International standards for test methods and reference sera for diagnostic tests for antibody detection

P.F. Wright

Canadian Food Inspection Agency, National Centre for Foreign Animal Disease, 1015 Arlington Street, Suite T2300, Winnipeg, Manitoba R3E 3M4, Canada

Summary
International standards are necessary to ensure that the diagnostic tests used by trading partners meet a minimum standard of diagnostic performance. The tests that are used to qualify animals for international movement must provide a degree of confidence that those animals which give negative test results (and therefore qualify for movement) are free of a particular infectious disease agent. Without the development of standards for diagnostic tests, no level of international harmonisation can ever be achieved. International standards for both test methods and sera serve to establish a baseline on which inferences can be made with respect to the diagnostic performance of a given test in a given laboratory. International standard test methods set baseline analytical and diagnostic performance requirements for new reagents or methodologies. International standard sera are primary reference standards and serve as both reference materials for the calibration of test methods and reagents and prototypes for the production of national and working standards. For most serological test methods, three standard sera should be established: a strong positive, a weak positive and a negative standard. These standards are used to calibrate the detection range and analytical sensitivity of the test method. Assays calibrated in this way are more likely to exhibit a diagnostic sensitivity and specificity which parallels that of the standard test method, thus resulting in a higher degree of international harmonisation.

Keywords
Agar gel immunodiffusion - Complement fixation tests - Enzyme immunoassay - International trade - Reference standards - Sensitivity - Serological techniques - Specificity - Virus neutralisation.

Introduction
One of the primary mandates of the Office International des Epizooties (OIE) is the harmonisation of regulations for trade in animals and animal products among Member Countries. The diagnostic tests used to qualify animals or animal products for international movement form an integral part of these regulations. Two categories of diagnostic tests are recognised by the OIE for the purposes of international trade. In the first category are ‘prescribed’ tests, which are required by the International Animal Health Code (13) for the testing of animals before they can be moved internationally. However, it is not possible at present to have prescribed tests for all of the OIE Lists A and B diseases (13). Therefore, the second category contains ‘alternative’ tests which are suitable for the diagnosis of disease within a local setting, and which can be used in the import/export of animals after bilateral agreement. Irrespective of the test category, there is a need for international standard methods and sera for diagnostic tests. This paper focuses on some of the major considerations pertinent to the development of these international standards for diagnostic tests for the detection of antibody to infectious disease agents.
Diagnostic performance

Predictive value

International standards are necessary to ensure that the diagnostic tests used by trading partners meet a minimum standard of diagnostic performance. The tests that are used to qualify animals for international movement must provide a degree of confidence in the probability that those animals which give negative test results (and therefore qualify for movement) are free of a particular infectious disease agent; confidence in what is known as the 'predictive value' of the test result. While both positive and negative test results are important and have their own predictive values, the predictive value of a negative test result is of paramount importance in trade. This represents the probability that an animal with a negative test result is free of infection.

The calculated predictive value (positive or negative) of a test is influenced by the following three factors:

a) diagnostic sensitivity
b) diagnostic specificity
c) prevalence of disease (1, 12).

Diagnostic sensitivity and specificity are characteristics of the test. Prevalence of disease is obviously related to the disease situation in the exporting country in which the test is being applied.

Diagnostic sensitivity

The diagnostic sensitivity of a test is calculated as the proportion of known infected animals which give positive test results, i.e., 'true positives'. Infected animals which yield negative results are considered to be 'false negatives'. Estimates of diagnostic sensitivity require the testing of as many samples as possible representing the broadest range of expected antibody activities in an infected population (12). This is a most difficult estimate to achieve with any degree of confidence because of the infinite range of antibody activities that may be encountered.

Diagnostic specificity

The diagnostic specificity of a test is calculated as the proportion of known uninfected animals which give negative results, i.e., 'true negatives'. Uninfected animals which give positive results are considered to be 'false positives'. As a general rule, accurate estimates of diagnostic specificity require the testing of thousands of samples from known uninfected animals (12). False positive results are often infrequent and unpredictable, and may occur at rates of one or two per thousand samples tested.

Predictive values, whether positive or negative, are calculated based on estimates of diagnostic sensitivity and specificity and disease prevalence. Errors in these estimates are compounded in the calculation of predictive value. Therefore, these estimates should be as accurate as possible.

In turn, diagnostic sensitivity and specificity are influenced by two factors, as follows:

a) analytical sensitivity
b) analytical specificity.

Both of these factors are inherent in the test and are in turn influenced by the reagents chosen and the diagnostic test method developed for their use.

Analytical sensitivity

The analytical sensitivity of a serological test refers to the smallest amount of antibody that can be detected in a test sample. Antibody is a functional term which refers to any immunoglobulin elicited in response to infection with a particular disease agent which is capable of binding to an antigen present on, or derived from, the agent. The term includes immunoglobulins of all isotypes (e.g., IgM, IgG1, IgG2, IgA, etc.). Depending on the diagnostic method, different isotypes may not necessarily be detected with equal efficiency. Analytical sensitivity is often assessed by end-point dilution analysis of test samples. In some cases, analytical sensitivity is inferred from the earliest time, post infection, at which antibody is detectable.

Analytical specificity

The analytical specificity of a serological test refers to its ability to detect antibodies of particular diagnostic relevance to the infection being questioned, in other words, the specificity for antibodies which uniquely identify the causative agent and not antibodies that may be elicited in response to cross-reacting antigens on other related or unrelated organisms. There are a number of ways in which the analytical specificity of a test can be modulated; most commonly by using unique sub-unit antigens, by designing test methods to detect antibody isotypes of particularly high specificity and relevance, or by using highly specific monoclonal antibodies in competitive assays. Analytical specificity is determined by testing panels of samples from animals infected with agents known to express cross-reacting antigens.

All of the parameters described above are obviously interrelated and interdependent; the parameters influence the overall diagnostic performance of a given test and ultimately, the confidence in a given test result. Without the development of standards for both test methods and reference materials, no level of harmonisation can ever be achieved. International standards for methods and sera serve to establish a baseline on which inferences can be made with respect to the diagnostic performance of a given test in a given laboratory.

Standard test methods

The diagnostic test methods for antibody detection that are described and designated as prescribed tests in the OIE Manual of Standards for Diagnostic Tests and Vaccines (hereafter referred to as the OIE Manual) (11) represent the...
current international standards. Prescribed tests provide an acceptable degree of confidence in the probability that those animals which give negative results are free of a particular infectious disease agent. Alternative tests may also be used for trade under special bilateral agreements. These tests may not provide the same degree of confidence as prescribed tests or they may not be validated to the full extent required by the OIE Standards Commission.

To say that all prescribed tests provide the same degree of confidence would be an overstatement. Not all prescribed tests offer the same level of analytical or diagnostic performance. However, a prescribed test method currently accepted as the standard is a valuable tool in the diagnosis of infectious disease. For many diseases, there are no prescribed tests available. This serves to underline the fact that diagnostic tests alone are not the sole consideration in any decision to allow the international movement of animals.

International standards for test methods are important in this discussion because these are the test methods which will set the baseline for diagnostic sensitivity and specificity, and for the development and characterisation of international standard sera. Some of the classical serological methods, such as complement fixation (CF), agglutination, precipitation and virus neutralisation (VN), have been prescribed tests for trade for many decades. With refinements to reagents and advances in supporting technologies, the analytical and diagnostic performance characteristics of some of these test methods have improved over the years. For all of the effort that has gone into the development of primary binding assays, especially enzyme immunoassay methods, surprisingly few are prescribed for international trade. However, this is now changing with the development of international standards for this test method in particular, and for the validation of test methods in general (17). These are just two of the reasons for which the OIE has established a network of Reference Laboratories and Collaborating Centres; to develop and standardise new test methods, and to co-ordinate the validation of test methods at the international level.

**Documentation**

An international standard test method should be fully documented and/or referenced. All biological reagents should be described in detail with respect to composition, purity, biological activity and biosafety or biobrane considerations. Where biological reagents are derived from reference cultures or strains, these should be identified. Pertinent references should be cited with respect to the production and standardisation of any biological reagent. Commercial sources of acceptable biological reagents should be identified. All titration procedures necessary for the optimisation of biological reagent activity should be detailed. All chemical reagents and buffers should be described as to their chemical grades, formulation, pH and ionic strength. All physical parameters of time, temperature, incubation, etc. should be described fully. Requirements for equipment and disposables should be described fully, especially if substitution may affect assay performance. Guidelines should also be given for internal quality control procedures and for calculation of results and interpretation of data. For the most part, descriptions of standard test procedures and for calculation of results and interpretation of data. For the most part, descriptions of standard test procedures presented in the various chapters of the OIE Manual fulfil the above requirements.

**Validation**

Standard methods must also be properly validated. This subject is discussed in the OIE Manual (12) and in a detailed presentation elsewhere in this edition of the Scientific and Technical Review and will not be reiterated here (5). However, it is pertinent to point out that the Standards Commission of the OIE requires that any test method being proposed for consideration as a prescribed test must be fully validated. This means that data must be submitted on the analytical sensitivity and specificity of the method. Estimates of diagnostic sensitivity and specificity must be presented based on statistically sound field trials. Data must be presented on assay repeatability, i.e., agreement between replicate samples within and between runs of the test method. To be internationally accepted, data must also be provided on reproducibility, i.e., the ability of the test method to provide consistent results when applied to aliquots of the same sample at different laboratories in several different countries.

The fact that standard test methods are of crucial importance cannot be overemphasised. For international trade, the standard method sets the baseline for diagnostic performance, and gives laboratories developing their own reagents and protocols a standard of comparison. It also sets the minimum acceptable standard for the development of new methodologies with respect to analytical and diagnostic performance requirements.

**Standard sera**

International standard sera are primary reference standards and are necessary for the standardisation and harmonisation of diagnostic test methods for the detection of antibody to infectious disease agents. These sera act as reference materials for the calibration of test methods and reagents and as prototypes for the production of national and working standards (2).

At a Meeting of Consultants sponsored by the Joint Food and Agriculture Organisation of the United Nations/International Atomic Energy Agency (FAO/IAEA) Division of the IAEA, convened in Vienna in January of 1992, one of the issues tabled related to international standard sera for enzyme immunoassay (EIA) techniques for antibody detection (17). The consensus of opinion was that three standard sera should be established: a strong positive, a weak positive and a negative standard. The positive reference standards should be selected from animals which exhibit a typical humoral response to the organism in question. Negative reference
For most diagnostic applications, samples are tested at a single dilution, but its guidelines for the preparation of international standard sera have been adopted. Using three standard sera to calibrate an assay by defining the minimum acceptable dose-response characteristics provides greater confidence that the assay will exhibit a diagnostic sensitivity and specificity which closely parallels that of the standard test method originally used to characterise the standards. Assays calibrated against a single positive standard serum alone may not provide the same level of confidence. The OIE Standards Commission has adopted this approach in its guidelines for the preparation of international standard sera and recommends that three standard sera be established for all EIAs and most other tests prescribed for trade. (6) and recommends that three standard sera be established for all EIAs and most other tests prescribed for trade.

Standard sera for enzyme immunoassays

For most diagnostic applications, samples are tested at a single dilution in EIAs rather than in a dilution series. At a single dilution, EIAs are semi-quantitative in terms of expression of antibody activity. In competitive or blocking EIAs, antibody activity is expressed as percent inhibition of a specific competing antibody. In indirect enzyme-linked immunosorbent assays (i.e., indirect ELISAs), it has been proposed that antibody activity be expressed as percent positivity relative to an antibody positive reference standard (17). This serves to unify the scale (0% to 100%) on which normalised data is expressed in all EIAs.

A working model for the selection and definition of antibody standards for the indirect ELISA using bovine brucellosis as an example has been described (18). A positive serum is selected as one which gives a typical dose-response curve when doubling dilutions are first prepared in negative serum and then tested at the normal diagnostic dilution by the standard ELISA method. After defining the linear portion of the dose-response curve, the strong positive standard is selected as that dilution which represents antibody activity midway between the upper and central points of the linear portion of the curve, and the weak positive standard as that dilution which falls midway between the central and lower points. As the strong and weak positive reference standards are prepared as a dilution in negative serum, that serum is included as a negative reference standard. The strong positive standard represents 100% positivity and is used to normalise all other raw data. The weak positive standard is used to define the minimum analytical sensitivity of the test. The analytical sensitivity of an indirect ELISA is dependent on a number of variables including the antigen, the antiglobulin and the enzyme label chosen. The weak positive reference standard must be unequivocally positive when tested at its specified dilution, irrespective of any variable encountered in assays of this type.

The selection and definition of antibody standards for competitive or blocking EIAs should be approached in a similar fashion. The key to any assay of this type is the competing antibody which is usually a monoclonal antibody preparation. Each monoclonal antibody preparation will be unique in terms of analytical specificity and sensitivity. As for the indirect ELISA, a positive serum is selected that exhibits a typical dose-response when doubling dilutions are first prepared in negative serum, and is then tested at the normal diagnostic dilution by the standard test method. In general, the dose-response curves of these assays tend to be much steeper than for indirect ELISAs, and the linear portion of the curve often spans only a very few dilutions. It is recommended (17) that the strong positive reference standard should represent the highest dilution which consistently results in maximum inhibition (i.e., 100%) of the reference monoclonal antibody. The weak positive reference standard should represent a dilution which results in an inhibition of greater than 50% and less than 100%. Depending on the assay and the steepness of the dose-response curve this standard is more difficult to define, but it should be unequivocally positive at all times.

Standard sera for limiting dilution assays

For limiting dilution assays, such as VN or CF tests, a positive serum standard should be representative of a typical dose-response profile when titrated in the standard test method. In the case of the CF test for detection of bovine antibodies to Brucella abortus, the 2nd International Standard B. abortus Serum (2nd ISABS) is currently the international serum standard (8). This serum exhibits a typical dose-response profile and its antibody activity has an assigned value of 1,000 international complement fixation test units (ICFTUs) per ml. At a dilution of 1/200, this serum is used to titrate the optimal antigen concentration to be used in the CF test and in doing so, calibrates the analytical sensitivity of the test. As undiluted serum, the 2nd ISABS is analogous to the strong positive serum standard recommended for EIAs. At a dilution of 1/200, it is analogous to the recommended weak positive serum standard except that it is prepared in diluent buffer and not negative serum by the user. The 2nd ISABS also forms the basis for normalisation of CF test data into ICFTUs for the purposes of international trade.
Standard sera for precipitation assays

Standard sera for precipitation assays, such as the agar gel immunodiffusion (AGID) test, should produce typical reactions when applied undiluted in the standard test method. As for the EIA, these standards should represent strong, weak and negative sera. The strong positive standard should result in a distinct precipitin line when made to react with an optimal dilution of antigen. It represents the reference line by which all test sample reactions are judged. The weak positive standard should result in the minimum detectable reaction in the assay. Wherever possible, the weak positive standard should be a whole serum. If not, it should be prepared by predilution in the negative serum standard. The weak positive standard is of major importance in AGID tests, as its reactivity defines the diagnostic sensitivity of the assay.

Precipitation assays use soluble, sub-unit antigens which, depending on the method of preparation and purity, may contain more than one type of antigen. Using enzootic bovine leukosis (EBL) as an example (3, 9), antibodies to two soluble antigens are readily detectable in the AGID test: anti-gp51 and anti-p24. Antibodies to gp-51 generally appear earlier than those to p-24, and are more consistent. The antigen preparation used in the standard test method contains predominantly gp-51, but may also contain p-24. Two standard reference sera are used. Serum E1 does not have sufficient anti-p24 antibody to be detected in the standard AGID test. This reference serum is used to calibrate the antigen to determine its optimal concentration for use in the AGID test. Serum E4 contains antibody to both antigens. This reference standard is used to calibrate the analytical sensitivity of the AGID test. When diluted 1/10 in a negative serum by the user, E4 should give an unequivocally positive reaction. In the EBL example, the two positive standards have been prepared from different sources. Undiluted, E1 is analogous to the strong positive serum standard recommended for EIA. When diluted in negative serum, E4 is analogous to the weak positive serum standard.

Defining typical sera

Throughout the foregoing discussion on calibration, constant reference is made to the term 'typical'. Standard sera should be representative of the majority of samples that would be routinely analysed in the test method. Typical with respect to the normal composition or matrix of serum, given that there are a number of host-related variables that may affect the matrix (12). Typical with respect to the isotypes, specificities and range of antibody titres expected during the course of natural infection (16). 'Typical' is difficult to define. Not all antibody isotypes fix complement, neutralise viruses, agglutinate bacteria or precipitate soluble antigens. Not all antibody isotypes are equally represented in terms of specificity, concentration or avidity at any given stage of infection. What is important is that the antibodies of diagnostic relevance are represented in a fashion which is typical of the host humoral immune response to infection with the organism in question. For some applications, selection of serum standards which are representative of the earliest stages of infection may be necessary. However, given the requirements for quarantine and other mitigating factors stated in import/export health certificates, 'typical' should be interpreted as an average or median response to infection (18). If more than one type of test method is available and the candidate serum demonstrates typical reactivity in each test, then this adds to the confidence that the serum is representative of the course of natural infection. In some cases, it may be necessary to experimentally infect or to immunise animals with inactivated organisms. However, when characterising the immune response, only sera which mimic the profile of natural infection should be selected.

Negative standard sera should also be typical with respect to the normal composition or matrix of serum. Negative sera should be checked for the presence of antibodies which cross-react with the organism in question. The reactivity of these sera should be typical of a negative reaction in the test method.

Both the positive and negative standard sera may be derived from a single source or may be pools of sera. If pools are used, each serum should be checked individually beforehand and after being pooled to ensure a typical profile.

To avoid any bias in the selection and characterisation of standard sera, the reference laboratory should send candidate sera to collaborating laboratories for testing. For international standards, these collaborating laboratories should represent the national authority in different countries. Each laboratory should use the same standard test method if possible, or at least tests of similar diagnostic sensitivity and specificity. Each laboratory should characterise the candidate sera in the same manner and provide data on repeatability. The co-ordinating reference laboratory should assess the reproducibility across all participating laboratories.

If a single standard test method has not been widely adopted, it may be necessary to use a consensus approach to establishing international standard sera. This has been the case with the establishment of standard sera for the VN and EIA techniques used for the serological diagnosis of infectious bovine rhinotracheitis (IBR). Representative descriptions of these techniques are found in the OIE Manual (10). These are prescribed tests and by default represent the standard method. However, these assays are performed with various modifications as routine tests in a number of diagnostic laboratories. In a European study, notable differences were observed in the analytical sensitivities of many of these IBR assays (14). As a consequence, a project was undertaken to establish three international standard sera based on a consensus of results from collaborating laboratories, each using its own modification of the standard method (15).
Safety and stability

According to the guidelines developed by the OIE Standards Commission (6), international standard sera should be prepared so that they are free from infectious material. It is recommended that the sera be produced in specific pathogen-free or gnotobiotic animals whenever possible. To ensure that the standards can be moved between countries, irradiation at 25 to 30 kilograys (kGy) (2.5 to 3.0 Mrad) is recommended. Bovine sera should be from a BSE-free source. To ensure stability, it is recommended that the sera be freeze-dried, and the inclusion of sterile diluent for reconstitution would be advantageous. After freeze-drying, several bottles of the standard should be reconstituted and re-evaluated.

National and working standards

Normally, international reference standards are prepared in large batches of small aliquots (0.5 ml to 1 ml). These should be regarded as a valuable commodity, and should only be used for specified purposes (4). As described above, international standard sera should be used to calibrate test methods in the national reference laboratory. Once the method has been calibrated, these standards should be used as prototypes for the production and cross-standardisation of national standards. National or secondary standards should be prepared by the national authority and used to calibrate test methods in regional, provincial or state laboratories. Working or tertiary standards should be cross-standardised against the national standard. These standards are applied routinely in the test method under the normal daily routine of the diagnostic laboratory. They may be used to establish standard curves or to act as internal controls as part of a quality assurance programme.

Data sheets

According to guidelines developed by the OIE Standards Commission (7), the following information must be supplied to facilitate the selection and/or preparation of national standards:

a) description of donor animals for the positive and negative sera, including species, age, reproductive status and origin, i.e., natural production, specific pathogen-free, gnotobiotic
b) nature of the antibody response, i.e., natural infection, experimental infection, immunisation
c) details of the organism used to elicit the immune response, i.e., source, strain, serotype
d) details of experimental infection or immunisation protocols, i.e., route, dose, immunisation schedules, method and time of sample collection
e) reference test methods used to select positive and negative reference serum candidates and to characterise the antibody response, i.e., ELISA, AGID, VN
f) sample of titration profiles of positive sera and criteria for selection of appropriate dilutions of defined activity
g) presence of heterologous antibodies, if known, and tests used in detection
h) description of the sterilisation methods, including type of irradiation and dose and condition of sample at time of sterilisation, i.e., liquid, frozen, freeze-dried
i) batch number and date of production
j) recommended reconstitution, handling and storage conditions.

Conclusion

International standards for test methods and reference sera are necessary to ensure that the diagnostic tests used by trading partners meet a minimum standard of diagnostic performance. These standards serve to establish a baseline on which inferences can be made with respect to the diagnostic performance of a given test in a given laboratory. Without such standards, no level of international harmonisation can ever be achieved. This paper has focused on the essential considerations which must be addressed in the development, evaluation and documentation of standard test methods and reference sera for diagnostic tests for antibody detection – considerations which should result in a higher degree of international harmonisation amongst diagnostic tests and in a higher degree of confidence in test results produced by trading partners.
References


