CHAPTER 2.2.7.

VALIDATION OF DIAGNOSTIC TESTS FOR INFECTIOUS DISEASES APPLICABLE TO WILDLIFE

INTRODUCTION

The WOAH Validation Recommendations in Section 2.2 Validation of diagnostic tests *of this* Terrestrial Manual *provide detailed information and examples in support of the WOAH Validation Standard that is published as Chapter 1.1.6* Validation of diagnostic assays for infectious diseases of terrestrial animals. *The Term "WOAH Validation Standard" in this chapter should be taken as referring to that chapter.*

Diagnostic testing of wildlife for infectious diseases is becoming increasingly important as interest grows in diseases that occur in wildlife that may have an impact on wildlife populations and biodiversity, as well as on the health of humans and domestic animals. For the purposes of this standard, "wildlife" will be defined as animals belonging to one or more of the following groups:

i) *Wild animals: Those animals that do not live under human supervision or control and do not have their phenotype selected by humans.*

ii) *Captive wild animals: Those animals that live under human supervision or control but do not have their phenotype selected by humans.*

iii) *Feral animals: Those animals that do not live under human supervision or control but do have their phenotype selected by humans.*

Wild animals generally are susceptible to infection with the same disease agents as domestic animals and, in some cases, the tests developed and validated in other species may have utility for wildlife species. However, diagnostic testing of wildlife can be more challenging than in domestic animals for a variety of reasons, including difficulties in animal and sample accessibility, poor sample quality, poor knowledge of pathogenesis/epidemiology of the disease in wildlife, and local or international regulations limiting or prohibiting possession and/or international shipment of samples. Affordability of tests is a key consideration because wild and feral animals do not have owners who pay for testing. Hence, low cost may be a critical factor in test selection for use for a designated purpose.

Many routine diagnostic tests that have been developed and are currently used for detecting or confirming diseases in domestic animals have not been validated for wildlife. The question remains as to whether there are any essential differences in diagnostic sensitivity or specificity of these tests when they are applied to wildlife samples.

Diagnostic tests can arbitrarily be divided into two categories: direct and indirect identification techniques. Direct diagnostic test methods to identify agents include microscopic examinations, culture – commonly used to isolate bacteria, viruses, fungi and some protozoa; and molecular techniques – including polymerase chain reaction (PCR) amplification of the agent's genetic material and sequences coding for immunoproteins. Importantly, these direct agent identification diagnostic techniques should theoretically not be affected by the species of the host, i.e. domestic animal or wildlife. However, there may be some species variation in the proliferation rate or amplification of the agent, which may affect the amount and distribution of pathogens and their products in the hosts. Indirect test methods are based on the detection of an animal's cellular or antibody immune response against a pathogen. In contrast to direct methods, detection of the immune response often requires species-specific reagents, which complicates this diagnostic approach in wildlife in species that do not benefit from existing validated tests in a closely related

species. Determination of the actual infection status of animals identified as infected or exposed in a serological test requires confirmation by a validated direct detection assay.

Validation of diagnostic tests for individual wildlife species presents challenges, including the accessibility of adequate sample numbers and volumes to be used in the validation process. The underlying principles and stepwise approach to the validation of a diagnostic test are outlined in the WOAH Validation Standard. The purpose of the standard described in sections 1 and 2 of this document is to present information specifically for validation of diagnostic tests for wildlife species that will be recognised by WOAH (completion of Stages 1, 2 and 3 of the Validation Pathway). However, recognising that completion of this process may not be necessary, or even possible, in all circumstances, guidance is provided for following the Validation Pathway to a point where the test can be provisionally recognised to provide confidence in results and for use in specific applications in a regional or national context (see section B.2.6 of the WOAH Validation Standard for more details).

Provisional recognition of an assay for wildlife species requires completion of Stage 1 (analytical characteristics) and Stage 2a (preliminary estimates of Diagnostic Sensitivity [DSe] and Diagnostic Specificity [DSp]) of the Validation Pathway. Stage 2a evaluation using a panel of positive and negative reference samples to evaluate diagnostic performance is considered essential because of the diversity of species within taxonomic families, varying host factors that may influence pathogenesis of infection, and different disease ecologies. Details about Stage 2a evaluation are provided in section 2 of this standard. Stages 2b and 3 of the pathway also need to be completed for a provisionally recognised assay to achieve full validation, through the WOAH pathway, for its originally intended purpose.

1. Test validation principles

Validation is a process that determines the fitness of an assay, which has been properly developed, optimised and standardised, for an intended purpose. Ideally, the validation process of tests for wildlife should be conducted in the same way as for tests for domestic animals (presented in the WOAH Validation Standard). However, as explained above, wildlife diagnostic testing often meets difficulties, which also place limitations on the prospects for full validation. Therefore, in cases in which full validation is not feasible, the best possible alternative may be to evaluate the wildlife assay's fitness in a reduced number of reference samples. The preliminary estimates of a test's performance may provide sufficient information for government authorities to agree that a test can be provisionally recognised for testing of animals being moved or translocated or for pathogen surveillance within a country.

In many cases, pre-existing diagnostic tests validated for one species may be adapted to and evaluated in other species with minimal or no modifications. In other cases, new tests for wildlife may need to be developed. In all cases, the intended purpose(s) and application(s) of the test should be established and defined before it is developed and validated as this may have an impact on selection of appropriate reference samples and ultimately, generalisability of the validation results.

The development of rapid and easy to perform field tests (animal-side or pen-side tests) for disease diagnosis in domestic animals has been well received by end users and these tests are becoming increasingly popular for use in wildlife. The use and interpretation of field tests is often the sole responsibility of the veterinary personnel attending to cases in the field without laboratory support. Therefore validation of these tests through Stage 2a by the manufacturer is essential to facilitate correct interpretation of the test results. Test kits used in the field rather than under laboratory conditions should be evaluated for reproducibility of results under different environmental conditions (temperature, humidity, etc.).

1.1. Fitness for purpose

A list of purposes for diagnostic testing is provided in the WOAH Validation Standard. More specifically for wildlife testing, the main purposes to develop and apply a diagnostic assay are:

1) Screening wildlife populations for the presence of infectious agents, for example:
   a) for surveillance (e.g. early detection, evaluation of trends in prevalence or incidence)
   b) to estimate prevalence of infection or exposure.

2) Screening or testing vectors or environmental samples for the presence of infectious agents.

3) Confirming a diagnosis of suspect or clinical cases (includes confirmation of positive results from a screening test).

4) Certifying freedom from infection or presence of the agent in individual animals or products, for:
   a) movement or translocation
   b) human consumption.

5) Monitoring of the geographical distribution and prevalence changes due to management interventions (including determining immune status of individual animals or populations).

6) Studying agent, host and environment factors associated with disease occurrence.

1.2. Reference samples and sample quality

Reference samples should represent the target condition of interest, e.g. clinically diseased, subclinically infected. Experience indicates that selection of inappropriate positive reference samples from clinically affected animals when the test will be used to detect subclinical infection results in overly optimistic estimates of sensitivity and specificity. Experimentally infected animals may be the only source of reference samples in some cases but their use should be supplemented with samples from naturally infected animals, wherever possible.

By definition, all reference samples should be well characterised in terms of the host and its source population, and the infectious agent involved. Although the same description details would be desirable for reference samples from wildlife as compared with domestic animals, the relevant information is often not available. In such cases as many details as possible should be recorded. Minimum requirements to adequately characterise a reference sample are: a) the precise host species and subspecies when possible, b) tests used for confirmation of the presence or absence of the pathogen/antibody, c) geographical location with reference to known disease free or infected areas/regions, d) the date of sample collection and e) specimen type. Wherever possible, information on sex, age category (juvenile, sub-adult, adult), absence or presence of clinical signs, and a description of the signs will add value.

1.2.1. Pooling of reference samples

Ideally, reference samples should be obtained from individual animals and aliquoted into smaller volumes (weights) for subsequent testing. However, when animals of small body mass are the source of the reference samples or when very few animals are infected with the particular agent of interest, pooling of samples is acceptable to obtain a reference sample. Preferably the stage of infection of the individual animals should be known. A strongly positive sample of good quality can be diluted with the same sample matrix, for example faeces or serum, from the same host species to generate a series of samples with decreasing concentrations of the agent or products of the immune response. If certain stages of infection are not available, this should be documented.

In cases where only a limited volume of a suitable sample of good quality is available, it can be used as a reference sample in support of a well-planned set of test runs (e.g. for a repeatability study).

1.2.2. Negative reference samples and samples of unknown infection status

If negative reference samples are not available to determine diagnostic specificity in terms of certain agents known to cause cross-reactivity, this should be documented.

Latent-class statistical models for estimating diagnostic sensitivity and specificity in the absence of a perfect reference standard (sometimes termed a gold standard) are appealing for validating diagnostic assays for wildlife. This approach may be particularly useful for evaluating the sensitivity of nucleic-acid detection assays compared with viral, bacterial, and parasite isolation. Latent class analysis models have inherent assumptions and require a thorough description of the source population(s), which may be difficult or impossible to obtain for free-ranging wildlife (see the WOAH Validation Standard and Chapter 2.2.5 *Statistical approaches to validation* for details).

1.2.3. Sample quality

The sampling environment for wildlife is often suboptimal and may lend itself to cross-contamination. In addition, opportunistic sampling constitutes an important aspect in screening and monitoring wildlife populations for infectious agents. This often results in the collection of samples with compromised integrity (e.g. contamination, advanced autolysis). Therefore, investigators are responsible to determine the suitability of such samples for test validation, but given the overall scarcity of samples for certain conditions or from certain host species (e.g. endangered species), great care should be taken to ensure maximum utilisation of samples of sub-optimal quality. A qualitative assessment of sample quality (e.g. good, poor, autolysed) should be recorded in databases documenting the characteristics of reference samples.

Therefore it is deemed useful and necessary to validate appropriate tests for a range of sample condition criteria such as changes in detectability over time, under different storage temperatures, during autolysis, etc. However, this step in the validation process should be conducted after the test has been provisionally recognised.

2. Test validation pathways and stages for wildlife

2.1. Introduction

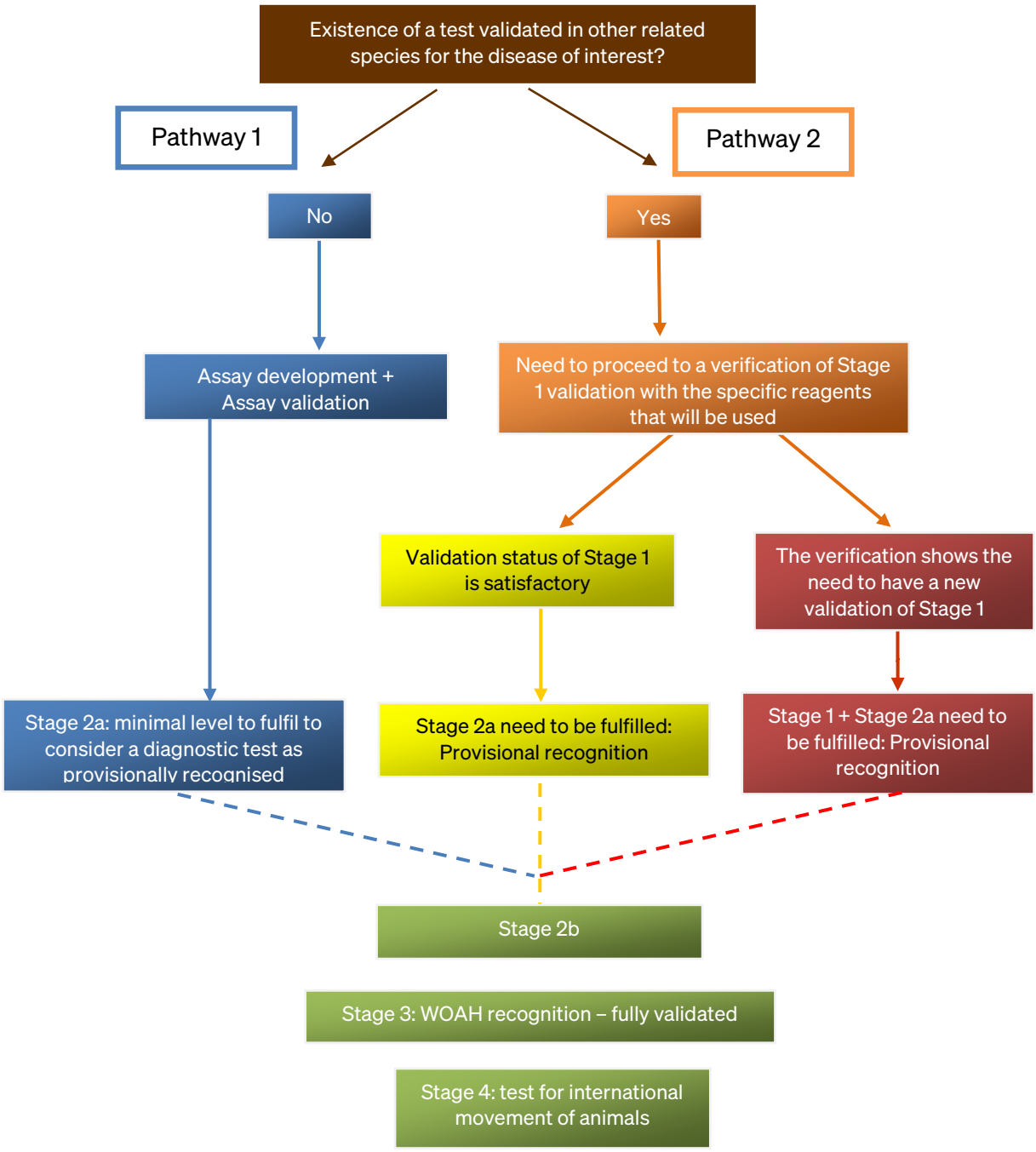
The two scenarios, considered in this Standard, involve the lack of availability (Pathway 1) or the availability of a validated test in another related species (Pathway 2) for the same pathogen. The flowchart (Figure 1) and the Table 1 show the stages in the validation process. Corresponding requirements to meet validation criteria and estimate test performance characteristics are shown in Table 1. Taxonomic relatedness of species should be a primary consideration when choosing the applicable pathway (Pathway 1 or 2, see Table 1), particularly when indirect test methods are applied. Other criteria such as behaviour of the animals, variation in pathogen strains or ecology of the diseases should also be taken into consideration. In most cases involving wildlife, Pathway 1 is appropriate because of the lack of validated tests in closely related species. When Pathway 2 is chosen, justification for its use should be provided by documenting the existence of a validated test.

Table 1. Validation Pathway: Steps required to meet validation criteria described in the WOAH Validation Standard and to estimate test characteristics. Requirements in the different stages need to be fulfilled with an acceptable outcome

Validation pathway: WOAH Validation Standard	Pathway 1: *No validated test in related species*	Pathway 2: *Validated test in related species*
Stage 1	**Stage-1 verified in new target species**	**Stage-1 verified in new target species**
Analytical specificity	Yes	Yes
Analytical sensitivity	Yes	Yes
Repeatability	Yes	No
Reproducibility (preliminary)	Yes	No
Stage 2	**Stage 2a (Provisional recognition)**	**Stage 2a (Provisional recognition)**
Diagnostic sensitivity	Yes (minimum of 30 positive reference samples)	Yes (minimum of 10 positive reference samples)
Diagnostic specificity	Yes (minimum of 30 negative reference samples)	Yes (minimum of 10 negative reference samples)
Cut-off determination	Yes (total of 60 samples)	Yes (total of 20 samples)
Reference sample description	Yes	Yes
	Stage 2b	**Stage 2b**
Diagnostic sensitivity	Yes	Yes
Diagnostic specificity	Yes	Yes

Validation pathway: WOAH Validation Standard	Pathway 1: *No validated test in related species*	Pathway 2: *Validated test in related species*
Cut-off determination	Yes	Yes
Reference sample description	Yes	Yes
Stage 3	**Stage 3**	**Stage 3**
Reproducibility	Yes	Yes
Repeatability	Yes	Yes
Stage 4	**Stage 4**	**Stage 4**
Predictive values (populations)	Yes	Yes

Fig. 1. Flowchart of pathways and stages of test validation in wildlife when a previously validated test exists or does not exist.



2.2. Additional considerations with regards to the general validation pathway as described in the WOAH Validation Standard

2.2.1. Stage 1 – Estimation of analytical characteristics

Estimation of analytical characteristics should follow recommendations set out in the WOAH Validation Standard.

Depending of the existence of diagnostic test methods validated according to the WOAH pathway for another species, the characteristics requiring validation will differ. If there is no validated diagnostic test method, all the characteristics of the Stage 1 should be evaluated (Pathway 1). If there is already a validated diagnostic test method, repeatability and reproducibility (preliminary) will not need to be reassessed until Stage 3 (Pathway 2).

Because the diversity of cross reacting organisms is often unknown, evaluation of analytical specificity can be more difficult than in domestic animals.

2.2.2. Stage 2 – Estimation of diagnostic characteristics

Estimation of diagnostic characteristics should follow recommendations set out in the WOAH Validation Standard.

For purposes of wildlife testing, Stage 2 is proposed to be divided into Stages 2a and Stage 2b. Stage 2a needs to be completed for “provisional recognition”, as previously described. In Stage 2a, the assumption is made that the pathway based on an existing validated test in a related domestic animal disease (compared with no validated test) was based on at least 10 positive reference samples and 10 negative reference samples and estimates of DSe and DSp are similar, if not identical, in the 2 species. These samples provide “credit” towards the reduced sample size (Pathway 1 vs. Pathway 2 in Table 1). Selection of the pathway with reduced sample size (Pathway 2) should be justified based on sample size and evidence of comparability (e.g. same test cut-off value and reagents) provided in peer-reviewed publications.

The selected sample size for completion of Stage 2 (Stage 2b) should be based on expected values for diagnostic sensitivity (DSe) and specificity (DSp), the desired confidence level and error margin as shown in Table 1 in the WOAH Validation Standard. For example for an expected DSe or DSp of 90%, a sample size of 138 is required to yield an error margin of 5% with 95% confidence (see right panel of Table 1 in the WOAH Validation Standard). However, it is acknowledged that this number of truly positive and negative reference samples may be difficult to obtain for some wildlife species and could potentially only be achieved when data from multiple testing laboratories using the same test in a standardised way are combined over time. Consequently, the initial number of samples tested may be lower than recommended numbers in Table 1of the WOAH Validation Standard.

If numbers of reference samples (positive and negative) are lower than numbers in Table 1 of the WOAH Validation Standard, the calculated error margins on estimates (typically represented as 95% confidence intervals) for DSe and DSp, respectively, will be wider than those on which the table was based. Consequently, small sample sizes increase the uncertainty in test performance characteristics. Use of reference samples that are representative of the target condition is critical to achieve an unbiased (and practically useful) estimate of DSe and DSp that will stand up to scrutiny over time. For example, samples should be obtained from subclinically infected animals if the test undergoing validation is to be used in apparently healthy animals. Obtaining and using representative samples of the target condition is therefore of greater importance than sample size.

The net effect of a lower sample size is greater uncertainty in estimates unless the prior information about the DSe and DSp in the related species is formally incorporated through a Bayesian analysis. Table 2 shows the effect of use of 140 or fewer known positive samples when the DSe estimate (90%) was calculated after field samples were collected and tested.

Table 2. Approximate error margins and 95% confidence intervals for diagnostic sensitivity (DSe) for decreasing numbers of positive reference samples

No. positive reference samples	No. positive	DSe (%)	Approximate error margin on estimate of DSe	95% exact binomial confidence interval for DSe (%)
140	126	90	± 0.05	83.8–94.4
100	90	90	± 0.06	82.4–95.1
60	54	90	± 0.08	79.5–96.2
30	27	90	± 0.10	73.5–97.9
10	9	90	± 0.18	55.5–99.7

Calculations for 95% confidence intervals for DSp are affected similarly by the number of negative reference samples that are used.

2.2.3. Stage 3 – Reproducibility

Generally the recommendations set out in the WOAH Validation Standard for the evaluation of the reproducibility are applicable, meaning that a minimum of 20 samples should be tested by three different laboratories in three distinct regions or countries. In cases where a particular test in wildlife is performed by very few laboratories or countries, or where the exchange of wildlife samples across international borders may be regulated by the CITES[1] agreement, the evaluation of the reproducibility may be postponed to a later stage when the test has been adopted by sufficient laboratories or the relevant CITES permit could be obtained.

2.2.4. Stage 4 – Interpretation of test results

Interpretation of test results (predictive values) in all species is dependent on knowledge of prevalence in the targeted population. This is difficult to know a priori in most free-ranging wildlife populations and even in captive populations where the population size is known, there may be substantial variations in prevalence among populations. Hence, it may be unreasonable to expect that predictive value calculations can be made with certainty in most wildlife populations. In the limited situations in which true prevalence can be determined, the predictive values of test results in these populations should not be extrapolated to other populations.

2.2.5. Monitoring assay performance after initial validation: modifications and enhancements – considerations for changes in the assay

Modifications in the protocol of the validated test may have an important impact on the performance of the test. Examples include: the use of body fluids collected from live or dead animals (such as ascites fluid, lung extract or pleural fluid) for an antibody detection test validated for serum; a change in the nature or source of reagents, and a change in cycling parameters of a PCR protocol.

Any modification will therefore require a limited re-evaluation of the analytical characteristics (Stage 1). If the characteristics are comparable with the initial protocol, with no significant change, the validation process can continue from the point where the change occurred. If the analytical characteristics change significantly, Stages 1 and 2a should be repeated in full.

*
* *

NB: FIRST ADOPTED IN 2014.

1 CITES: Convention on International Trade in Endangered Species of Wild Fauna and Flora