The quality assurance of proficiency testing programs for animal disease diagnostic laboratories

Ann L. Wiegers

Abstract. Laboratory data credibility has 3 major components: 1) valid methods, 2) proficiency testing (PT) to verify that the analyst can conduct the method and to compare results of other laboratories using the same method, and 3) third-party accreditation to verify that the laboratory is competent to conduct testing and that the method validation has been done within the environment and requirements of an effective quality-management system. Participation in external PT programs by a laboratory is strongly recommended in International Organization for Standardization/International Electrotechnical Commission International Standard 17025. Most laboratory accreditation bodies using this standard require that laboratories participate in such programs to be accredited. Internal PT is also recommended for each analyst. Benchmarking, or comparison between laboratories using PT or reference materials, is also recommended as part of the validation and evaluation of test methods. These requirements emphasize the need for proficiency test providers to demonstrate their competence. Requirements for competence are documented in national and international standards and guidelines, and accreditation is available for providers. This article discusses the activities and the components that are necessary and recommended for PT projects and programs for animal disease diagnostic testing. These are based on the requirements of the national and international standards, which address this subject, and on the experience of the author. The accreditation of external PT programs is also discussed. Organizations that accredit PT providers or that provide PT programs are listed. Existing references, guidelines, and standards that are relevant to PT in veterinary diagnostic laboratories are discussed.
Proficiency testing may also be used to assign values to reference materials or to determine their suitability for use. Proficiency test samples may, in many cases, also be considered reference materials. A type of reference standard, a “reference material” is a substance of 1 or more properties that are sufficiently well established to be used for calibration, assessment, or assigning values (e.g., a known positive, something from your reference collection such as a fixed specimen or a picture, a reference that comes with the diagnostic equipment, or a reference culture, sometimes referred to in animal health diagnostic laboratories as a “standardized diagnostic reagent”). Therefore, PT may be useful in harmonizing methods and in resolving disparities in the results obtained by different laboratories. For this reason and for its usefulness in ensuring and verifying that a laboratory can produce a valid test result, PT can be critical to the acceptance of test results for trade and regulatory purposes. The International Cooperation for Laboratory Accreditation (ILAC), and therefore laboratory accreditation bodies (ABs), has produced comprehensive guidance and requirements for providers of PT programs, and PT programs are now considered to be an integral part of a national laboratory system. Hereinafter, PT samples are referred to as PT “test materials,” “test items,” or “panel items” to clearly distinguish these from actual “sample submissions.” Note that a test material, as with a reference material, may be prepared and held in bulk, but an “item” is a single unknown.

Proficiency testing for animal disease (veterinary) diagnostic laboratories is essential to effectively facilitate and document the transfer of the test method from the development phase to its use at the bench, transfer from one laboratory to another, to validate, evaluate, and compare methods, to verify successful training, to assess precision, and to facilitate trade and regulatory approvals and decisions. Although PT captures only a moment in time, it is an important measure of a laboratory’s capability in many areas. Participation in PT is a strong guideline of International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) International Standard 17025. Accreditation bodies in the United States currently view this participation as a requirement. Therefore, veterinary laboratories seeking accreditation will be required to participate in PT programs.

In discussing the QA of PT, it is important to note that relevant national and international standards emphasize the following 4 issues. The PT project or program designer and quality manager should be ever mindful of these issues when implementing QA in PT.

1. Proficiency tests (test materials, their values, and criteria for pass or fail) are produced using test results.
2. Test results used to create panels should be done by a laboratory considered to be competent and technically expert.
3. Proficiency test providers and their programs produce a distinct product: proficiency tests. As for any product, certain elements and activities are critical to the consistent production of a satisfactory result. These elements and activities must be considered and addressed with appropriate specifications, procedures, and policies. Therefore, the PT program itself should operate under a quality-management system, which covers both technical and operational requirements. For a client to have adequate confidence in PT results, or for a user laboratory to have confidence in a PT program, third-party accreditation of the PT program may be necessary.
4. The laboratory producing PT results has created, in the PT test material, a reference material. Several international standards discuss and apply to the production of reference materials.

**Determination of the needs and uses of the PT project or program**

The first QA step in the design of a product is to accurately determine client (user) needs and expectations. In PT program design, a determination of use will assist in the determination of needs. To implement the desired level of quality in the design and in the production process, it is necessary to know exactly how the product will be used. Taking the time to identify all the issues, formulate the right questions, and determine the answers are the most important QA activities to be done in PT design. The measurements made by the PT will only be as useful as the validity and suitability of the measurement system, calculations, and the statistics used to create the PT test materials, the PT panels, and the policies and the procedures used to execute the PT program. “Panels” are the uniquely identified set of test items that are to be used in the PT.

The following should be considered, especially with regard to the design aspects of frequency, number, composition, and arrangement of panel items:

- The specific activity for which the PT is to be used (e.g., verification of training, transfer of work from one laboratory to another, verification of proficiency to conduct the test, assessment of precision among technicians, detection of bias or nonrandom error, method evaluation or comparison, official analyst approval, official laboratory approval, official method approval).
The acceptable risk, or the level of due diligence expected by the client of the PT (the user of PT results).

- The impact of the risk being realized as an adverse event, for example, the number of test materials needed per panel, how frequently the panels are sent out, and what constitutes a pass or fail on the panel testing.
- Other kinds of evaluations required, for example, whether the laboratory is accredited or whether it has its own internal PT program.
- The data needed and whether these will be adequate for the statistical methods used to obtain the statistics used in formulating the needed results and answers.
- The statistical methods used to obtain the answers from the data, for example, a PT panel designed for a specific use is sometimes adopted for another use, for which the panel is not suitable. One might want a number of positives and negatives for a training PT, but one may not need negatives for a laboratory transfer of a quantitative test when one wishes the PT to detect bias between the laboratories. Negatives have nothing to measure and in many cases should be treated as a separate population from positives, especially when establishing pass or fail criteria for PT panels.

There are many more questions, not as directly affecting panel design, that should be asked. These include the level of technical complexity of the test method and whether the test is performed routinely, constantly, seasonally, or sporadically. If the test is done seasonally, then panels should be sent just before the testing season. If it is also advisable to send panels throughout the testing season, if resources permit. If the test is done sporadically or infrequently, PT before the commencement of testing is essential. Proficiency testing becomes even more critical if the test is done infrequently or if the number of test materials in the panel is small. Sometimes the laboratory must maintain a constant state of readiness in anticipation of samples that may never come, or that come seldom, unpredictably, and sporadically. Such a constant state of readiness means that regular panels will have to be supplied on a regular basis and some must be held, ready for use in emergencies. Consider also the costs associated with each panel design option.

**Proficiency test panel design**

Having determined needs and uses, the design of the panel can be effectively discussed. Panel design includes the following issues:

- Selection of the test material (e.g., serum, spiked muscle, preserved cultures, artifact [used in calibration PT], fixed slide).
- Composition of the panel (e.g., number of test items, number of positives and negatives, titer, or reaction range).
- Arrangement (replications of each item).
- Characterization of test materials (i.e., determination of the correct result by multiple analyses).
- Determination of homogeneity and stability.
- Determination of limits, pass or fail, per test material.
- Determination of limits, pass or fail, for the entire panel.
- Determination of pass or fail criteria based on diagnostic result.
- Frequency of sending.
- Changes in composition and arrangement of panel to eliminate bias.
- Whether to include reference materials (e.g., standardized diagnostic reagents).

What constitutes an appropriate panel design is a trade-off between costs, risks, resources, the technical aspects of the method, and the analytical, administrative, or regulatory problem to which the PT is to be applied. Requirements should be framed within the context of many varying priorities and issues.

**Choosing the test material**

Having decided on the panel design with respect to presence, absence, or concentration of analyte (i.e., the substance or constituent for which the laboratory conducts testing: e.g., antibody, cell, infectious agent, activity, chemical, or property), one must decide on the matrix that will be used to produce the test material. A matrix is all components of a material system, except the analyte. For this discussion, the matrix is the substance requiring analysis, or the test material. For a matrix such as serum, this is a relatively simple task, and one may decide on whether one will use only field samples or dilutions of serum known to be positive in comparison with dilutions of negative serum. For other matrices, there will be many other considerations, particularly if the matrix is to be spiked with analyte (i.e., a measured quantity of analyte is added to a measured amount of the matrix). In choosing the type of material to be used on the panel, one should consider the characteristics of the test and the characteristics or numbers that must be experimentally obtained to characterize the test materials. One should also consider the results the test method generates. These considerations include:

- Characterization of test materials.
- Determination of homogeneity and stability, including shelf life.
• Determination of matrix effects (e.g., the detection and effects of interfering substances and inhibitors).
• Determination of ecological effects (e.g., effects of other microorganisms in the test material on the microorganism of interest).
• Determination of handling (e.g., processing, preservation, storage, shipping, and thawing or resuscitation [or both]) effects.
• Determination of limits, pass or fail, per panel item.
• Determination of limits, pass or fail, for the entire panel.
• Determination of pass or fail criteria based on the diagnostic result (e.g., positive or negative, over or under regulatory cutoff).

Time, costs, and availability of materials will affect what can be done. Once it has been decided what can be done, it is necessary to determine what effect this will have on the applicability of the PT and on the uncertainty of the established “true” value of the test material. For example, if one sends cultures with which to spike a matrix at the laboratory under test, one should note that one is limiting or changing what abilities the PT is actually verifying or measuring. Similarly, a PT consisting of the microscopic examination of and interpretation of what is seen on slides will only measure the proficiency of reading the slide and not the proficiency of slide preparation.

Creation of the test material and panel items

Creation of the bulk test material may be accomplished in many ways, including spiking the matrix with the analyte of interest. A full discussion of the methods that are used to create test materials is beyond the scope of this article. There are several very important QA activities that must be done. These include the criteria listed in “Choosing the Test Material,” above.

• Characterization of test materials: Each panel item must be assigned a true value. This value must be determined using data obtained by valid testing, using the applicable method, by proficient technicians, and using a statistically appropriate number of test runs per test material. Careful thought must be given to the logistics of obtaining this value because it will be used to determine the pass or fail value. If possible, it is advisable to use more than 1 technician and to do at least 10 runs per test material per technician, with a total of 40 results per material. This will establish a mean value having a reasonable confidence interval while still allowing for differences among technicians. If this is not possible, one must know how this contributes to the measurement uncertainty. For qualitative materials (positive or negative), which are weak positives, it is advisable to have a positive result at least 95% of the time to call the material “positive.” In such a case, 20 results would be the minimum necessary to establish this percentage.
• Determination of homogeneity and stability: This important step is required by ILAC-G13. Such determinations include analyte distribution (confirmation of even distribution), shelf life, and shipping, handling, preservation (e.g., lyophilization), and storage/usage (e.g., freezing and thawing) effects.
• Determination of matrix effects: Matrix effects are the influence of a material property, other than the analyte, on the measurement or the physiochemical effects of the matrix on the method’s ability to accurately measure an analyte. Examples of matrix effects are interferences or cross-reacting substances. When the usual matrix must be processed or altered in a way that is not normal for field samples, such effects must be determined. This is particularly important when spiking a matrix. For example, when one spikes hot dogs with Listeria spp., processing the test material (spiked hot dogs) may break down the matrix (hot dogs), releasing enzymes that affect resuscitation or recovery (or both). The National Committee for Clinical Laboratory Standards (NCCLS), which has published several useful documents concerning PT, discusses the determination of matrix effects in detail.
• Determination of ecological effects (e.g., effects of other microorganisms in the material on the microorganism of interest): This is particularly important in microbiological PT because the presence or absence of other microorganisms can affect the isolation rate or behavior of the one of interest.
• Determination of handling (e.g., processing, preservation, storage, shipping, freezing and thawing, and resuscitation) effects.
• Determination of limits for pass or fail per panel item: It is important that these limits be defined in a scientific, statistically valid, and usable manner. The limits must address the diagnostic issues and must reflect the needs and expectations of the program. Standard deviation determinations are useful for continuous data (e.g., optical density). A determination of ±2 SD from the mean is often used for positive items. For some negative items, as much as 3 SD may be used. The use of limits larger than ±3 SD is meaningless. Some programs use outlier tests such as the Youden test to define outlier laboratories. Typically, no misses are allowed except in the case of weak positives. Quadrupling weak positives (labeling these as separate panel items) and allowing some misses within the quadruplicates (e.g., 1 miss of the 4 results if the test material was positive at least 95% of the time in the expert lab-
oratory that produced the test material) may assist in a reasonable determination of proficiency.

- Determination of limits for pass or fail for the entire panel.
- Determination of pass or fail criteria based on the diagnostic result (e.g., positive or negative, over or under regulatory cutoff).

**Labeling, packaging, storage, and distribution of PT materials, items, and panels**

All PT item manipulations, including the processing activities of spiking, aliquoting, and preserving, should be done with techniques in environments that will ensure that the integrity of the PT is not compromised in any way. Manipulations must be performed to protect the PT items from contamination, misidentification, deterioration, and degradation.

Each PT material, item, and box must be labeled. Labels must be legible, durable, and securely attached to the container. Each bulk material label should contain the following information:

- The identity of the test material and panel type (e.g., proficiency panel for the bluetongue virus neutralization test - bovine serum).
- The value (e.g., titer of 1/8).
- The amount (e.g., 100 mL)
- The producer (e.g., the name of the producer laboratory or provider [or both], with addresses and phone numbers).
- Expiration date.
- Batch number.
- Any other information that would allow each bulk material to be uniquely distinguished, referenced, and traceable to source materials and data.

In panel sets, item numbers should be randomly assigned. This coded number should be on the item (e.g., vial) label. Item labels may have more information, as space permits. The amount of item is useful. Traceability of each item should be by item number and box numbers to the test material identification and value.

Panels should be boxed (placed in appropriate storage containers) and the box sealed with tamper-evident tape or seal(s). The box label should include the following information:

- The identity of the panel (e.g., proficiency panel for the bluetongue virus neutralization test).
- The producer (e.g., the name of the producer laboratory or provider [or both], with addresses and phone numbers).
- The item numbers, but not the values. Do not use the material number.
- Storage conditions.
- Expiration date.
- Number of the box (e.g., 1 of \( n \)).

In shipping the PT, the provider must ensure conformity with transport requirements and that item integrity is maintained. Accompanying each panel should be instructions for panel handling and use, including the specific identity of the method(s) to be used, if applicable, and instructions and forms for reporting the results. Such instructions should also indicate any potential problems, especially safety concerns, with its use.

More guidance on some of these items may be found in ILAC-G12 and in ISO Guides 30 through 35. The effect of the above activities on stability should be determined and shelf life determined and documented.

**Statistical analysis and interpretation of PT data**

Proficiency testing produces data. Data must be analyzed correctly, and the choice of statistical design and methods will determine how the results of these analyses are interpreted or applied, and therefore, the usefulness to the PT program. Statistical analyses are chosen to objectively establish and justify whether an analyst or laboratory passes or fails. Such analyses determine pass and fail criteria per item or per panel, detect bias, and determine outlier laboratories. Such criteria define whether a laboratory can or may do diagnostic work. Therefore, the choice of appropriate statistical methods is a critical QA step in PT and is essential to the usefulness of PT results.

An in-depth discussion of the statistics that may be used is beyond the scope of this article. The statistical analyses used will be determined by the amount and quality of data produced, by the method(s) used to generate the data, by the desired risk, and by how the PT is being used (e.g., analyst certification, method comparison, laboratory comparison, and reference material production). ILAC-G13 discusses how statistics are used in PT. It also has an appendix and a bibliography that are pertinent to the use of statistics for PT. The OIE has a useful general discussion of considerations for the choice of statistical analysis. ISO and the NCCLS have also published many useful statistical references.

**Application of the analysis and determination of action criteria**

Effective and appropriate analysis and action are necessary for the effective and correct application of PT results, particularly when the PT is used in laboratory approval or accreditation programs. In establishing action criteria (action to be taken on a laboratory’s PT results), the provider should consider how many kinds of “failure” there are and what might be
the impact on a test result. The PT provider must also consider and clarify what “pass” really means (e.g., how closely the PT actually mimics testing of submitted samples, and other limitations of the PT data to verify or predict capability).

Where applicable, PT providers must logically establish and clearly and unambiguously document what criteria and actions are to be taken for each type of pass or fail. If bias is an issue, the provider must establish and document the action that will be taken in each case. Such actions might include resending a PT; retraining, suspension, or revocation of the laboratory’s or analyst’s approved status; or a notice being sent to laboratories to work together to discover the cause of bias and remove it. The PT providers are expected to provide expert commentary on the performance of participants.²

Records

Accurate, legible, indelible, complete, unambiguous, objective, secure, and retrievable records, with established and recorded retention times, are an essential QA item. The PT provider must be able to recreate all events relating to the preparation of the test material, the panels, and the program. Record-keeping and archiving systems are also needed. For PT programs, the records to be kept include all individual measurement observations, including test results and records to be kept for any submission to be tested by a particular method (e.g., test worksheets). In addition, records should include:

- References to procedures used to do the work.
- Retired or superceded technical procedures.
- Relevant communications.
- Sample/test material records.
- Sampling records/sourcing records.
- Records of meetings, actions, and decisions.
- Calculations.
- External and internal test reports.
- Calibration certificates.
- Clients’ notes, papers, and feedback.
- Contracts.

Records for the PT program itself include:

- Calculations.
- Derived data, including statistical treatments.
- Results from participants.
- Program reports.
- Identity and traceability of test items, materials, and panels.
- Results of homogeneity and stability testing.
- Quality-management system records.
- Superceded or retired procedures.

Procedures

Required procedures must be documented, clear, unambiguous, and described in sufficient detail. They must be created, approved, distributed, revised, and archived according to a documented system of document control. In addition, where “ensure” is used in this article in the context of ensuring a particular outcome, this indicates that documented policy or procedure (or both) describing how this will be accomplished must be in place, and that these procedures, by the validity or logic of the systems or tasks described therein, will accomplish the desired result. When no longer in use, procedures are records.

Prevention of PT analyst bias

An important QA activity in PT is ensuring against the prevention of analyst bias. Because of its importance and because it contains elements of other topics discussed, this area warrants a separate discussion. People want to do well on tests. Some may have to pass a PT to retain their job. People and laboratories can share the results of PTs and other information that can bias the results of the PT. The value of a PT will be seriously compromised or even lost if people believe they can guess at the result of an unknown. Panels should be constructed so as to negate this human tendency. There are several ways in which this can be done.

- Recode each and every panel (“box” of PT “samples”) separately: Each test item should be given a randomly selected number or code that is traceable to the test material and test value. Each item is then arranged in the box from lowest to highest number, as for a submission. The coding should be different for each laboratory or analyst for a particular panel set per exchange. In certain cases, the composition of each panel with regard to items and results may have to be different, particularly in the case when 2 analysts operate in the same laboratory.

- The use of blind panels is essential. Single-blind panels are identified in a manner that ensures that the laboratory under test does not know the result (i.e., coded). However, the laboratory knows that these test items are a panel. The use of double-blind panels (laboratory under test does not know that the test items are not normal submissions) yields more accurate information about proficiency but is usually neither practical nor possible because the use of double-blind panels may involve the falsification of submission paperwork. This can be illegal unless appropriate arrangements are made. Although difficult and costly to set up, double-blind panels may be very helpful in some situations because the comparison of single-blind and double-blind results for
the same test materials from the same laboratory can be enlightening.

- Avoid the use of a test material or item that has a distinct appearance when the item is to be used repetitively or replicated. For example, for a panel composed of serum, it may be easy to spot a particularly lipemic, hemolyzed, pale, flocculent, or dark yellow item, if the same item is used often. The use of a coloring agent or other additives, where possible, may help avoid this problem. Some test items may be processed before testing and aliquoting for use on a panel to ensure a particular uniformity of appearance.

- Change the diagnostic makeup of a panel each time the panel is used: For example, vary the number of positives and negatives in an unpredictable way. For sets of weak positives and negatives, do not use a 50% positive, 50% negative composition. A 30:70 ratio works better in avoiding bias, while maintaining the ability of the panel to assess, with adequate numbers of test items, the capability to distinguish a weak positive reaction from no reaction. In some cases, particularly when small panel sets are tested frequently, it can be useful to have panels that are all negative, or mostly negative with 1 positive (i.e., ratios greater or smaller than 70:30). Such extensive changes will remove bias, but these will also remove data that tell the PT result user (e.g., the regulator) how well the laboratory under test can distinguish weak positives from negatives.

- Change the numbers of replicates in the panel each time the panel is used: As stated, using quadruplicates in a PT set can provide valuable information about within-laboratory precision and, in particular, whether there has been a change in sensitivity. Also, as previously stated, quadruplicates can aid in spotting test material and therefore sample-associated differences. For example, if a test fails to transfer because of too many panel items out of limits, decoding the quadruplicates can assist in determining whether this is associated with particular items. If only positives, and only certain positives, are affected in the same way, this can point to antigen-processing or handling differences. However, if a laboratory knows that all the test items are quadruplicated (i.e., there are 10 test materials quadruplicated, for a total of 40 items as 40 separate unknowns to analyze), this knowledge will bias the laboratory in the assignment of a diagnostic value to a single test item. This is particularly true if the test materials have a distinct appearance. For example, the analyst may not be able to decide whether the answer is weak positive or negative. If the analyst believes that the panel items are quadruplicated, then the analyst can look to see which set of quadruplicates is missing a result, and use this, rather than scientific observation, to determine the reported result.

### Choosing a PT provider

Because PT results are critical to the success of a laboratory’s internal QA program, to its ability to be accredited, and possibly to its approval to do work, it is important to a laboratory to choose a provider whom it can trust. There are therefore QA issues to be addressed in the choice of a provider. It is recommended that the provider be accredited according to the ILAC scheme. If this is not possible, then the provider should be a recognized expert (e.g., an OIE reference laboratory or a national reference laboratory) in the testing for and diagnosis of the disease of interest or in the analytical technique used, if easily applicable to a wide range of organisms, substances, or pathologies (e.g., some histopathological examinations). Registration of a provider’s quality system as meeting the requirements of ISO 9001 is a commendable accomplishment but does not necessarily imply technical competence.

Regarding the accreditation of PT providers, ABs will accredit the providers of PT programs. Although AB requirements for accreditation may vary slightly, they are based on ILAC-G13. This document is based on the requirements of ISO Guide 43-1, the relevant elements of ISO/IEC 17025, and the relevant elements of ISO 9001. These guidelines are very comprehensive and complex, overlap with those of ISO/IEC 17025, and could easily be merged into an extant ISO/IEC 17025 management system. ILAC-G13 lists technical and management requirements (e.g., document control, corrective action system). These requirements include a quality-management system that requires policies and procedures designed to protect confidentiality and ensure ethical actions with regards to control of documents and other information, review of contracts, use of collaborators (e.g., providers of test material), procurement, client feedback, control of nonconforming test materials or nonconforming work, corrective and preventive action, internal audits, management review, document control, and records. Specific operational and staff requirements are also listed.

Technical requirements for accreditation include policy and procedures for test item preparation, management, storage, homogeneity and stability testing, packaging, labeling, and distribution. Accreditation of PT programs is also dependent on demonstrably acceptable procedures for the analysis and interpretation of data, communication with participants, coordination and operation of the program, choice of test methods or procedures, and dealing with collusion and falsification of results. ILAC-G13 also presents the requirements for staff and collaborators (e.g., laboratories or...
other parties subcontracted to produce test material). These collaborators must have the necessary technical, statistical, and administrative expertise and the necessary authority and resources to do the work. Staff requirements also include a program coordinator, technical manager, and quality manager. In addition, ILAC-G13 requires a program plan, program reports, data-processing equipment, documented and valid statistical models and data analysis and interpretation techniques, and instructions for the PT participants. Because the production of PT test materials requires testing, it would be desirable for collaborator laboratories to be accredited to ISO/IEC 17025 (i.e., verified as competent to conduct testing).

Accreditation bodies have very helpful information on their Web sites for those interested in establishing a PT program that meets national and international requirements for such activities. These ABs include the American Association for Laboratory Accreditation, Standards Council of Canada, and the National Association of Testing Authorities. Some ABs operate PT programs. There are many accredited providers. Those accredited may be found on AB Web sites. The Web sites of these organizations will be helpful in describing the services they offer. Some providers may operate under a quality system registered or certified as meeting the requirements of ISO 9001. Although this is a desirable attribute, it is not equivalent to the requirements used by ABs, which also, in addition to quality system requirements, specify technical requirements.

Higher levels of PT

As further discussion of the usefulness of PT programs, the following information may be particularly useful to laboratories performing regulatory animal health diagnostic testing. Because federal animal health diagnostic laboratories are often required to develop or operate laboratory approval programs that use PT as part of the approval process or to participate in PT for trade agreements, the following information will be particularly useful.

Proficiency tests can be arranged by ABs among the laboratories each accredits to assess the technical equivalence of their accreditations. Such PT activities are often arranged by the body that recognizes the ABs as competent to award accreditations (e.g., National Cooperation for Laboratory Accreditation [NACLA], or a national standards body). Proficiency testing among laboratories accredited by different ABs is often done to support Mutual Recognition Arrangements of the ABs, in which the ABs agree to accept each other’s accreditations. Governments may arrange PTs among themselves to facilitate a particular trade or treaty issue, or to support a Mutual Recognition Agreement. For example, the National Institute of Standards and Technology of the United States Department of Commerce arranges such projects. National Reference Laboratories often participate in international-level PTs to fulfill ISO/IEC 17025 requirements and to confirm their diagnostic equivalence to other laboratory systems.

Conclusions

Proficiency testing, in itself an important QA and conformity assessment activity, has QA and quality-management issues of its own. These issues cover technical, operational, and managerial aspects of PT. Proficiency testing has become an essential component of test result verification, method validation and harmonization, laboratory approval by regulators, third-party laboratory accreditation, and management of trade issues. The initiation of accreditation programs for PT providers verifies the importance of the ability to demonstrate competence in this area by meeting required and relevant national and international standards. It will become increasingly important for the providers of PT to have third-party recognition of their competence to carry out PT activities and to have this recognition based on international standards and guidelines. It will become even more important for the users of such programs to have accredited providers available. All providers of PT programs, especially those using PT in the approval of laboratories, should carefully consider the need for implementing the requirements specified in relevant standards, as well as the need for obtaining third-party accreditation.

Acknowledgements

I thank Dr. David Scarfe of the American Veterinary Association, Schaumburg, Ill.; Dr. Randall Levings, Dr. Richard Trowbridge, Dr. Janet Payeur, and Ms. Lynn Lynch of the National Veterinary Services Laboratories, Ames, Iowa; and Dr. Paul Hauer and Ms. Kathy Clark of the Center for Veterinary Biologics, Ames, Iowa, for their assistance in writing this article.

References