity.
Precision	Repeatability minimum of 3 reference standards in concentrations spanning the assay range.	A single reference strain at 1 log ₁₀ more concentrated than the LOD. [†] 5 replicates over 6 d. ^{24,32}
Accuracy	Schematic in one chapter ⁵⁷ indicates 8 negatives and 24 positives in duplicate. Chapter on reference panels ⁵⁶ suggests 30 positives and 30 negatives. Positives should span the dynamic range.	30 positives, 30 negatives, well-characterized samples tested by a reference method with known infection status of animal. ^{‡§}
Reproducibility	Panel for interlaboratory comparison necessary if the assay is deployed to another laboratory.	Same.

LOD = limit of detection; LTC = Laboratory Technology Committee.

* LTC acceptance criteria: at least 90% AE, and $R^2 \geq 0.99$ over full range for linearity. Candidate method should be comparable or more sensitive than the reference method.

[†] Calculate standard deviation³² or coefficient of variation.²⁴

[‡] Diagnostic samples should be comparable to or more sensitive using the candidate method. Visualize using a Bland–Altman plot.

[§] Refers to an internal or interlaboratory panel of well-characterized samples previously compared to an accepted reference assay.

and DSp of an assay. As such, the findings should be summarized. Should the modified assay result in a more sensitive assay, the analyst could sequence the amplicon to resolve whether the additional detections are specific to the target, as discussed in a companion article in this focus issue and elsewhere.^{14,15}

Verification

Most guidelines, whether ASCVP, OIE, or FDA, have essentially the same definition of verification; however, different amounts of testing are required to complete a verification study. Verification studies typically demonstrate that an assay performs as expected when newly adopted without changes in a laboratory. Implementation of a commercial kit or of a published assay are examples of verification. Verification studies are sufficient provided the assay was fully validated by the company or the authors of the published article. The original validation data for a kit should be available from the company; published articles should have followed one of several accepted guidelines for validation studies.^{6–8,11,53} To perform a verification, analytical sensitivity (LOD) is recommended by the LTC and performed by testing 10-fold dilutions of a single reference strain in matrix; no further analytical sensitivity or specificity testing is suggested. Accuracy for verification as interpreted by the LTC involves testing a minimal number of well-characterized diagnostic samples, specifically 30 positives and 30 negatives previously assessed by a reference method.⁵⁶ Although a verification study is appropriate to introduce a USDA-regulated commercial kit into use in a laboratory, it should be

noted that these guidelines are not intended to be used to validate changes to commercial kits that are licensed by the CVB-USDA. The kit must be used as approved by the CVB.

Note that verification as defined by AAVLD accreditation requirements includes studies performed to obtain objective evidence to show that a changed standard method can be performed properly. The LTC recommendations concur that verification is needed for a changed standard method that comes from a governing body. However, if the modification is initiated within the laboratory, then additional work, more in line with method comparability, is recommended.

Special cases of comparability

Addition of an internal control

It is highly recommended that rtPCR assays incorporate either an exogenous or endogenous internal control. That said, many laboratories do not currently have these incorporated in their assays for economic reasons. Use of an internal control is a special case of comparability testing because the assay is modified from a singleplex to a duplex assay. An example of incorporation of an internal control is given in a companion article.⁵⁸

Conversion of singleplex to multiplex

Real-time PCR assays with multiple targets provide cost savings and can be performed without a loss of sensitivity. The researcher can assess the process (extraction and PCR assay) starting with stocks of pathogens or can assess the assay itself

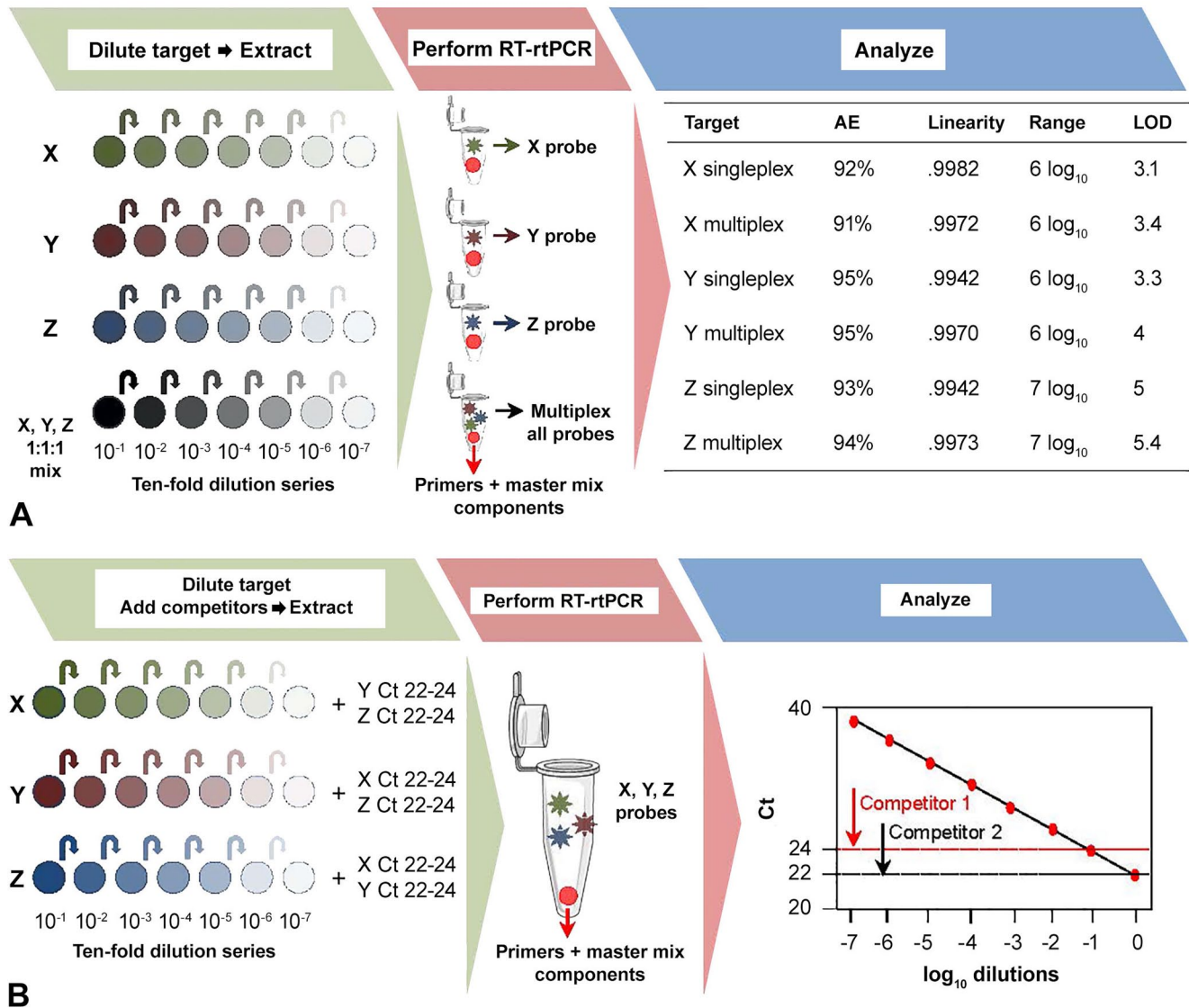


Figure 2. **A.** A method comparability study format for the conversion of 3 singleplex assays for detection of targets X, Y, and Z into a multiplex assay. Representative data is presented from 10-fold dilution series created from viral or bacterial stock for each target alone and in combination. The plate with dilutions is extracted, and the purified nucleic acid is used to perform reverse-transcription real-time PCR (RT-rtPCR). Results from the assays as a singleplex and in multiplex format are compared using performance criteria of amplification efficiency, linearity, operating range, and limit of detection (genome copies). **B.** To test the effect of concentrated co-targets, 10-fold dilution series of each stock target are spiked with a pool of competitor targets added at a strong concentration (20–24 Ct). The plate is extracted, and RT-rtPCR is performed. A linear regression is shown to better illustrate the setup; however, all assay performance characteristics (i.e. AE, R^2 , range, LOD), with and without competitor targets, should be compared.

starting with nucleic acid. Here we give an example starting at the extraction step. Careful choice of extraction kits allows co-extraction of DNA and RNA agents. Inclusion of a reverse-transcription step allows co-amplification of DNA and RNA agents as well. Each assay is validated initially as singleplex and then combined in a multiplex. The critical parameter to test is whether a single component assay is still functional across its established range in the presence of concentrated amounts of potentially competing assays. Each singleplex assay is compared with the corresponding assay in the multiplex format and

also compared to an assay with competing targets and assessed for performance criteria of dynamic range, linearity, and LOD (Fig. 2). Precision is determined by testing replicates of a single concentration in the multiplex series using the dilution closest to 33–34 Ct for all targets (e.g., ~1 log more concentrated than the LOD). Sufficient volume of the chosen dilution should be prepared to complete all necessary testing, and aliquots should undergo the same storage and freeze–thaw conditions.³² Each mixture is tested 5 times within each of 6 independent runs and analyzed as described previously.³²

Table 5. Multiplex assays modifications requiring method comparability.

Type of modification	Mechanism
Combining additional validated singleplexes into existing multiplex	Perform a de novo validation of the singleplex (if unvalidated). Then perform a method comparability between the candidate assay as a singleplex and as part of a multiplex. Also, assess the impact of the original multiplex to the modified multiplex.*
Performing multiplex assay without 1–2 singleplex assays component	Method comparability performance of the partial multiplex to the complete multiplex.*
Change in PCR chemistry, platform, or robotics, or slight variation in nucleotide sequence or fluorophore	Method comparability of original multiplex to candidate multiplex.*

* Can also utilize an inter- or intra-laboratory reference panel to demonstrate accuracy for the modified multiplex assay.

A reference set of well-characterized diagnostic samples, ideally 30 positive samples with various target concentrations and 30 negative samples per analyte,⁵⁶ is used to assess the diagnostic performance of the multiplex assay. The number of reference samples can be reduced if specimens from animals with natural polymicrobial or multiple viral infections are tested to verify the multiple-pathogen detection capability of the assay. Also, to be congruent with the AAVLD standard, an internal or interlaboratory panel of well-characterized samples can be used without the necessity of a side-by-side comparison. The recommendation from the OIE guidance is to test a subset panel (5–8 of each species or matrices) when faced with a potentially large sample set.⁵⁷ The modified assay is considered acceptable if the results of the subset tested with the candidate method are equal to or better in performance compared to the reference method. Methods for resolution of discordant samples should be identified at the beginning of the study as described earlier. Secondary assays such as sequencing can be used to confirm specificity of positive samples. Cutoffs should be the same for each assay in singleplex or multiplex format. Multiplex assays developed by AAVLD member laboratories using the OIE and LTC guidelines are referenced.^{34,35}

Modification of multiplex assays

Special cases of multiplex re-validation may arise (Table 5). Changes made to a validated multiplex assay, such as PCR chemistry, platform, minimal nucleotide changes, etc., may be assessed as described previously using a methods comparability study.³² Host species and matrix introductions may require a more extensive re-validation, which can first be assessed using a reference panel of well-characterized samples.⁵⁷ Dependent on the modification, the testing may utilize a methods comparability assessment of candidate multiplex to the original multiplex assay. Other types of re-validation include incorporation of a newly developed singleplex into a validated multiplex, incorporation of additional validated singleplex(s), or performance of a partial multiplex.

Method modification guidelines from FDA

Most state VDLs provide testing for foodborne disease pathogens in animal-derived products or feeds and work

Box 2. FDA levels of validation.

- L1: 1 Laboratory emergency
- L2: 1 Laboratory non-emergency
- L3: ≥ 2 Inter lab validation
- L4: 8–10 Inter lab validation

closely with their public health counterparts in FERN and Vet-LIRN. Both of these organizations utilize methods found in the U.S. FDA's Bacteriological Analytical Method (BAM) manual. The FDA Food Programs Governance Board regulates the methods used in food safety and is a different branch of government than that which regulates human diagnostic laboratory procedures.

As noted throughout the description of de novo validation, the LTC recommendations closely follow those of the OIE. The U.S. FDA recommendations, however, are different, largely because of the nature of "samples" with a nearly unlimited number of matrices. With food items, spiking in order to simulate a contaminated food is the only option, given that no archive exists that could accurately represent all of the food products subject to spoilage and subsequent outbreak testing needs. Compared to the OIE and LTC guidance, requirements for validation of an assay are more exacting, whereas requirements for modification such as extension to a new matrix are less prescriptive.

The most likely situation in which a veterinary laboratory may need to follow U.S. FDA guidance is when testing a food product associated with a submission from an outbreak, but the food product is not a validated matrix. Recent guidelines⁴⁶ outline requirements for validation by a single laboratory during emergency situations (L1), or non-emergency (L2) situations, and validation for use in ≥ 2 laboratories (L3) or 8–10 laboratories (L4; Box 2). Most VDLs would modify BAM methods for use with a new food matrix using level 1 or 2 criteria. A comparison between the LTC recommendations for assays targeting animal pathogens and FDA requirements for a modification to BAM methods in a single laboratory (L2) is shown (Table 6) for prokaryotic pathogens and with non-microbial targets (FDA equates "microbial" to prokaryotes only). Acceptance of the modification would

Table 6. U.S. FDA requirements for studies to assess performance for modified BAM assays.

Performance characteristic	U.S. FDA requirements
Analytical sensitivity	Not required.
Analytical specificity	Not required.
Precision	Use pure cultures to spike samples. Very specific about reference strains for spiking.* 2 spiked concentrations (at 50% LOD and 1 log more concentrated) and 1 uninoculated. Prokaryotic targets: 20 at LOD ₅₀ of agent in reference method (BAM) 5 for 1 log above 5 for each un-inoculated matrix Non-microbial† targets: test ≥6 replicates if 1 matrix; ≥2 replicates if 2 or more matrices tested.
Evaluation of accuracy	Optional to perform an interlaboratory comparison. May be required if the modified assay has a broad distribution that would increase the validation level to L3 or L4.

BAM = Bacteriological Analytical Method manual.

* U.S. FDA: <https://www.fda.gov/media/83812/download>

† U.S. FDA definition of microbial equates to prokaryotic (bacteria); non-microbial would be parasites, fungi, viruses, etc.

require at least 50% of the replicates at the LOD, that all of the replicates at the higher concentration were detected, and that no target was detected in the uninoculated samples. Note that adherence to the FDA guidelines are requirements not suggestions. Adhering to both FDA and the AAVLD standard is possible but ultimately will be determined by an individual laboratory's quality system.

American Society for Veterinary Clinical Pathology

The American Society for Veterinary Clinical Pathology (ASVCP) also provides guidance for clinical pathology laboratories that are usually housed within a veterinary school. Disciplines covered include hematology, cytology, clinical chemistry, toxicology, and serology. ASVCP has been regularly active in providing guidelines for their members;^{16,17,43} however, the disciplines targeted are quite different from the primarily infectious disease testing that a state VDL performs, with very little rtPCR performed in a clinical pathology laboratory. Genetic testing for breed identification or tumor markers is within the purview of clinical pathology; however, few if any laboratories abiding by ASVCP guidelines provide this service. Guidelines have been adopted for canine clinical genetic testing laboratories from human genetic testing and are in a nascent stage,³⁶ with genetic testing usually performed in private laboratories.

Alignment with AAVLD accreditation requirements

The AAVLD requirements for validation are summarized here. We reiterate the comparison with emphasis on how the LTC recommendations comply or extend beyond the AAVLD accreditation requirements.

First, all test methods validated and in use in a laboratory must be constantly monitored by use of a traceable reference

control through ongoing documentation of internal or inter-laboratory performance. The LTC recommendations are in accordance with this requirement by requiring a reference PAC for each target assay.

Second, one of the following 3 criteria, paraphrased here, must also be met:

- endorsed or published protocol by a reputable technical organization;
- published peer-reviewed article with sufficient documentation; or
- comparison to an accepted method through internal or interlaboratory comparison.

Implementation of the USDA National Animal Health Laboratory Network (NAHLN) assays are an example of the first bullet point. The LTC recommendation and AAVLD accreditation standard concur in that a reference panel, such as a required proficiency test, would be sufficient for verification prior to implementing a NAHLN assay. The second bullet point, adoption of a published assay, would require verification, which could be accomplished by testing an inter- or intra-laboratory panel of well-characterized samples or by using another process that can objectively demonstrate accuracy of the method in obtaining the expected results. The third bullet point applies to either de novo or method comparability. The LTC recommendations again fit with the AAVLD accreditation requirements as long as testing of an inter- or intra-laboratory panel of well-characterized samples used an accepted methodology and generated objective evidence confirming fitness for purpose.

Acknowledgments

We thank the number of individuals who contributed to the discussions and the development of these guidelines over a number of years. We specifically thank Amy Glaser, Deep Tewari, Beth

Harris, Kate Schumann, Susanne Hinkley, Pam Ferro, and Megan Schroeder. We also thank Mauricio Navarro for his expert help with Figure 2.

Declaration of conflicting interests

The authors declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.


Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Kathy Toohey-Kurth  <https://orcid.org/0000-0002-6212-3263>

Rebecca L. Tallmadge  <https://orcid.org/0000-0002-7466-5449>

Laura B. Goodman  <https://orcid.org/0000-0002-8327-3092>

Beate M. Crossley  <https://orcid.org/0000-0003-2932-7229>

References

- Agresti A. Kappa measure of agreement. In: *Categorical Data Analysis*. 2nd ed. Wiley, 2002:434–435.
- American Association of Veterinary Laboratory Diagnosticians, Inc. Requirements for an Accredited Veterinary Medical Diagnostic Laboratory, AC1, Version 2021-01. [cited 2020 Aug 10]. <https://www.dropbox.com/s/bda59w546nr1g3k/AAVLD%20Requirements%20for%20an%20Accredited%20Veterinary%20Medical%20Diagnostic%20Laboratory%20AC1%20V%202021.01.Final.WebPost.pdf?dl=0>
- Armitage P, Colton T. Equivalence trials. In: *Encyclopedia of Biostatistics*. 2nd ed. Wiley, 2005:1735–1740.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310.
- Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135–160.
- Bossuyt PM, et al. Toward complete and accurate reporting of studies of diagnostic accuracy. The STARD initiative. *Am J Clin Pathol* 2003;119:18–22.
- Bossuyt PM, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology* 2015;277:826–832.
- Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. *Clin Microbiol Rev* 2010;23:550–576.
- Burdick R, et al. Design of gauge R&R experiments. In: *Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random and Mixed ANOVA Models*. Society for Industrial and Applied Mathematics, 2005:53–63.
- Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* 2004;15:155–166.
- Bustin SA, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 2009;55:611–622.
- Bustin SA, Huggett J. qPCR primer design revisited. *Biomol Detect Quantif* 2017;14:19–28.
- Bustin SA, Wittwer CT. MIQE: a step toward more robust and reproducible quantitative PCR. *Clin Chem* 2017;63:1537–1538.
- Cortes-Hinojosa G, et al. Development and validation of a quantitative PCR for rapid and specific detection of California sea lion adenovirus 1 and prevalence in wild and managed populations. *J Vet Diagn Invest* 2017;29:193–197.
- Crossley BM, et al. Guidelines for Sanger sequencing and molecular assay monitoring. *J Vet Diagn Invest* 2020;32:767–775.
- Flatland B, et al. ASVCP quality assurance guidelines: control of general analytical factors in veterinary laboratories. *Vet Clin Pathol* 2010;39:264–277.
- Flatland B, et al. Analytical quality goals—a review. *Vet Clin Pathol* 2018;47:527–538.
- Genzen JR. Regulation of laboratory developed tests—a clinical laboratory perspective. *Am J Clin Pathol* 2019;152:122–131.
- Goris N, et al. Validation of two real-time RT-PCR methods for foot-and-mouth disease diagnosis: RNA-extraction, matrix effect, uncertainty of measurement and precision. *J Virol Methods* 2009;160:157–162.
- Hoffmann B, et al. A review of RT-PCR technologies used in veterinary virology and disease control: sensitive and specific diagnosis of five livestock diseases notifiable to the World Organisation for Animal Health. *Vet Microbiol* 2009;139:1–23.
- Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics* 1980;36:167–171.
- International Organization for Standardization (ISO). General requirements for the competence of testing and calibration laboratories. ISO/IEC standard 17025:2017. 3rd ed. ISO, 2017. <https://www.iso.org/standard/66912.html>
- Jia B, et al. Validation of laboratory tests for infectious diseases in wild mammals: review and recommendations. *J Vet Diagn Invest* 2020;32:776–792.
- Kralik P, Ricchi M. A basic guide to real time PCR in microbial diagnostics: definitions, parameters, and everything. *Front Microbiol* 2017;8:108.
- Longford NT, Nelder JA. Statistics versus statistical science in the regulatory process. *Stat Med* 1999;18:2311–2320.
- Liao JJZ, Capen R. An improved Bland-Altman method for concordance assessment. *Int J Biostat* 2011;7:1–17.
- Lin JS, Wei LJ. Linear regression analysis based on Buckley-James estimating equation. *Biometrics* 1992;48:679–681.
- Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;45:255–268.
- McCullagh P, Nelder JA. Dilution assay. In: *Generalized Linear Models*. Chapman & Hall, 1989;2:11–14.
- Qu Y, et al. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics* 1996;52:797–810.
- Raso A, et al. Troubleshooting fine-tuning procedures for qPCR system design. *J Clin Lab Anal* 2011;25:389–394.
- Reising MM, et al. Methods comparison. *J Vet Diagn Invest* 2014;26:480–487.
- Rodríguez A, et al. Design of primers and probes for quantitative real-time PCR methods. *Methods Mol Biol* 2015;1275:31–56.

34. Schroeder ME, et al. Development and performance evaluation of calf diarrhea pathogen nucleic acid purification and detection workflow. *J Vet Diagn Invest* 2012;24:945–953.
35. Schroeder ME, et al. Development and performance evaluation of a streamlined method for nucleic acid purification, denaturation, and multiplex detection of bluetongue virus and epizootic hemorrhagic disease virus. *J Vet Diagn Invest* 2013;25:709–719.
36. Shaffer LG, et al. Quality assurance checklist and additional considerations for canine clinical genetic testing laboratories: a follow-up to the published standards and guidelines. *Hum Genet* 2019;138:501–508.
37. Smith B. The equivalence problem. In: *Encyclopedia of Biopharmaceutical Statistics*. Marcel Dekker, 2000:45–46.
38. Steinrigl A, et al. Bluetongue virus RNA detection by RT-qPCR in blood samples of sheep vaccinated with a commercially available inactivated BTV-8 vaccine. *Vaccine* 2010;28:5573–5581.
39. Suslov O, Steindler DA. PCR inhibition by reverse transcriptase leads to an overestimation of amplification efficiency. *Nucleic Acids Res* 2005;33:e181.
40. Tallmadge RL, et al. Characterization of a novel *Mycoplasma cynos* real-time PCR assay. *J Vet Diagn Invest* 2020;32:793–801.
41. Thornton B, Basu C. Real-time PCR (qPCR) primer design using free online software. *Biochem Mol Biol Educ* 2011;39:145–154.
42. Thornton B, Basu C. Rapid and simple method of qPCR primer design. *Methods Mol Biol* 2015;1275:173–179.
43. Tomlinson L, et al. Best practices for veterinary toxicologic clinical pathology, with emphasis on the pharmaceutical and biotechnology industries. *Vet Clin Pathol* 2013;42:252–269.
44. Toohey-Kurth K, et al. Best practices for performance of real-time PCR in veterinary diagnostic laboratories. *J Vet Diagn Invest* 2020;32:815–825.
45. U.S. Department of Health and Human Services, Centers for Medicare and Medicaid Services. Clinical Laboratory Improvement Amendments of 1988 (CLIA). *Fed Regist* 1992; 57:7002–7186. [cited 2020 Aug 10]. <https://www.govinfo.gov/content/pkg/FR-1992-02-28/pdf/FR-1992-02-28.pdf>
46. U.S. Food and Drug Administration (FDA). Guidelines for the Validation of Analytical Methods for Nucleic Acid Sequence-Based Analysis of Food, Feed, Cosmetics and Veterinary Products. 1st ed. FDA, 2019. [cited 2020 Aug 10]. <https://www.fda.gov/media/121751/download>
47. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics* 1985;41:959–968.
48. Vardeman S, Jobe J. Gauge R&R studies. In: *Statistical Quality Assurance Methods for Engineers*. 1st ed. Wiley, 1999:20–23.
49. Walz PH, et al. Virus detection by PCR following vaccination of naive calves with intranasal or injectable multivalent modified-live viral vaccines. *J Vet Diagn Invest* 2017;29:628–635.
50. Warg JV, et al. Detection and surveillance of viral hemorrhagic septicemia virus using real-time RT-PCR. I. Initial comparison of four protocols. *Dis Aquat Organ* 2014;111:1–13.
51. Warg JV, et al. Detection and surveillance of viral hemorrhagic septicemia virus using real-time RT-PCR. II. Diagnostic evaluation of two protocols. *Dis Aquat Organ* 2014;111:15–22.
52. World Organisation for Animal Health (OIE). *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. 8th ed. OIE, 2019.
53. World Organisation for Animal Health (OIE). Principles and methods of validation of diagnostic assays for infectious diseases. Chapter 1.1.6. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. 8th ed. OIE, 2019.
54. World Organisation for Animal Health (OIE). Development and optimization of nucleic acid detection assays. Chapter 2.2.3. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. 8th ed. OIE, 2019.
55. World Organisation for Animal Health (OIE). Statistical approaches to validation. Chapter 2.2.5. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. 8th ed. OIE, 2019.
56. World Organisation for Animal Health (OIE). Selection and use of reference samples and panels. Chapter 2.2.6. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. 8th ed. OIE, 2019.
57. World Organisation for Animal Health (OIE). Comparability of assays after changes in a validated test method. Chapter 2.2.8. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. 8th ed. OIE, 2019.
58. Yan L, et al. Inhibition monitoring in veterinary molecular testing. *J Vet Diagn Invest* 2020;32:758–766.