Appendix 1.1.4.3

Development and Optimisation of Nucleic Acid Detection (NAD) tests

Introduction

An ever increasing number of Nucleic Acid Detection (NAD) tests are now being used for diagnosis of infectious diseases in various species of animals and man. The most common methods are the Polymerase Chain Reaction (PCR) of which there are a number of variations and Isothermal Amplification Methods (such as LAMP). In addition, solid-phase or liquid phase microarrays are increasingly appearing as new tools of biotechnology-based diagnosis. The amplification techniques employed in these assays make them highly analytically sensitive. The product of the amplification reaction can be detected in a number of ways, for example, by visualisation in agarose gels, using labelled probes (e.g. TaqMan probes) to detect accumulation of product in real time or by using arrays where specific probes are captured on a solid-surface matrix or beads (8, 9). Different PCR assays can be multiplexed together to detect several targets in one tube or to combine targeted analytes and controls that generate different amplification products, all in one reaction vessel. Whilst this has obvious advantages, great care must be taken during optimisation to ensure that assay performance is not compromised. Similarly a multiple outcome PCR can be created using one set of primers, but employing tagged probes, which bind to different target sequences in the different species or strains detected by the PCR (13,14).

NAD amplification techniques are usually based upon the principle that there is an exponential amplification of the specific sequence targeted in the reaction. This makes the assay highly analytically sensitive but can, under some circumstances, result in contamination of other samples with the products of previous reactions. This is more likely to be a problem where reaction tubes are opened in the laboratory for further processing, for example, to run gels or to perform nested assays. To avoid such contamination, strict laboratory protocols should be employed involving separate rooms or cabinets for particular stages of the assays, changes of laboratory gowns and gloves, and stringent cleaning programmes. (9, 10). For these reasons, tests based on closed tube systems are generally more suitable for diagnostic assays (3). As with all assays, it is important to use appropriate controls to prove that the assay is performing as expected. All samples used in assay development should have well established provenance and the development should be carried out within the framework of a quality system to ensure appropriate levels of training, equipment maintenance and monitoring etc (5).

A. Assay Development Pathway

1. Definition of the intended purpose(s) for an assay

The first consideration in assay development is to clearly define the specific purpose and application of the test to be developed and to understand how it will be applied because this informs many of the decisions of the development pathway. For example one might choose to develop a screening test to detect all avian influenza (AI) strains in birds for which an inclusive and broadly reactive test is necessary, or to determine the haemagglutinin type, in which case a more specific test is needed. For some tests the requirement is to detect a group of viruses e.g. the pan peste virus PCR that detects Bovine Viral Diarrhoea (BVD), Border Disease and Classical Swine Fever (CSF) viruses. For other tests, the requirement may be to detect a single virus, or sometimes even to allow a DIVA approach. An example of such methods are the recently published real-time RT PCR assays for CSF virus which were developed for the genetic differentiation of naturally infected from vaccinated wild boar (12).

Purpose of the assay:
- Is it for a screening or confirmatory test, or both?
- Is it for detection of a group of pathogens?
- Is it for detecting a single virus or strain?
- Is it for discriminating between vaccinated *and infected animals*?

2. Assay Development

a) Quality Assurance

It is important that assays are developed in laboratories where high standards of quality assurance and control are employed (see Terrestrial Manual Chapter 1.1.3). The validation data for test performance and accuracy determined during the development and validation phase must be robust as it will form the basis for interpretation of disease status and consequent actions when the assay is routinely used.

b) Reference Materials

Sample selection (see Terrestrial Manual Chapter 1.1.1. Collection and shipment of diagnostic specimens)

For most diseases, the samples required for NAD assays are likely to be similar to those used for current detection methods such as bacterial or virus isolation. For example, detection of AI by real time RT PCR and virus isolation use the same samples. Recently, swabbing of fresh cuts in organ samples proved to be a practical approach, replacing the laborious homogenisation of organ samples. Similarly, there is a trend to use samples, which can be obtained by non-invasive methods, such as saliva for detection of PRRS in pigs (15) or bulk milk samples for

- What samples will be used from the target population?
- What are predilection sites for the agent in the host?
- What sample collection, storage and transport method is anticipated and what are the possible effects on results?
- Will samples be single or pooled?
- Will samples in the validation panel be representative of the target population

determination of herd BVD status. Prior to the selection of pooled samples, such as bulk milk, test developers need to consider and assess the implications of dilution of the target analyte on diagnostic sensitivity and incorporate this into the validation plan (see below).

It is important to understand the biology of the pathogen concerned and the nature of the sample collection devices. As in the case of AI, different strains of organisms have different predilection sites in host birds. For example cloacal samples are appropriate for some strains, whereas buccal samples are acceptable for others. Therefore, a test for use in a surveillance programme would need to incorporate both sampling sites and be validated using both types of sample. Another significant factor is the matrix in which the analyte resides in the host. Cloacal samples are more likely to contain PCR inhibitors than buccal samples. Another potential confounding factor is the type of swab used to collect the samples; some contain materials that inhibit PCR reactions. Therefore, it is very important to precisely specify the preferred sample material and to fully describe the swabs and swabbing protocol, including the preferred buffers or transport media and storage conditions.

Consideration also should be given to whether samples should be tested singly or in pools and whether pools should be of different samples from a single or multiple animals. Any pooling strategy should be precisely defined and validated prior to use. Finally, if the target population is "birds", the validation study should cover a large representative population of different species to demonstrate that the assay is widely applicable, whilst concentrating on the most prevalent species or those used as sentinels of infection.

c) Design of test Method

i) Choice of test

For most surveillance activities large numbers of samples may be tested first by a screening assay. In the above example for AI, the screening assay described must detect all known strains of Influenza A, either in a particular region or throughout the world. The test must be highly sensitive so it does not miss true positive samples, and analytically specific (inclusive) for detection of all viruses in the Influenza A group.

- Is the test for use only in a particular region or world-wide?
- Is it sufficiently sensitive and is the analytical target inclusive enough to not miss positive samples?
- Consider the impact of a high proportion of false positive results that cannot be confirmed…then what?
- Are rapid results possible and/or necessary?
- Will confirmatory tests (an analytical tool) be used to determine which strain?
- Will determination of the strain have important bearing on the action taken?

Avian Influenza is a high profile disease. If a new test should generate a high proportion of false positive results that cannot otherwise be confirmed, the infection status of the birds would be difficult to resolve. Exclusion of closely related agents that are not of interest within context of the intended purpose of the assay is essential.

It is also important for the test to produce rapid results so that disease control measures can be swiftly applied. The logical choice of tests would be a real time PCR using TaqMan probes or a field based assay. For a TaqMan assay, primers and probes are likely to be based upon the Matrix (M) gene which is known to be present in all influenza type A isolates. It might be important to develop confirmatory methods for use in conjunction with the screening test to determine which particular strain of AI is present and whether the strain is highly pathogenic, because this has an important bearing on the control measures that may follow. Such confirmatory methods, that further characterise samples, would be classed as analytical tools and only are applied to the subset of samples that test positive using the screening assay.

ii) Method design

The method should be carefully designed to fulfil the purpose that was defined at the outset. Important factors to consider will include application, types, and numbers of samples to be tested. This, in turn, leads to logistical and practical considerations for the candidate assay, such as whether it will be conducted in a laboratory, with or without automation to achieve high throughput, or in the field using a pen-side assay. All available information such as publications, sequences deposited in databases and in house sequence data should be utilised. Sophisticated software is available to optimise primer and probe design; computer modeling of probable sequences for use in the assay is a first step.

Has the design been shaped by the intended purpose of the assay?
- What is the specific application?
- What are the types and statistically relevant numbers of samples to be tested? (see annex 1.1.4.5).
- Is automation required?
- Should the test be field or lab based?
- Has exhaustive evaluation been done on publications of sequences?

d) Feasibility study

Before embarking on the validation of a newly developed assay, a feasibility study should be carried out using a small panel of approximately six or eight well-characterised samples to assess whether the system is viable. The panel should consist of samples distributed across the operating range of the assay, i.e. at least two negatives, at least two unambiguously positives, and ideally samples falling mid-range. Ideally these samples should come from different animals. The assay should achieve as wide-as-possible separation of test results for the high positive and negative samples in this panel. It may be prescient to test a dilution series (analyte diluted in matrix) at this point to assess relative analytical sensitivity if the test is a potential replacement for another method where sensitivity is an important criterion (See Appendix 1.1.4.6 on Methods Comparability).

- Was the assay first evaluated on a small sample panel (6-8 samples) to assess viability of the approach?
- Was there good separation between test results of negative and positive samples?
- Was a preliminary test on a dilution series in matrix to preliminarily assess relative analytical sensitivity made (if that is an important criterion)?

e) Development & optimization

The aim of this stage is to define and optimise the method that will be used to carry out the test in future routine testing. This includes description of appropriate facilities to carry out the assay. It is important to consider both the method of sample extraction required to prepare samples for the test and the assay procedure. For any PCR assay, definition of clean room protocols is required to minimise contamination. If large numbers of samples are anticipated, selection of an automated extraction procedure may be essential (6).

- Optimal location? Use a clean room or cabinet system to minimise contamination
- Optimal method of sample extraction to prepare samples? Done manually or by automation?

Usually, it is necessary to assess a number of different test methodologies and vary concentrations of reagents, template additions and reaction times to optimise the extraction and the assay. It is important to either change only one variable at a time or use a multi-factorial design and analysis (see Validation Chapter, section A.2.f on Robustness). It is important to identify which factors have only a narrow range of optimal activity, as these are critical points in the assay procedure and may affect an assay's robustness. Generally, the assay stages are relatively simple to optimise, but care needs to be taken to ensure that a robust extraction method has been developed which is suitable for routine laboratory application. (5)

- Have concentrations of reagents been tested for optimal reactivity?
- Has extraction been optimised?
- What are the template additions and reaction times to optimize the assay?
- Which factors have a narrow range in which they perform optimally?
  - Have you defined that range ?

Controls for PCR are many and varied. It is important to include appropriate controls to show that the assay is performing as expected. The various controls that need to be considered include:

i) **A host-species control**. This control demonstrates that the swab had been in contact with a target species and that the sample contained "bird" nucleic acid, which is available for extraction using the defined protocol. The most suitable target for this control is a housekeeping gene such as beta actin, which is present in the target host species. This is relatively simple if the host species is a single species e.g. chickens, but it will be more challenging to find a suitable target for all "bird" samples. This type of control may not be required if a sample is added directly to the extraction process, such as a piece of tissue or blood, but is recommended for "indirect" samples such as those collected on swabs; in this case, a the host-species control should be used for every sample tested.

ii) **No-template control.** This control reveals whether contamination of the sample has occurred, resulting in amplified product when no amplified product should be present, as in a sample containing no targeted sequence. Consideration should be given to the number and placement of no template controls in the assay set up template. Generally, a number (approximately 5% of wells) of no-template controls are distributed randomly over the plate.

Have you included all necessary controls to prove the assay is performing as expected?

- template control?
- Inhibition control?
- Positive sample control?
- For RNA assay – reverse transcription control?

iii) **Analyte-positive control.** This positive control, having Ct activity within the defined operating range of the assay, is used on each plate. The most suitable choice for this control may be a plasmid containing the target sequence, which can be used to check for the expected level of amplification in the assay. However, this does not assess the efficacy of the extraction process, which requires a known field sample or its equivalent.

iv) **Inhibition control.** This control is needed to detect possible inhibitors of the PCR reaction. If an inhibition control yields a negative result, this infers that the sample contains inhibitory substances. and that a negative test result for the test sample cannot be interpreted as "negative" because of likely inhibition. Certain samples such as faeces and semen often contain inhibitors whereas this is less of a problem when testing blood samples or cultured organisms. (4). Data collected during the validation process will allow for a risk-based decision as to whether an inhibition control should be included for each sample or the test system is unlikely to be affected by inhibition. If inhibitory substances are a significant problem an inhibition control must be included for each test sample.

- Did you consider all of the possible pitfalls in use and interpretation of results for this inhibition control?
- Is the purpose and application of the inhibition control clearly specified in the assay protocol?

There is considerable debate about what to use as the most suitable and effective inhibition control. Generally this could be one of the following:

- An artificial target, such as a length of DNA contained in a plasmid, which is added to the extracted sample and amplified with the same primers as the test target, but is of a different size or contains a different internal sequence so that it is identified as the internal control when the detection method is applied. The advantage of this approach is that it utilises the same primers employed in the test, and the control can be added at precise concentrations. However care must be taken during assay optimisation, so that the analytical sensitivity of the assay is not detrimentally affected because of competition for primers. The other disadvantage is that as an added component, it only controls the assay stage of the test and does not act as a control for the extraction stage.

- An alternative strategy is to amplify a housekeeping or structural gene such as β-actin which always is present in the target tissue, and thus the sample. If the housekeeping gene has been inhibited by substances in the sample, the inference is that amplification of the gene targeted by the test may also be inhibited. This conclusion is not always warranted. Housekeeping genes are often present in abundance, so sometimes can be detected even in the presence of inhibitory substances. But the more limited amounts of the assay's targeted sequences may be inhibited from amplification. In this case, the amplified and detected housekeeping gene was not a sufficient control for inhibitors, resulting in a false negative inference for the test result. However, if properly used, an inhibition control, which is naturally present in the sample controls can be a useful control for inhibitors for the whole assay including sampling, storage and extraction.

- For assays based on RNA targets, a control can be included to assure that reverse transcription of RNA to cDNA has proceeded appropriately. This control should be included during the validation process to assess whether reverse transcription is likely to be an issue. A risk-based decision can be made following analysis of the validation data as to whether the control should be included routinely when the assay is run.

f) Inhibitory factors in the sample matrix

Generally for NAD methods, pure cultures, blood and most tissues are relatively easy to work with and extraction of amplifiable NA is generally successful. Faecal samples, semen and autolysed tissue can be more challenging because they generally contain more inhibitors. It is vital to have a robust and repeatable sample extraction procedure, which is appropriate for the numbers of samples to be handled (automated if necessary) and to utilise inhibition controls if necessary (see above section).

g) Operating range of the assay

The operating range of the assay should be determined by diluting out a high positive sample and plotting the range of results obtained vs known amounts of nucleic acid (concentration, dilution, number of genomic copies, etc). This reference sample must be in the same matrix as the test sample, i.e. it is not appropriate to determine the operating range of a sample diluted out in buffer, if the usual matrix is blood.

- For determining the operating range of the assay, did you dilute the samples in matrix of samples for which the test is intended?
- Does the operating range of the assay conform with the expected norms for such an assay?

h) Robustness

An assay should tolerate small changes in concentrations of reagents and/or slight variations in processing times and temperatures used for different stages of the assay. This can be determined during assay optimisation when critical stages are identified. Such stages that tend to be variable should be well described in the assay protocol so that particularly exacting processes are assured for carrying them out. This is a laborious process that ultimately is monitored for precision and accuracy by internal and external quality control samples run in the assay.

- Does the assay tolerate small changes in reagent concentration or in physical parameters?

i) Calibration of the assay to reference samples

Ideally, international or national reference standards should be used to calibrate the assay. However, these are not always available so it may be necessary to create an in-house reference standard. A working standard(s), for inclusion in all runs of the

- Have you calibrated the assay to external reference standards?
- If external reference standards not available, have you created an in-house reference standard?
- Have you made working reference standards in sufficient amounts for use in all development and validation experiments?

assay, needs to be created, aliquotted, and stored in sufficient quantities for use in every run of the validation process and for routine use after the validation has been completed. The working standard(s) could be multiple aliquots of a particular sample, which can be used within each assay run. They could also consist of a plasmid containing the sequence of interest, spiked into sample matrix. Use of the latter allows the test developer to determine the number of genome copies that can be detected by the assay. In some instances test sample results are "normalised" by comparison to the working standard sample(s) included in each run of the assay. This allows direct comparison of data between runs (reference 7 and Validation Chapter 1.1.4).

B. Assay Validation Pathway

Once the protocol for the assay has been developed and optimised, it must be fixed and held constant while being evaluated through the stages of the validation pathway.

1. Stage 1 - Analytical Performance criteria

a) Repeatability

- Has intra- and inter-assay repeatability been determined?
- Is repeatability within the accepted range of coefficient of variation (CV) limits of 15% to 20%?

Repeatability of the assay is a measure of agreement between results (within and between runs) using the same test method in one laboratory. Usually a small panel of three (preferably 5) samples covering the operating range of the assay is selected and tested using the entire assay procedure (including nucleic acid extraction). Within assay (intra-assay) variation is determined using multiple (at least 5) replicates of each sample in this panel in one assay run. Between run (inter-assay) variations are determined by testing these samples over several days, using several operators and at least 20 runs. The repeatability panel should be tested treating all samples and each of their replicates exactly like individual diagnostic samples, subjected to every step from sample preparation to data analysis. Accordingly, every replicate of every sample is subjected to an independent extraction. This allows for determination of repeatability, both with and between runs of the assay that mimic future runs of the assay when implemented for diagnostic use. Minimal variation in repeatability is important, particularly near the cut-off(s) that establish positive, inconclusive and negative ranges, because higher variability can result in incorrect interpretations (see Appendix 1.1.4.4 on measurement uncertainty).

For real-time PCR, a variation of 1-2 Ct values for a single retested sample is considered acceptable. Repeatability can also be expressed as a coefficient of variation; in this case a CV of 5-10% is acceptable. The assay should be designed so that the decision point (cut-off) lies on the steepest part of the real-time PCR Ct curve. If this is achieved, repeatability will be optimal at the critical point of the assay. (Larger CV's at the clearly negative and highly positive ends of the operating range of the assay do occur and have little impact on test result interpretation.)

b) Analytical specificity (ASp)

Depending upon the intended purpose of an assay, its analytical specificity is determined by the selected genetic sequence(s) of an organism(s) targeted by the assay. The assay can be designed to be highly selective, with analytically specificity for a single genetic sequence that is not known to be present in other organisms or strains of the targeted organism. Such an assay is said to exhibit exclusivity and connotes a confirmatory assay of high ASp.

Alternatively, the assay may be designed to target a conserved genetic sequence that is common to several strains of a given species, or several species of a genus. Such an assay has an ASp that exhibits inclusivity, making it useful as a screening test. For an inclusive screening assay, the analytical specificity should be determined by testing all lineages, strains, species, etc., that the assay is expected to detect. The assay should then be evaluated for its capacity to exclude related organisms such as non-pathogenic strains that are not of interest to the intended purpose of the candidate assay.

- Has a panel of as many well-characterized isolates of the target pathogen as possible, including isolates from a variety of geographical areas and hosts, been tested?
- Have related organisms and pathogens, which cause similar clinical syndromes, been tested?

In the example of an AI screening test, the assay should be run against as many well-characterised isolates of AI virus as are available to assure that all strains from a variety of geographical areas and hosts are detected (i.e., to demonstrate inclusivity). This is generally done using laboratory strains/cultures or nucleic acid. For AI, different lineages of the virus exist such as the Asian, European and North American strains. It is important to consider how the assay will be used and in which geographical regions. That will aid in determining whether it is necessary to evaluate some or all of these lineages. Another important factor is that viruses can change rapidly and mutations can render a test sub-optimal. An example of this is the appearance of the recent pandemic strain of H1N1. The analytical sensitivity of the PCR for the traditional M gene was impaired for this strain because the sequence that rendered the assay specific had mutated. In most countries new primers have been introduced and used either as a new test or in combination with the traditional M gene primers as a combination test.

The discriminatory power of the assay should be checked by testing organisms related to the AI virus. These would include pathogens which cause similar clinical syndromes such as Newcastle Disease, Infectious Bursal Disease etc, and other organisms, which are likely to be found in the target sample (i.e., to demonstrate exclusivity).

c) Analytical sensitivity (ASe)

- What is the detection limit of the assay?
- How does the limit of detection compare to current standard methods?
- Is the ASe sufficient that it will be fit for its intended purpose?

There are two common approaches to assessing analytical sensitivity, also known as the limit of detection. The first is to use a dilution series of the target pathogen (in this case AI virus) diluted in sample matrix and not buffer. The dilution series is usually tested using the new assay under validation and a standard method. For AI, the standard method could be virus isolation or another in-house standard method of detection. This approach yields a comparative measure of the two methods. The second approach is to use a

plasmid construct, containing the target sequence and test this as a dilution series in sample matrix. In this manner, the number of genome copies detectable by the test method can be estimated.

d) Standard test method for comparison with the candidate test method (see Appendix 1.1.4.6 on Methods Comparability)

- Has the new test performed in a satisfactory manner compared to a standard method of comparison?
- Is preliminary reproducibility data acceptable
- Does the assay merit proceeding to studies on full validation (Stages 2 – 4)?

On some occasions it is not possible to complete a full validation exercise either because samples of known analyte status are scarce (e.g. exotic diseases) or, when an emergency situation arises, the assay is required for use before it can be fully validated. Provisional recognition can be achieved provided that results through Stage 1 of the validation process compare favourably with results of a standard test method or a known established and preferably published method. Another choice for a standard method is the one used routinely in the laboratory. It is important to recognise that different methods identify different morphological or functional entities of the organism. It is therefore possible that comparison between a standard culture-based method and a new NAD technique will give rise to discrepant results (see section B.2, Stage 2, below, for discussion on resolution of discrepancies). The choice of assessment panel is also very important and it should be as extensive as possible (see Appendix 1.1.4.7 on reference panels). If only a small panel of samples is available for evaluation, it is useful to determine if the assay withstands the rigors of use in other laboratories. This requires that both laboratories use the same protocol, and same reagents, same panel of samples and similar (if not identical) equipment.

If lack of samples prevents continuing through the next stage of the validation pathway (Diagnostic Performance of the assay), it is acceptable to use a NAD assay that has been provisionally accepted by having been thoroughly validated through Stage 1 of validation (Validation Chapter 1.1.4, section B.2.g). Acceptance of provisional validation is fully dependent upon approval by local authorities, or through bilateral agreements between countries.

e) Analytical accuracy

Analytical tools, designed to provide information to characterise samples that are detected using a screening assay, only require validation of their analytical performance characteristics. So, there is no requirement to determine diagnostic sensitivity or diagnostic specificity in such cases. Examples of such approaches include PCR typing assays to determine whether a matrix-positive influenza A strain is H5 or H7, or methods to determine antibiotic resistance which are only applied to cultured bacteria. A nucleic acid based technology employed for such tests includes nano-array based methods, which introduce their own challenges due to the large amount of data generated for each sample tested. Before implementing such assays, consideration should be given to how this analysis can be accomplished. A simpler method approach is to compare the results of a new analytical tool to a standard tool and permit its use as long as the results of the new and existing technique compare favourably (references1, 2, and Appendix 1.1.4.6).

2. Stage 2 - Diagnostic Performance Criteria

Diagnostic Sensitivity and Diagnostic Specificity

Diagnostic sensitivity (DSe) and specificity (DSp) are the primary performance indicators of a diagnostic assay. Approaches for the use of reference standards (methods) in defining the infection/disease status of potential reference animals or populations and their limitations in determining estimates of DSe and DSP are discussed in more detail in Appendix 1.1.4.1, Section B.2.b.i.

When determining these estimates it is vital to select sufficient numbers of samples which are relevant to the target population for the test under assessment. It can be difficult to obtain large numbers of samples (particularly positive samples) for some exotic diseases, which are not commonly found. In such cases, with few samples available, the amount of error allowed in estimates of DSe and DSp, of necessity, may be rather large (see Validation Chapter 1.1.4, section B.3, Table 1).

Negative reference samples are often selected from animals living in regions where the disease is not present, while positive samples are usually obtained from animals with clinical signs which have been confirmed in the laboratory. This can lead to overly optimistic estimates of DSe and DSp because the samples do not represent the whole spectrum of the disease process, ranging from non-clinical animals which may have pathogen loads that are much different than animals experiencing fulminant or chronic disease.

Samples are often categorised using current test methods such as virus isolation (VI) or bacterial culture. However, this can be problematic when validating new molecular tests, because the basis of the two test systems is different. For example, a positive bacterial culture is dependent upon the presence of a viable organism whereas nucleic acid based methods detect genomic sequences of both live and dead organisms as long as the nucleic acid is still present in the sample. VI methods can be particularly susceptible to inhibitors and contaminants present in the sample matrix, leading to an underestimate of “true positives”. This can result in apparent discrepancies where samples are positive using the new molecular test and negative by traditional methods. Various strategies for resolving such anomalies include, but are not limited to, sequencing (which can demonstrate that the pathogen of interest was present in a particular sample), or testing using another molecular approach.

An alternate method is available if samples from animals of known infection status are difficult to obtain. If an appropriate sampling design can be employed and different independent test methods used to test the samples, it is possible to obtain estimates of DSe and DSp by using Bayesian methods (latent class models), under certain circumstances. This has been particularly successful for resolving the lack of samples of known infection status (see Appendix 1.1.4.5).

- **Threshold and cutoff** are considered to be synonymous. A cutoff is the test value selected for distinguishing between negative and positive results on a continuous scale of test values.
- **Indeterminate, intermediate, suspicious, borderline, grey zone or equivocal** are terms used synonymously for a zone of test values falling between the positive and negative cutoffs.

To calculate DSe and DSp estimates of the candidate assay, the test results first must be reduced to categorical (positive, negative, or indeterminate) status. This is accomplished by insertion of one or two cut-off points (threshold or decision limits) on the continuous scale of test results. For example, in some circumstances it is appropriate to use a cut-off for a real time PCR assay in the region of 35Ct, which means that some samples that produce a higher Ct value are categorised as negative or inconclusive. For a different PCR assay, however, any sample which merely registers a Ct may be categorized as positive. The performance of a particular real-time PCR, comparative validation data, the ultimate application of the results generated, and any relevant veterinary information should be taken into account when considering the use of a cut-off.

3. Stage 3 - Reproducibility and augmented repeatability estimates

Reproducibility is a measure of the agreement between results obtained in different laboratories using the same protocol, similar (preferably the same) equipment and the same panel of samples. Ideally the panel would consist of 20-30 samples including a few which are present as quadruplicates. The panel should consist of samples that cover the dynamic range of the test with several that have activity close to, and on either side of, the test-cutoff(s). The same panel used for determination of repeatability could be used for this evaluation, but with enhanced numbers of replicates. Measurements of precision can be estimated for both the repeatability and reproducibility data (see appendix 1.1.4.4 on the Measurement of Uncertainty for further explanation of this topic and its application). Appendix 1.1.4.5 provides further information about the selection and use of reference panels.

4. Stage 4 - Programme implementation

a) Interpretation of test results

Best practices for programme implementation are general to all assay types (see Validation Chapter 1.1.4). For PCRs, an inherent advantage is the possibility of follow-up genomic sequencing to resolve apparent false positive results. Assays including PCRs are often validated using similar numbers of positive and negative samples. However, in surveillance programs, they are often applied to affirm the absence of the disease in question in locales where disease prevalence is very low and often approaching zero. In such circumstances, false positive results can be a significant problem even if the diagnostic specificity of a particular assay is high. If the DSp of an assay is 99.5% this means that one in 200 test-positive results will be false if the prevalence is close to zero. If a large number of samples are tested from a population of zero or very low prevalence, such false positive results can significantly out-number true positive results (see section B.5.a – Stage 4 of chapter 1.1.4 for further explanation of predictive values of test results as a function of disease prevalence). For NAD assays used in such circumstances, it would be good practise to confirm PCR-positive results by sequencing.

5. Monitoring of assay performance after initial validation.

a) Monitoring the assay

Monitoring of repeatability by charting the values obtained for working standard control samples provides re-assurance that the assay is performing as expected. Similarly, participation in proficiency testing schemes issued by external providers of quality assurance samples provides evidence of ongoing reproducibility and also allows comparison of test accuracy if a reference standard(s) is included in each run for "normalization" of data. Re-testing of a proportion (usually in the range of 1-5% depending on throughput) of retained samples is also employed by some laboratories to demonstrate that the assay is performing consistently between runs.

b) Minor modification of an existing validated assay

In time it may be necessary to modify the assay because the target analyte has changed, e.g., if the assay for avian influenza is to be applied in another part of the world or if new strains or lineages of a virus have emerged (see section B.1.b, above, on the evolution of the new pandemic lineage of H1N1). RNA viruses evolve rapidly and point mutations can occur, so it is advisable to regularly confirm the nucleotide sequences of the primer and probe sites to ensure that they remain appropriate.

i) Technical Modifications

Modifications of an assay are likely to be required, over time. For example, use of different equipment, use of a different extraction protocol or automation of particular stages will minimally require comparison of the original validated assay with the modified version (see Appendix 1.1.4.6). If results of the modified version fall outside of the operating range of the original assay, a revalidation may be necessary.

ii) Replacement of depleted reagents

It is important to assign unique identification numbers to all batches of reagents and to record the components used for particular assays. The most critical components in NAD assays are probes and primers. Current and new batches should be tested in parallel prior to their introduction. However, for other reagents it is usually sufficient to monitor batches to inform troubleshooting, should that become a necessity.

c) Major change in assay requiring revalidation

Upon occasion, application of the assay may need to be extended beyond the scope of the original intended purpose of the assay. Examples are inclusion of another host species or a population of animals from a different geographical area. In such cases it is important to revalidate the assay because of new biological considerations with their many associated variables. The precise details will depend on the extent of the change. Moving the assay into a new geographical area might mean that the analytical characteristics of the assay are still valid but that the diagnostic criteria need to be re-defined. Similarly modifications may be made to primer or probe sequences to allow detection of new strains. It will then be necessary to demonstrate how the new reagents behave in terms of analytical and diagnostic accuracy compared with the previous version of the assay.

REFERENCES

1. Anjum, MF; Mafura, M; Slickers, P; Ballmer, K; Kuhnert, P; Woodward, M.J; and Ehricht, R (2007). Pathotyping of E coli by using miniturized DNA microarrays. App. & Env. Micro. 73 (17) 5692-7

2. Batchelor, M; Hopkins, K.L; Liebana, E; Slickers, P; Ehricht,R; Mafura, M; Aerestrup, F; Mevius, D; Clifton-Hadley, FA; Woodward, M.J; Davies, R.H; Threlfall, E.J and Anjum, M.F (2008) Development of a miniaturised microarray based assay for the rapid identification of antimicrobial resistance genes in Gram-negative bacteria. International Journal of Antimicrobial Agents 31: 440-51

3. Sawyer,J; Wakeley,P; West,D; Fearnley,C. Anderson,S; Flowers,M. Webster,K;, Erington,J and Williams, R (2006) practical Experiences of moving molecular diagnostics into routine use at the Veterinary Laboratories Agency. Dev. Biol. (Basel) 126: 89-97

4. Ballagil-Pordány, A. & Belák, S. (1996). The use of mimics as internal standards to avoid false negatives in diagnostic PCR. Mol. Cell. Probes, 10, 159–164.

5. Burkhardt H.J. (2000). Standardization and quality control of PCR analyses. Clin. Chem. Lab. Med., 38, 87–91.

6. Jungkind D. (2001). Automation of laboratory testing for infectious diseases using the polymerase chain reaction – our past, our present, our future. J. Clin. Virol., 20, 1–6.

7. Huggett J., DhedaA K., Bustin S. & Zumla A. (2005). Real-time RT-PCR normalisation; strategies and considerations. Genes Immun., 6, 279–284. (Review).

8. Lauerman L.H. (2004). Advances in PCR technology. Anim. Health Res. Rev., 5, 247–248. (Review).

9. Viljoen G.J., Nell H & Crowther J.R. (2005). Molecular Diagnostic PCR Handbook. IAEA-FAO, Springer, Dordrecht, The Netherlands.

10. Subcommittee of Animal Health Laboratory Standards (SCAHLS) Veterinary Laboratory Guidelines for nucleic acid detection techniques (http://www.scahls.org.au/

11. Vilcek, S. Nettleton,P. and Paton,D (2003) J Virol Methods 114 (2): 167-70

12. Liu L, Hoffmann B, Baule C, Beer M, Belák S, Widén F (2009) Two real-time RT-PCR assays of classical swine fever virus, developed for the genetic differentiation of naturally infected from vaccinated wild boars. J Virol Methods.159(1):131-3.

13. Wakeley, P. R., J. Errington, S. Hannon, H. I. Roest, T. Carson, B. Hunt, J. Sawyer, and P. Heath. 2006. Development of a real time PCR for the detection of Taylorella equigenitalis directly from genital swabs and discrimination from Taylorella asinigenitalis. Vet Microbiol 118:247-54.

14. Wakeley, P. R., N. Johnson, L. M. McElhinney, D. Marston, J. Sawyer, and A. R. Fooks. 2005. Development of a real-time, TaqMan reverse transcription-PCR assay for detection and differentiation of lyssavirus genotypes 1, 5, and 6. J Clin Microbiol 43:2786-92.

15. Prickett, J., Christopher-Hennings, J., Yoon, K-J., Evans, R. B., Zimmerman, J.J. (2008) Detection of *Porcine reproductive and respiratory syndrome virus* infection in porcine oral fluid samples: a longitudinal study under experimental conditions. Journal of Veterinary Diagnostic Investigation 20: 156-163.

Further readings (a suggestion):

Belák, S. Thorén, P.: Validation and quality control of polymerase chain methods used for the diagnosis of infectious diseases. Office International des Epizooties (OIE), Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (mammals, birds and bees). Sixth Edition, 2008, Chapter I.1.5. 46-55.http://www.oie.int/eng/normes/mmanual/2008/pdf/1.1.05._VALID_PCR.pdf

Belák, S., Thorén, P., LeBlanc, N., Viljoen, G.: Advances in viral disease diagnostic and molecular epidemiological techniques. Expert Reviews for Molecular Diagnostics, 2009. 9. 367-381.