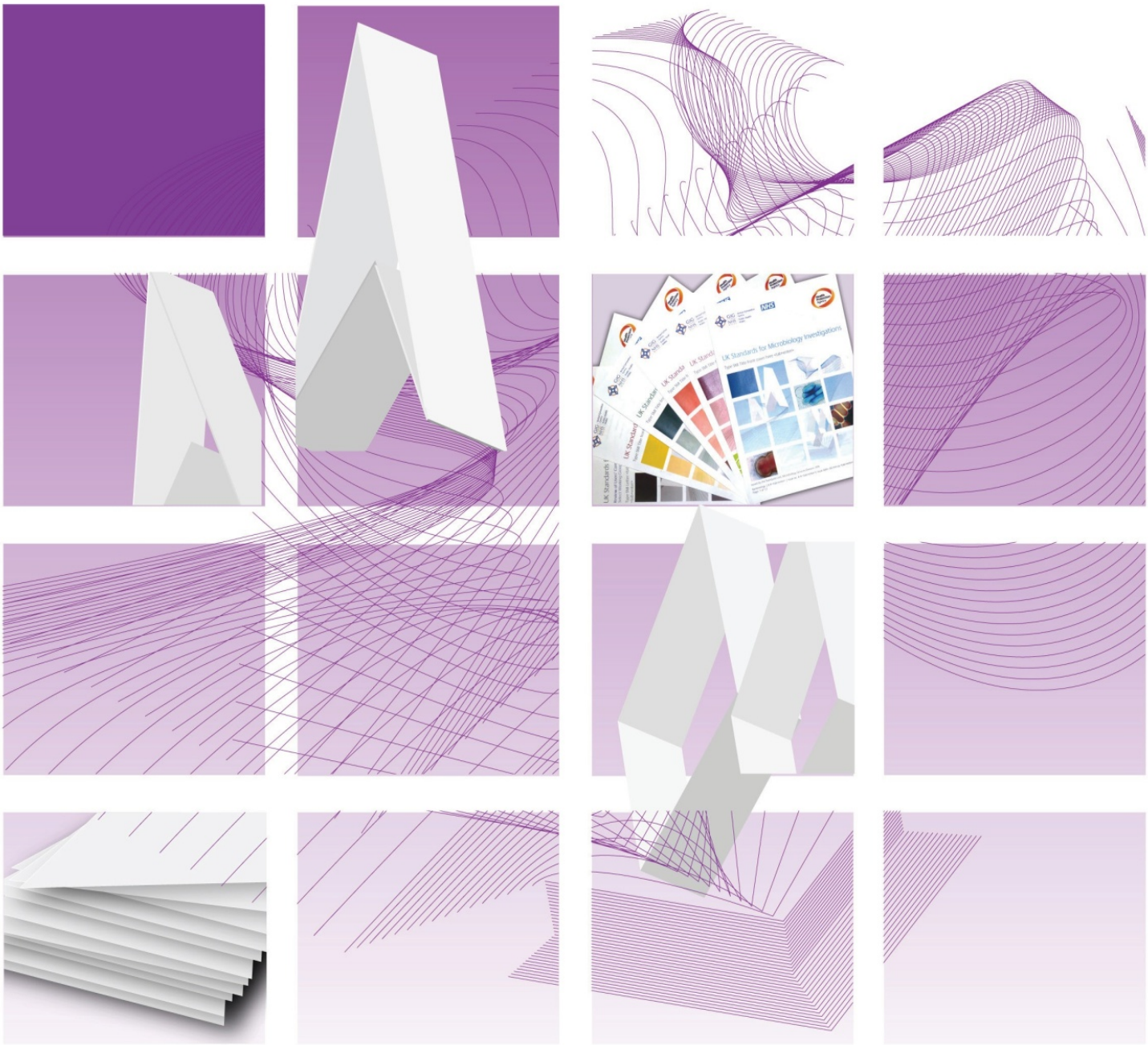
UK Standards for Microbiology Investigations

Guidance on the Development and Validation of Diagnostic Tests that Depend on Nucleic Acid Amplification and Detection



Issued by the Standards Unit, Microbiology Services Division, HPA

UK Protocols | P 4 | Issue no: 1| Issue date: 01.02.13 | Page: 1 of 49

Acknowledgments

UK Standards for Microbiology Investigations (SMIs) are developed under the auspices of the Health Protection Agency (HPA) working in partnership with the National Health Service (NHS), Public Health Wales and with the professional organisations whose logos are displayed below and listed on the website http://www.hpa.org.uk/SMI/Partnerships. SMIs are developed, reviewed and revised by various working groups which are overseen by a steering committee (see http://www.hpa.org.uk/SMI/WorkingGroups).

The contributions of many individuals in clinical, specialist and reference laboratories who have provided information and comments during the development of this document are acknowledged. We are grateful to the Medical Editors for editing the medical content.

We also acknowledge Professor Nick Saunders for his considerable specialist input.

For further information please contact us at:

Standards Unit
Microbiology Services Division
Health Protection Agency
61 Colindale Avenue
London NW9 5EQ
E-mail: standards@hpa.org.uk

Website: http://www.hpa.org.uk/SMI

UK Standards for Microbiology Investigations are produced in association with:

UK Standards for Microbiology Investigations[#]: Status

Users of SMIs

Three groups of users have been identified for whom SMIs are especially relevant:

- SMIs are primarily intended as a general resource for practising professionals in the field operating in the field of laboratory medicine in the UK. Specialist advice should be obtained where necessary.
- SMIs provide clinicians with information about the standard of laboratory services they should expect for the investigation of infection in their patients and the documents provide information that aids the electronic ordering of appropriate tests from hospital wards.
- SMIs also provide commissioners of healthcare services with the standard of microbiology investigations they should be seeking as part of the clinical and public health care package for their population.

Background to SMIs

SMIs comprise a collection of recommended algorithms and procedures covering all stages of the investigative process in microbiology from the pre-analytical (clinical syndrome) stage to the analytical (laboratory testing) and post analytical (result interpretation and reporting) stages.

Syndromic algorithms are supported by more detailed documents containing advice on the investigation of specific diseases and infections. Guidance notes cover the clinical background, differential diagnosis, and appropriate investigation of particular clinical conditions. Quality guidance notes describe essential laboratory methodologies which underpin quality, for example assay validation, quality assurance, and understanding uncertainty of measurement.

Standardisation of the diagnostic process through the application of SMIs helps to assure the equivalence of investigation strategies in different laboratories across the UK and is essential for public health interventions, surveillance, and research and development activities. SMIs align advice on testing strategies with the UK diagnostic and public health agendas.

Involvement of Professional Organisations

The development of SMIs is undertaken within the HPA in partnership with the NHS, Public Health Wales and with professional organisations.

The list of participating organisations may be found at http://www.hpa.org.uk/SMI/Partnerships. Inclusion of an organisation's logo in an SMI implies support for the objectives and process of preparing SMIs. Representatives of professional organisations are members of the steering committee and working groups which develop SMIs, although the views of participants are not necessarily those of the entire organisation they represent.

[#] UK Standards for Microbiology Investigations were formerly known as National Standard Methods.

Microbiology is used as a generic term to include the two GMC-recognised specialties of Medical Microbiology (which includes Bacteriology, Mycology and Parasitology) and Medical Virology.

SMIs are developed, reviewed and updated through a wide consultation process. The resulting documents reflect the majority view of contributors. SMIs are freely available to view at http://www.hpa.org.uk/SMI as controlled documents in Adobe PDF format.

Quality Assurance

The process for the development of SMIs is certified to ISO 9001:2008.

NHS Evidence has accredited the process used by the HPA to produce SMIs. Accreditation is valid for three years from July 2011. The accreditation is applicable to all guidance produced since October 2009 using the processes described in the HPA's Standard Operating Procedure SW3026 (2009) version 6.

SMIs represent a good standard of practice to which all clinical and public health microbiology laboratories in the UK are expected to work. SMIs are well referenced and represent neither minimum standards of practice nor the highest level of complex laboratory investigation possible. In using SMIs, laboratories should take account of local requirements and undertake additional investigations where appropriate. SMIs help laboratories to meet accreditation requirements by promoting high quality practices which are auditable. SMIs also provide a reference point for method development. SMIs should be used in conjunction with other SMIs.

UK microbiology laboratories that do not use SMIs should be able to demonstrate at least equivalence in their testing methodologies.

The performance of SMIs depends on well trained staff and the quality of reagents and equipment used. Laboratories should ensure that all commercial and in-house tests have been validated and shown to be fit for purpose. Laboratories should participate in external quality assessment schemes and undertake relevant internal quality control procedures.

Whilst every care has been taken in the preparation of SMIs, the HPA, its successor organisation(s) and any supporting organisation, shall, to the greatest extent possible under any applicable law, exclude liability for all losses, costs, claims, damages or expenses arising out of or connected with the use of an SMI or any information contained therein. If alterations are made to an SMI, it must be made clear where and by whom such changes have been made.

SMIs are the copyright of the HPA which should be acknowledged where appropriate.

Microbial taxonomy is up to date at the time of full review.

Equality and Information Governance

An Equality Impact Assessment on SMIs is available at http://www.hpa.org.uk/SMI.

The HPA is a Caldicott compliant organisation. It seeks to take every possible precaution to prevent unauthorised disclosure of patient details and to ensure that patient-related records are kept under secure conditions.

Suggested Citation for this Document

Health Protection Agency. (2013). Guidance on the Development and Validation of Diagnostic Tests that Depend on Nucleic Acid Amplification and Detection. UK Standards for Microbiology Investigations. P 4 Issue 1. http://www.hpa.org.uk/SMI/pdf.

Contents

Amendment Table

Each SMI method has an individual record of amendments. The current amendments are listed on this page. The amendment history is available from standards@hpa.org.uk.

New or revised documents should be controlled within the laboratory in accordance with the local quality management system.

Amendment No/Date.	-/01.02.13
Issue no. discarded.	-
Insert Issue no.	1
Section(s) involved.	**Amendment.**

Scope of Document

This SMI describes a generic framework for in-house assay development and validation of new nucleic acid amplification assays including real-time Polymerase Chain Reaction (PCR). The framework may be used selectively but the relevant standards should be met fully.

This SMI should be used in conjunction with other SMIs.

Introduction

There are significant risks associated with the introduction of new or modified diagnostic tests, either through development in-house, or by adoption of methods developed elsewhere (whether available commercially or not) and therefore there is a need to adopt a consistent approach to assay validation and verification.

In order to comply with section f1.2 of the Clinical Pathology Accreditation (CPA) standards and with section 8 of the joint code of practice for research, there must be a method of assessing that test methods are "fit for purpose". This document gives guidance on development and describes how a validation file is produced. A test method may be a commercial kit, an in-house assay or reagent or a set of reagents bought separately and used to prepare an in-house assay.

The term 'validation' is often used very loosely and can cover a variety of different processes. From the manufacturing industry: "Validation is a quality assurance process of establishing evidence that provides a high degree of assurance that a product, service, or system accomplishes its intended requirements. This often involves acceptance of fitness for purpose with end users and other product stakeholders". Validation is an evidence-based process that requires proper planning in order to ensure that newly developed assays comply with laboratory standard systems and meet clinical governance and risk management requirements. There are a number of stages in this process including planning and inception, assay development and optimisation, assay validation, roll out and verification (ie a quality control process that is used to evaluate whether the assay complies with its specification), and finally implementation. Method validation can be used to judge the quality, reliability and consistency of analytical results. It is therefore an integral part of any good analytical practice. Annex 15 to the European Union (EU) Guide to Good Manufacturing Practice[1] which deals with qualification and validation provides a useful context.

Analytical methods need to be validated or revalidated before their introduction into routine use; whenever the conditions under which an original validation was done change (eg use of an instrument with different characteristics or samples within a different carrier matrix); and whenever the method is changed or modified beyond the original specification. The changes to a protocol that may be considered significant and that require revalidation with adequate evidence for equivalent performance before implementation depend on the specific details of the test. Various situations are likely to arise in which it is appropriate to repeat only a subset of validation tasks. For example, if the extraction method is changed it may not be necessary to carry out specificity checks but the sensitivity will require reassessment.

A validation file should be produced for all existing and new test procedures. The file may refer to existing data recorded in workbooks, papers and reports. Modifications to existing assays (including commercially available assays) will require an update to the existing validation file or creation of a new file.

It is essential to provide documentary evidence that any assay is suitable for its intended purpose. This may involve experiments to determine its accuracy, specificity, sensitivity,

reliability and reproducibility. Validation may be extensive, for example in the case of a newly developed in-house assay, or narrow in scope, for example in the case of a commercial assay already in use which has had minor modifications.

For nucleic acid amplification based assays, this document complements guidance on performing validations available in UK Standards for Microbiology Investigation Q 1 - Commercial and In-House Diagnostic Tests: Evaluations and Validations.

1 Project Personnel

All project personnel must have appropriate knowledge, competency and experience. Records of their qualifications must be available.

A designated Project Manager, most appropriately an individual at Team Leader, Unit Head or Laboratory Director level, has overall responsibility for the completion of the project and responsibility for signing off the completed validation file and SOP.

A Project Leader is assigned with responsibility for the project work which will include the laboratory activity needed for development of the test and its validation, data analysis, compilation of the validation file, report writing, presentation of data to review meetings, writing and maintaining the SOP and training of other staff to carry-out the new SOP.

All project personnel should have clearly defined lines of accountability.

The Project Leader may also be the Project Manager.

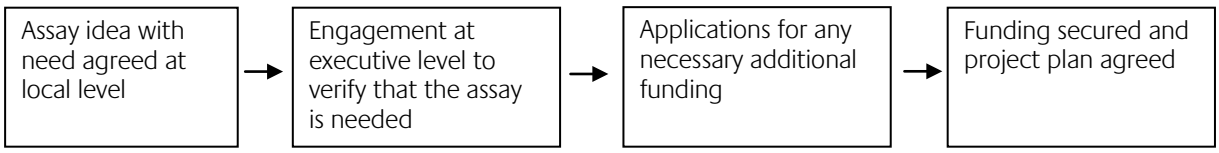
2 Equipment

All equipment used in the assay development and validation exercise must be maintained, serviced, calibrated and monitored as appropriate to ensure that it is suitable for use. This is essential to ensure that all conditions can be reproduced accurately during routine production of reagents and performance of the assay.

3 Planning and Inception: Establishment of a Project Review Team

The drivers for development of diagnostics are gaps in capability and capacity or opportunities for improved service presented by new knowledge and technology.

Figure 1: Initial stages of assay development



The first stages in the development process (Figure 1) are to secure high level management agreement for the work to proceed and approval of any necessary funding.

The aim of the planning phase is to produce a clear, agreed project plan. A suitable panel ('review team') should be established to review all aspects of the plan. This panel will also review the progress of development work, the validation plan, the validation study evidence that the test is fit for purpose and plans for post-deployment monitoring of test performance.

An example of a suitable team would be a Clinical Scientist, two clinical representatives (Medical Virologists or Microbiologists), a Biomedical Scientist and the local Quality representative. It is highly desirable that at least one of the clinical/medical members of the team should be an end user of the assay as this will be useful in providing input to design and validation parameters to ensure clinical utility. It is recommended that consideration is given to the inclusion of a statistician in the review team if none of the other members has adequate statistical training. Alternatively, a statistician should be consulted to advise on the validation study design. The project manager or project leader but not both, unless both roles are held by a single person, may be members of the review team.

The review team have the following responsibilities:

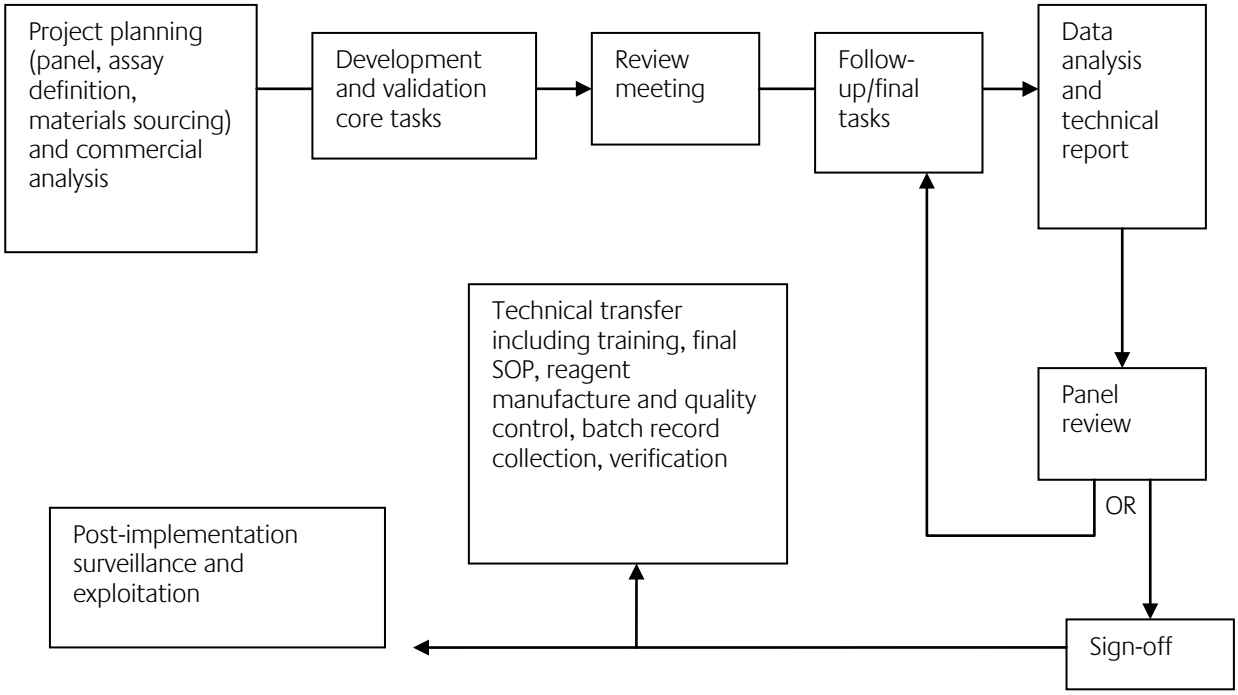
- To ensure that the purpose and objectives of the new assay development are well defined. Specifically how it is supposed to improve or fill in gaps in the current repertoire. This includes identification of the diagnostic needs, users, stakeholders and currently available alternatives.
- To ensure that a risk register is compiled. The risk register should include risks associated with project success/failure and with implementation of the assay (eg users might use inappropriate specimen types or misinterpret the results). Design actions to be taken in mitigation of the specified risks should be specified.
- To ensure that laboratory safety issues associated with inherent features of the developed test and various stages of the development and validation process have been assessed.
- Analysis of the business case including any procurement issues, the cost of performing the new assay, the price to be charged for the test taking into consideration any alternative assays that may be available. Consideration should also be given to any potential financial risks, for example, existing patents that may be associated with the proposed development.
- Performance of a commercial analysis including the opportunities for intellectual property protection and ownership, and the potential for commercialisation. If the assay or reagents are to be supplied to other organisations the cost of fulfilling regulatory requirements (eg IVD / CE marking) should be included in the analysis.
- Ensure that the engagement of collaborative partners to provide expertise or share costs has been considered.
- Ensure that HR aspects including training issues are addressed.
- Ensure that the means for efficient project management have been established including the nomination of advisors and reviewers as necessary.
- Approval and sign off (for example of suitable form see Appendix 7) of the detailed project plan.
- Approval and sign off (for example of suitable form see Appendix 7) of the validation plan.
- Review of the validation study data and recommendation on whether the assay is suitable for deployment.
- Review of deployment and post deployment plans.
- Ensure that the project dossier is available to the ADRG and is maintained.

4 Assay Development Plan

The objectives of assay development are to ensure the assay accurately and reliably measures only the analyte of interest in real (eg patient) samples, to ensure uniformity of assays and reagents over time and maintain full traceability of the results. Project objectives must be well defined eg development of a robust, convenient IVD assay for the mecA gene of *Staphylococcus aureus* in extracts from blood, urine, tissue and broth cultures, capable of being performed by clinical laboratory staff with basic molecular biology training and producing results by mid-to-late afternoon on day of sample extraction.

The project plan is developed and implemented as illustrated in Figure 2. The review panel and project lead should conduct a series of planning meetings, perform literature searches, appraise options and where possible consult with other centres that carry out the same or similar assays. The panel and project lead agree on the Assay Definition, which should include a name for the test and version control.

Figure 2: Assay development and validation plan



The flow chart illustrates the process. When the project plan has been agreed, the project lead will be responsible for performance of the laboratory tasks. A review panel meeting should then be held with the project lead presenting the data. Any follow up work and/or analysis, if necessary will be agreed, following which the project lead will generate the technical report. The technical report is then circulated to the review panel, who will then either sign-off the assay as ready for roll-out to routine use (technical transfer) or request further work is carried out before the assay can be signed off as suitable. Training of routine diagnostic staff is then undertaken.

The development plan should include a description of the technological details of the assay including information on the platforms and chemistry to be used. The reagents to be employed and the assay protocol should be described. Where practical the plan should minimise the number of separate reagent additions by combining reagents. For PCR an ideal is to use only two reagent mixes that can be stored ready for use. Whenever possible kits of finished reagents assembled to run the assay in a diagnostic setting should be formatted so

that it is cost effective to perform the test with batches of samples of the size generally encountered.

Where possible the assay protocol should be designed to minimise the processing time and technical complexity of the test. In general, the total assay time should be less than a working day so that results can be delivered on the day of sample receipt. The development plan should aim to minimise the number of operator intervention steps and the number of different volumes for additions or reconstitutions involved in the test. Ideally, volumes should be measured in round numbers (ie 5µL rather than 4.5µL) and be as large as practicable to facilitate operator or instrument precision.

The development plan must specify the requirements for success in detail including the desired analytical specificity, analytical sensitivity, diagnostic sensitivity, diagnostic specificity, clinical sensitivity, clinical specificity and reproducibility. It is important to consider the number of samples required for test validation. Table I gives the 95% confidence interval for a range of test sensitivities together with the required number of samples to be tested taken from Nature Reviews Microbiology[2].

Table I: Relationship between sample size and 95% confidence interval

Number of infected (non-infected) subjects required*	Estimated test sensitivity (or specificity)#					
	50%	60%	70%	80%	90%	95%
50	13.9%	13.6%	12.7%	11.1%	8.3%	-
100	9.8%	9.6%	9.0%	7.8%	5.9%	4.3%
150	8.0%	7.8%	7.3%	6.4%	4.8%	3.5%
200	6.9%	6.8%	6.4%	5.5%	4.2%	3.0%
500	4.4%	4.3%	4.0%	3.5%	2.6%	1.9%
1,000	3.1%	3.0%	2.8%	2.5%	1.9%	1.4%

*As defined by the reference standard test.

#95% confidence interval around the estimated sensitivity (+/- value in table)

The sample types to be evaluated must be specified together with the essential and desirable sample volumes. Sample preparation methods must be specified and for validation purposes form an integral part of the diagnostic test.

Ethical approval should be sought for the use of human clinical material. Guidance on research ethics and the Integrated Research Application System (IRAS) is available through the Department of Health website.

The controls to be used must be specified.

The planned shelf-life of the reagents to be used or the kit should be specified. Procedures for labelling for reagents mixed in-house should be specified. Labelling must include batch numbers and space for the insertion of an 'in use' date.

5 Core Tasks for Assay Development

Choice of oligonucleotide sequences and testing with control material

If new primers and probe(s) are required, a comprehensive survey of literature and sequence databases should be done to identify suitable sequences. Design of primers and probes should be conducted using appropriate parameters (eg those required for Taqman primers and probes) and design software should be exploited wherever appropriate eg by use of the latest version of Applied Biosystems Primer Express software. Several parameters need to be ensured during oligonucleotide design, namely unique target sequence specificity (the National Centre

for Biotechnology Information (NCBI) database search engine "BLAST" should be used to check candidate sequences). For genera with large numbers of subtypes, initial amplification should consider pooled consensus primers for subgroups as an alternative to degenerate primers to ensure optimum detection of all or as many as types possible for the subgroup.

In many circumstances it will be helpful in the process of primer and probe design to create an up to date alignment (eg by use of ClustalW or Bioedit) of a comprehensive collection of target sequences (ie the genes, open reading frames (ORFs) etc.) of different strains, subtypes or other relevant taxa. Sequences may be accessed via the comprehensive set of tools available on the NCBI website. The alignments should be used to check the conservative nature of the prototype oligonucleotides, ie if there are any nucleotide polymorphisms between significant strains (ie strains likely to be circulating in the sampled population for which the assay will be used), such sites should be avoided if possible and new candidate oligonucleotides chosen. If there is no choice, some polymorphisms may be acceptable if they do not significantly affect the hybridisation of the oligonucleotide in question (eg where guanine adenine base pairs are made) or if they can be overcome by use of non-canonical bases (eg inosine which allows indiscriminate hybridisation).

If the primers and probe sequences are already known, eg if already in a published paper, or they have been obtained from another molecular diagnostics centre, they should still be checked against databases and alignments as 'typographical mutations' (ie errors) may have occurred in critical sequence areas and recent studies may have revealed the existence of sequence variants.

Oligonucleotide sequences and the appropriate platform compatible fluorophors and quenchers must be chosen. The optimum concentrations of the oligonucleotides and the master mix components (ie the mixture containing Taq DNA polymerase, other enzymes if necessary, dNTPs, buffer etc.) must then be selected. This is done by testing different primer and probe concentrations against a dilution series of pure target nucleic acid (which extends below the nominal detection limit). The concentrations of oligonucleotides which allow optimal detection limits on a reproducible basis should be chosen. In general, it is helpful to use low concentrations of oligonucleotides since these minimise the occurrence of cross reactions and to reduce costs. Sufficient evidence must be collected that the primers amplify the expected target sequence.

Estimation of the PCR efficiency should be carried out by performing a real-time PCR amplification of a ten fold dilution series of control material. It may be suitable to carry this out during the assessment of the analytical sensitivity. The PCR efficiency is useful in assessing the design of the primers and probe and it is essential to identify minimum levels of PCR efficiency in developing quantitative PCR assays.

To calculate the PCR efficiency the Cq (quantification cycle ie generally the cycle at which the signal is determined to have become significant because, for example, it crosses a fluorescence threshold) values of each dilution are plotted against the known quantity or dilution factor. This can be achieved using the real-time instrument software. The gradient of the curve is obtained and for a ten fold dilution series would be -3.3 for a 100% efficient PCR assay. Efficiency is calculated from the equation E = 10[-1/slope] -1. Titration curve gradients of -3.1 and -3.6 giving reaction efficiencies between 90 and 110% are typically acceptable.

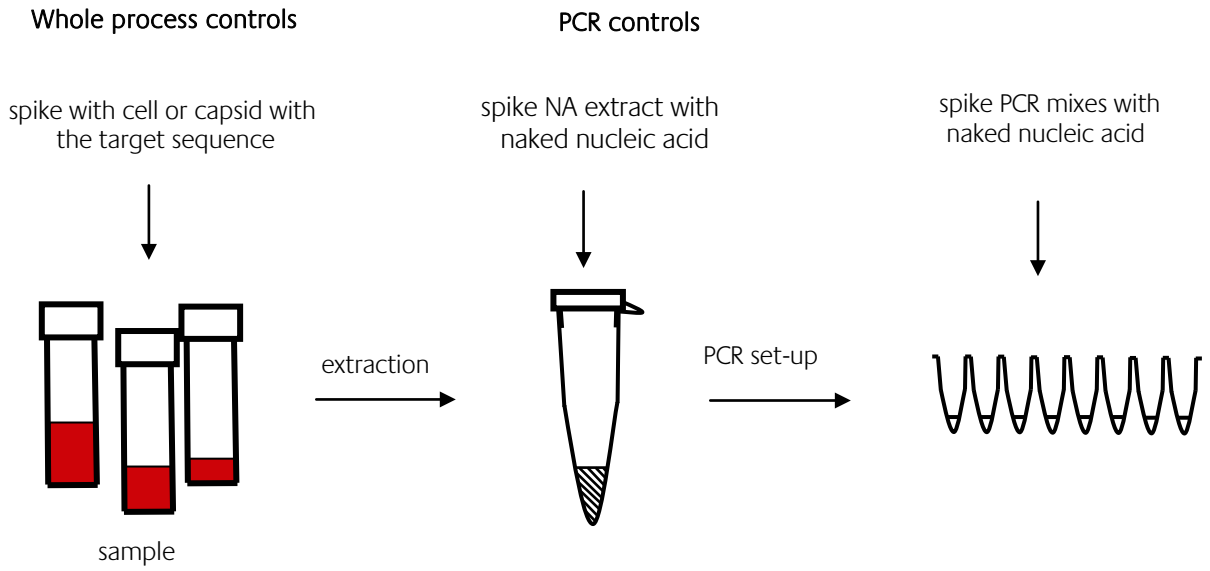
The following websites can be used to calculate efficiency once the slope of the curve has been determined.

- http://www.finnzymes.com/java_applets/qpcr_efficiency.html
- http://efficiency.gene-quantification.info/

Selection of PCR internal controls

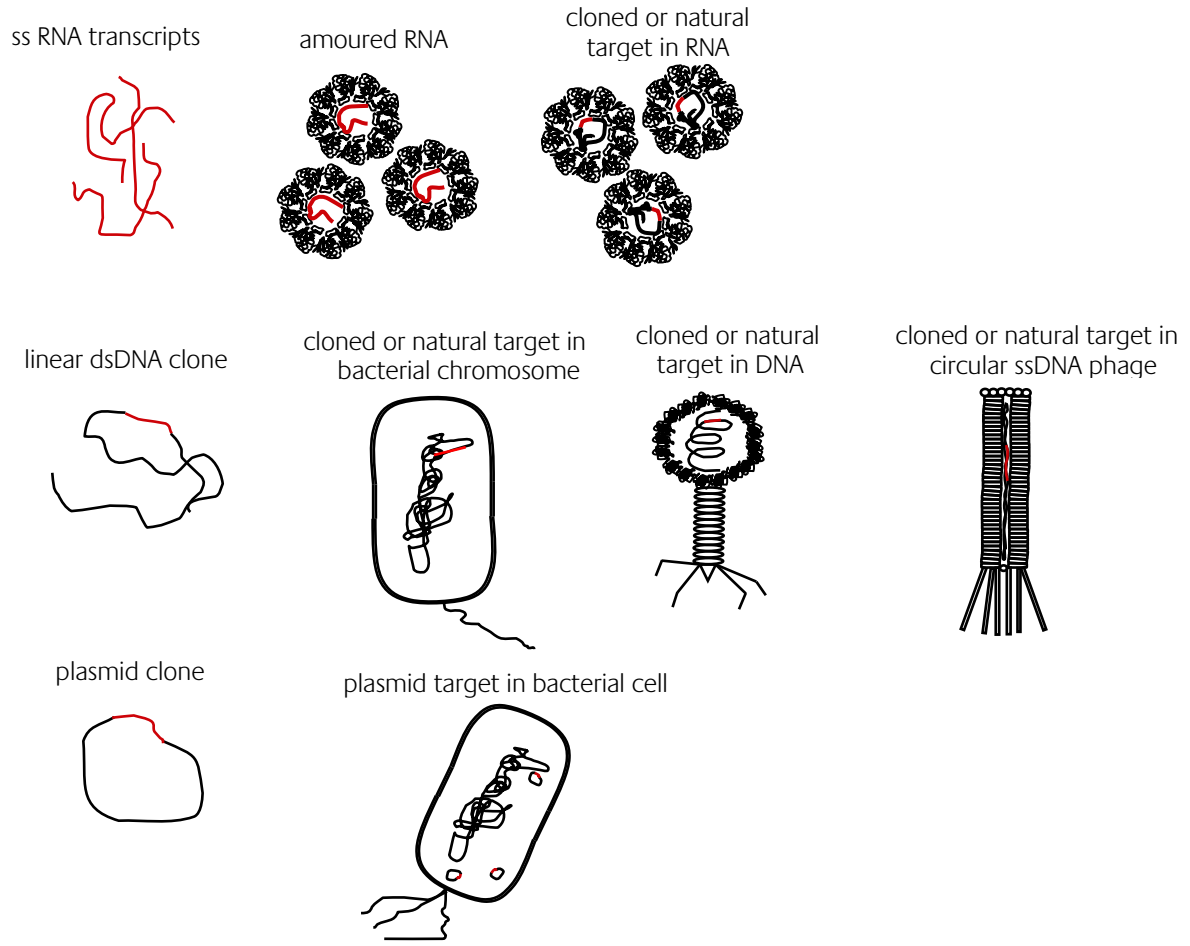
False negatives in diagnostic PCR may arise due to problems encountered at various stages of the diagnostic cycle. Although some of the problems are generic, also relating to other testing methods, some are specific to the design and use of diagnostic PCR. The problems of diagnostic PCR that are most intractable in terms of false negative results are extraction failure and reaction inhibition. The probability of false-negative results can be decreased by good practice and by following relevant laboratory guidelines. The use of appropriate controls aids in identifying false-negative results and greatly reduces the risk of incorrect diagnosis and reporting. Consequently, it is best practice to include a whole of process control (ie one capable of detecting both extraction failure and PCR inhibition). Figure 3 illustrates the difference between whole process and PCR controls.

Figure 3: Real-time PCR controls



It is essential to use an appropriate RNA or DNA internal control depending on target amplified. Whole process controls must be chosen to provide an appropriate challenge to the relevant extraction method ie optimally the nucleic acid target should be enclosed in a viral coat or bacterial cell wall. Figure 4 shows a selection of control materials. The internal control should be designed so that any adverse effect on the sensitivity of the assay is minimised. It should also be demonstrated that the internal control and target PCRs are similarly affected by the presence of potential inhibitory substances. This can be achieved by spiking the PCR reaction with potential inhibitors such as haemoglobin or reagents such as ethanol or phenol that may be carried over from the extraction process. The shift in crossing threshold values can be used to assess the relative impact of inhibition on the amplification of both the internal control and target organism.

Figure 4: Nucleic acid control targets for real-time PCR



Target sequence is shown in red and carrier nucleic acid in black.

Further advice on the design and application of internal controls is in Appendix 1 of this document.

External controls (ie where reference materials are tested in a parallel PCR reaction containing the sample) are less satisfactory although they are simple to design and implement. The risk that false negatives remain undetected, though greatly reduced, remains. However, for quantitative PCR it is advantageous that the control reaction cannot interfere with test accuracy.

Evaluation of DNA/RNA extraction protocol

The choice of extraction protocol will depend on the sample material, the numbers of samples to be extracted, the frequency of testing (eg one run every day or one run per week) and the cost. Other than primer and probe design, extraction is the most critical aspect of a molecular diagnostic assay. If an automated platform with a bespoke nucleic acids extraction kit is to be used, there may be a manual pre-platform processing step to consider, especially with recalcitrant sample materials such as faeces, sputum or human tissue. Sensible options need to be assessed. Elaborate protocols requiring extensive manipulations are unlikely to transfer robustly to the routine diagnostic setting.

Once candidate protocols have been identified, efficiency should be tested by spiking dilutions (extending below nominal detection limit) of whole cell or virus control into aliquots of sample material and proceeding with the extraction. As well as identifying the version of the protocol

which allows the best detection limit, attention should also be paid to any effects on the level of signal of the real-time PCR results. Partial inhibition may affect detection limits.

Optimisation of combined extraction and amplification assay

Results from the tasks outlined above should be analysed before deciding upon the need for optimisation of the assay. Sample volume, oligonucleotide concentrations and other parameters may need considering to optimise the assay. Dilutions of whole cell or virus control into aliquots of sample material should be used throughout optimisation to allow determination of performance improvement. A troubleshooting guide is appended to the guidance (Appendix 9).

6 Validation Technique

Using the list in Appendix 4 (includes a checklist based on the STARD Initiative[3]) and external literature including MIQE[4] the Project Leader must prepare a validation plan, considering the following as appropriate:

- Define the purpose and objectives of the investigation. For example, the assessment may be designed to validate the performance of a new assay or may aim to demonstrate that a significantly modified assay/variant protocol gives results within the tolerance of the original.
- Identify any training requirements to ensure everyone involved in the validation has suitable levels of competency. Ensure training records are up to date for procedures being carried out.
- Identify any risk assessments and COSHH assessments which need to be reviewed or written.
- Identify standard or reference materials where available to allow the assay to be standardised, facilitate assay comparison, permit stability of the test being validated to be determined over time and for use as controls.
- Identify any kit or reagent to be used for comparison of the assay undergoing validation, using the “gold standard” currently recommended for the subject in question.
- Design an analytical validation study to test the sensitivity and specificity of the assay using purified DNA. This will include DNA from wide range of strains/variants of your target organism and a range of unrelated strains or species that could be present in a sample, but that should not give a false positive result.
- Design a clinical validation study including the clinical context (eg surveillance, screening, clinical diagnosis), choice of comparators and ethical considerations. Choose the study population including type of patients, case definitions, inclusion/exclusion criteria and study settings.
- Identify the types (ie specimen, method of sampling, transport and processing) and numbers of samples to be tested. Consider the need to include known positives, known negatives, low and high positives and samples which are known or likely to be problematic (eg containing inhibitors or possibly cross-reactive markers) or representative of a particular population.
- Consider statistical techniques to select an appropriate sample size and avoid bias. It is essential to consider statistical requirements during project design to ensure that

results are statistically significant. Numbers required to ensure statistically significant results cannot be anticipated in this guidance document and will vary depending on a number of factors. Guidance is given in Appendix 2.

- The validation design should avoid discrepant analysis bias. Some samples may give discordant results with the new test compared to a gold standard. If only these samples are retested bias is introduced as there is a probability that the second analysis will give concordant results for some of these samples. To avoid bias, at least the same number of concordant as discrepant samples should be re-tested.
- Consider documenting an assessment of assay usability including method practicability, user feedback, barriers to implementation.
- Plan to review the validation study in a timely manner, minute meetings detailing attendees, consensus to changes to the study plan and all other agreed actions.
- Ensure that all SOPs related to the modified or new kits or reagents are current. It may be necessary to maintain new or revised SOPs as working drafts while their contents are being validated, ensuring COSHH and risk assessments are up to date. SOPs should be authorised as fully controlled documents when the validation study has been completed.
- In some instances assays may be designed to diagnose a disease for which relevant clinical material is hard to obtain or rare (eg the viral haemorrhagic fevers). In these cases use of the assays may be justifiable due to their potential diagnostic value even though the full validation criteria described in this document are not met. Deficiencies in assay validation should be documented in the development dossier together with justification for each instance where the data are inadequate. In such cases a plan for completion of satisfactory validation should be outlined. For example, it may be possible to obtain samples post-implementation or it may be reasonable to place greater reliance on the use of carefully designed simulated specimens.

7 Performing the Validation

Many of the below tasks can be carried out in concert or as part of the same assay runs with appropriate planning.

General principles

Accuracy can be expressed through sensitivity and specificity, positive and negative predictive values, or positive and negative diagnostic likelihood ratios. Determination of accuracy requires that the true value (ie as determined by a Gold standard where one exists) of the sample is known as indicated in Table II which defines these terms.

Table II Definitions

		Reference Test Results (Gold standard)	
		+ve	-ve
New test results	+ve	TP	FP
	-ve	FN	TN

TP=number of true positive specimens

FP=number of false positive specimens

FN=number of false negative specimens

TN=number of true negative specimens

Sensitivity=TP/(TP+FN)

Specificity=TN/(TN+FP)

Positive predictive value=TP/(TP+FP)

Negative predictive value=TN/(FN+TN)

Positive diagnostic likelihood ratio=sensitivity/(1-specificity)

Negative diagnostic likelihood ratio= (1-sensitivity)/specificity

Where possible, samples used should include high, medium and low positives (ie reactives) as well as negatives. It is not likely to be a fair test of sensitivity if only high positives are used. Appendix 2 gives information on the acceptable construction of validation experiments. Interpretation of the acceptability of an assay based on its validation data will necessarily depend upon the performance relative to alternative tests and upon the diagnostic situation, for example, a very high NPV will be required for a screening test.

Determination of analytical sensitivity (detection limits)

Analytical sensitivity of an assay is that assay's ability to detect a low concentration of a given substance in a biological sample. This type of sensitivity is expressed as a concentration (for example, in mg/mL or in gene copies/50 million cells). A lower detectable concentration shows a greater analytical sensitivity.

Analytical sensitivity is also known as "limit of detection" or "minimal detectable concentration" which is the lowest quantity of a substance that can be distinguished from the absence of that substance (ie a blank value).

When reference controls are available they should be used directly or indirectly via calibration of in house controls. The indirect route is acceptable since reference control material derived from a true biological source is often a precious resource.

For validation purposes the analytical sensitivity must be determined by the final version of the assay as developed and optimised. The detection limit should be expressed in acceptable units (copies, cfu, pfu, genome equivalents etc.) per mL or per g of sample material as appropriate. When a World Health Organisation (WHO) International Standard has been established, the units should be calibrated against this reagent. The detection limit is as achieved over at least four different assay runs with three replicates per assay. The detection limit may be different in different sample types. Conventionally, the limit of detection (LoD) is reported as the estimate of the detection limit that can be achieved with 95% confidence. This determination requires

Probit analysis involving testing of replicate samples around the end point of the assay, and processing the detection rate at each level through statistical analysis software[5].

Analytical specificity

Analytical specificity is the ability of an assay to exclusively identify a target substance or organism rather than similar but different substances (eg HIV-1 rather than HIV-2) in a sample or specimen.

This aspect of assay should be assessed as soon as possible after the prototype amplification assay has been shown to work. For the validation file, nucleic acid preparations of targets likely to be encountered in the samples for this assay should be tested. This should include as wide a variety of differing strains containing the target analyte (ie positives) as can feasibly be obtained, as well as negative target material (other organisms and nucleic acid likely to be encountered in sample material which should register as negative or not detectable by the assay).

Diagnostic sensitivity

The diagnostic sensitivity of a test is the assay's ability to detect persons with the condition of interest in a population, expressed as a proportion or percentage as defined above (Table II).

Diagnostic sensitivity may depend more on the ability to obtain the target substance in a processed sample from a person who has the condition than with the ability to detect very low concentrations of a substance.

Diagnostic specificity

Diagnostic specificity is the ability of an assay to correctly identify a person who does not have the disease in question. The diagnostic specificity of a test is the probability that a test will produce true negative results when used on a non-infected population.

Quantitative assays

Extra criteria are required to assess the performance of quantitative assays. These include measurements based on the standard curve formed using dilutions of control material. These measurements include the slope of the curve, the linear range and the X and Y intercepts. The efficiency of the PCR should be determined as described in the 'core tasks for assay development' section. The coefficient of variation across the linear range should be determined using quantitative values or standard deviation if using Cq values. This is important to establish clinical relevant variation and to establish lower level quantitative cut-off levels for an assay. Where they are available, international standards should be used for the calibration of quantitative assays. There are also commercially available quantified standards that may be appropriate to use (eg whole virus quantified controls from Accrometrix), however, the values assigned will not have the same standing as International Standards. Alternatively, in house quantified material can be generated synthetically (eg in plasmids) or from purified preparations of the organism. These in house standards can be quantified by appropriate methods such as measurement of A260 or fluorimetrically with dyes. The quantitative range should cover spread of results normally expected from clinical samples. Results should not be extrapolated beyond the established linear dynamic range of the assay. The standard curve should contain four or five points and the upper and lower values of the standards should be within 1 log of the reported quantitative range. Positive results obtained above or below the Limit of Quantitation (LoQ) should be reported as "Positive, greater than xx copies/mL" or "Positive, less than xx copies/mL" respectively.

Appropriate test controls should be established that may be used to monitor the performance of the quantitative assay when in use. These should be at a level that should be expected to be reliably positive but not at a concentration significantly higher than a typical clinical sample

in a typical real-time PCR reaction (eg as an indication the control should give a Cq value of 30). Test controls allow assay performance to be monitored over time and assessed by Westgard rules (see Appendix 8).

Linearity

The linearity of an analytical procedure is its ability (within a given range) to obtain test results which are directly proportional to the concentration of analyte in the sample. A linear relationship should be evaluated across the range of the analytical procedure. It is demonstrated by testing a series of dilutions of a known sample.

Linearity should be evaluated by visual inspection of a plot of signals as a function of analyte concentration or content. If there is a linear relationship, test results should be evaluated by appropriate statistical methods, for example, by calculation of a regression line by the method of least squares. Guidance from a statistician is advised if not familiar with the mathematical concepts associated with this process.

For the establishment of linearity, a minimum of four concentrations is recommended.

Measurement range

The range of an analytical procedure is the interval between the upper and lower limits (eg concentration, number of organisms or number of DNA copies) in the sample for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and linearity.

The specified range is normally derived from linearity studies and depends on the intended application of the procedure. It is established by confirming that the analytical procedure provides an acceptable degree of linearity, accuracy and precision when applied to samples containing amounts of analyte within or at the extremes of the specified range of the analytical procedure.

Precision

Precision can be determined by repeat testing of any sample. The precision of an analytical procedure expresses the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Note: precision is not the same as accuracy since an assay may be precise but inaccurate.

Precision may be considered at three levels: repeatability, intermediate precision and reproducibility. 'Repeatability' is the agreement between replicates within and between assay runs by the same operator over a short period of time.

'Intermediate precision' measures variation within a laboratory to include, for example, tests performed on different days, by different analysts and using different equipment. 'Reproducibility' includes the agreement between replicate tests performed in different laboratories.

Precision should be investigated using homogeneous, authentic samples, ie, behaving as much like real samples as possible. However, if it is not possible to obtain a homogeneous sample it may be investigated using artificially prepared samples or a sample solution.

The precision of an analytical procedure is usually expressed as the variance, standard deviation or coefficient of variation of a series of measurements.

Known positive samples should give the same acceptable range of results if assayed 4x or more in one run and over at least four different runs on different days. With sensible planning the task of collecting these data can be achieved when performing analytical sensitivity

testing. This level of validation does not include an assessment of reproducibility which is the closeness of agreement between independent results obtained with the same method on identical test material but under different conditions (different operators, different apparatus, different laboratories and/or after different intervals of time). The measure of reproducibility is the standard deviation qualified with the term 'reproducibility' as reproducibility standard deviation. Since there is less control of possible variables when determining reproducibility compared to repeatability, it should be expected that the reproducibility standard deviation will be higher than repeatability standard deviation. A comparison of precision between platforms should be considered, especially if more than one platform (extraction and/or amplification) is an option eg for contingency purposes such as during machine breakdown or surge capacity.

Reproducibility studies should be included in the validation file if it is proposed that the diagnostic method be 'rolled-out' to other laboratories.

Analytical accuracy

Known negative and positive samples (ie with known reference or accepted reference test values) should give same acceptable range of results if extracted and assayed 4 or more times in one run (with at least three replicates to allow intra-assay reproducibility to be measured too) and over at least four different runs on different days. With sensible planning this task can be achieved when performing analytical sensitivity testing.

Clinical accuracy

Validation requires comparison of the assay to an appropriate "gold standard" using sequential clinical samples. The number of samples tested will vary depending on the availability of suitable clinical material. They should, wherever possible, include a wide range of concentrations of positive samples as well as negative samples. A statistician should be consulted to discuss possible statistical analysis that may be carried out.

Robustness

The robustness of the assay may be evaluated by deliberately deviating aspects of the protocol that are perceived as being sensitive to the outcome, for example, by simulating pipetting errors affecting critical components such as enzyme and Mg++ or by extending the storage of PCR reaction mix at room temperature or at 4^{o}C before amplification. It may be important to include such data in the validation file if it is considered that it is likely to be useful for troubleshooting purposes.

Reference intervals for the quantities being determined and description of the appropriate reference population

The reference intervals, also referred to as the reference range, normal range or reference limits, are the upper and lower levels of analyte that you would expect to see in a normal population. Any values above or below that range are outside normal limits. Reference intervals are not relevant where the assay is designed to detect the presence of a target (ie a qualitative test) but is important when the purpose of the assay is to determine whether the patient or the sample is in or outside the normal range. Examples might include levels of antibody in a serology assay or Human Immunodeficiency Virus (HIV) viral load. Another example is an antibiotic sensitivity test, where it is important to know the concentration of antibiotic which will inhibit sensitive organisms, above which the organism would be deemed resistant.

Information needed for the control of known relevant interferences

This requirement needs an understanding of the biological sample that will be used when performing the test. What is present or might be present in the sample which could interfere

with the test result? It may be possible to control known interferences by pre-treatment of samples. It is essential to include appropriate positive and negative controls to ensure that the presence of possible interferences is detected. Presence of interferences may be identified during assay development and/or validation when it is used to test samples of known activity.

Limitations of the method

It is important to understand any limitations associated with the assay. The limitations may be based on general experience or have been identified during the validation. Examples might be lack of sensitivity compared to another method, or when certain specimen types are used; possible false results in contaminated specimens.

Clinical validation and implementation in diagnostic pathways

Analytical validation ensures that the test chemistry reliably detects the target. Diagnostic validation ensures that the test detects the target in clinical specimens and is designed to show that it is fit for use on clinical samples. The next step is then to integrate the test into the clinical routine as part of a pathway. The clinical pathway is initiated by the set of clinical indications leading to the patient being tested and includes the expected range of results and details of patient management depending on these results (ie to establish workflows so that the test is performed at an appropriate point in the diagnostic cycle, that performance is triggered by specific clinical questions and that the results are interpreted in a way that feeds back into patient management).

Clinical validation can be carried out either retrospectively or prospectively. It may be useful to establish a prospective evaluation once the assay is implemented to add data to the file. The validation file should include information on the clinical situations that would trigger use of the test and the appropriate clinical interpretation of the results in relation to identification of the pathogen. Details of appropriate reporting comments in relation to the results in particular clinical scenarios should be provided. High and low positives in various samples and in different clinical contexts may have different implications (eg Cytomegalovirus (CMV) may be present in a specimen but may not be causing disease). Thus the clinical validation will inform the decision on whether the finding is clinically relevant.

The US Patient-Centered Outcomes Research Institute (PCORI) Methodology Committee have published standards for PCOR[6] and standards for the 'Design, conduct and evaluation of diagnostic testing'[7]. These standards relate to clinical validation of assays and their appropriate integration into clinical pathways. They are a useful resource and should be taken into account when planning clinical validation studies.

Shelf-life of materials

All commercial diagnostic reagents have expiry dates on their labels. These have been determined by shelf life studies. It is important to know how long a reagent remains active. Activity may fall off gradually and so still give an acceptable result with Quality Control (QC) material but not be performing at its optimum.

The principle of shelf-life studies is straightforward. First determine the proposed storage temperature. It may be necessary to determine shelf-life at a range of temperatures, eg -20 ^{o}C+/-2 ^{o}C, +4 ^{o}C+/-2 ^{o}C and room temperature (ie 20 ^{o}C+/-5 ^{o}C). Then store samples of reagents at the required temperature(s) for a period of time, regularly testing samples using a known QC sample, to determine how long the reagents remain at their optimum. BS EN 13640[8] requires shelf-life testing to be performed on samples from three batches. There are two approaches that have proven acceptable. These are detailed in Appendix 3 which also shows an analysis of the advantages and disadvantages of the two methods of determining shelf-life

Shelf-life stress studies

Shelf-life studies determine the stability of the reagents under controlled conditions (eg 2-8 ^{o}C). However, during their life reagents may be subjected to other conditions, eg, ambient temperature during delivery to another location and/or storage on the bench; freeze-thawing once or more frequently during the reagents' life time. It is important to consider possible conditions that the reagents could be subjected to during their life time and attempt to recreate them in the laboratory to see if the performance of the reagents is affected. It is only necessary to perform stress studies on one batch of a reagent.

If the reagents could be freeze-thawed several times during their life then take different aliquots, subject them to a series of freeze thaw cycles and retest them. Samples can be tested after each freeze thaw or they can be collected and all tested at the same time to allow easy comparison and control.

If reagents are to be posted to another laboratory, then they could be left in a hot warehouse overnight during a heat wave. Therefore, you may wish to consider retesting a sample after subjecting it to 25 ^{o}C for 24hrs or whatever temperature and time is considered appropriate.

If, following stress studies, it is found that reagents are not stable following several freeze-thaw cycles then instructions for use must clearly state this and require that freeze-thawing is avoided. For reagents which need to be stored frozen, instructions may need to state that they should be discarded after use.

If stress studies show that reagents are not stable when stored outside 2-8 ^{o}C for example, then they must be delivered on ice and instructions must clearly state that they must be stored at the appropriate temperature when not in use.

A troubleshooting guide is appended (Appendix 9).

8 Risk Assessment

It is essential that a risk assessment is performed prior to using any reagent in a diagnostic test in order to minimise the risk to both users of the device and to patients[9]. The risk assessment should consider for example infection risk of materials where they contain biological substances; all combinations they might be used in, eg, different platforms, different ancillary reagents. Depending what risks are identified, further action may be required to ensure appropriate control measures are in place. This might include inactivating a reagent; reviewing stress studies; reviewing how reagents are stored eg, temperature or type of container used; providing more information in the instructions for use.

9 Data Analysis and Composition of Technical Report

The technical report and validation file should summarise assay validation, results and recommendations. This task also includes the setting up of the Westgard rules criteria using positive control results from runs conducted as per routine SOP, ie positive controls from the clinical accuracy runs should be sufficient.

For in-house assays, validation may be supported by the research and development carried out during the development of the procedures. Workbook records can be cross referenced if appropriate in the validation report.

The validation work should be completed in accordance with the design and results recorded in a designated file.

The results should be analysed as defined in the study design. Results are compared with expected values to determine whether or not the kit or reagent is suitable for use. This may include comparison with results using alternative reagents and for commercial kits sensitivity and specificity information provided by the manufacturer.

10 Review of Technical Report and Sign-Off

The validation data should be presented in a written report, so that the reviewers can make a decision as to whether or not the assay is fit for purpose (as agreed in the Assay Definition).

A conclusion should be documented based on the results analysis. This must include a formal declaration that the assay is suitable for use which should be signed by the Project Leader and the Project Manager.

The validation must be completed and signed off before any use of the kit or reagent for diagnostic or reference purposes. Any use prior to sign-off must be reported as research use only. Validation must be completed in a timely manner yet the need to meet deadlines must not compromise the thoroughness of the validation process.

11 Instructions for Use

All assays must have instructions for use. This may be the SOP that is in place or a product information sheet (PIS) supplied with the reagents when shipped to other laboratories. The IVDD includes very detailed requirements of what should be included in a PIS. Some flexibility on the IVDD requirements may be permissible for in house assays.

12 Labelling Requirements

The IVDD includes very detailed requirements of what should be included on reagent labels. Some flexibility may be permissible for in house assays. However, the minimum requirement is for the reagent name, batch number and expiry date to be clearly identifiable.

13 Production and Storage Records

Royal College of Pathologists (RCPath) guidelines state that batch records must be retained for a minimum of 10 years. Although the guidelines do not make it clear what batch records they refer to, batch records related to production of in house assays should be archived for that period of time. Batch records to be kept are details of the purchase and manufacture of every reagent, including reagent lot numbers, details of what equipment was used, how the reagent was made and all treatment undertaken. All reagents must be manufactured, stored and used within the laboratory's quality system and sufficient records taken to allow a full audit trail to be undertaken.

14 Validation of Assays Already in Use

Once satisfactory validation has been completed using a defined SOP it is essential that the assay is performed according to the SOP. Minor changes to the method that are immaterial to the performance specification of the assay should, as far as possible, be accommodated within the SOP provided evidence is available. For example, it may be better to specify the use of distilled, deionised or molecular grade water rather than a particular brand of molecular grade

water. Similarly, it may be reasonable to avoid specifying the particular model of PCR instrument to be used.

When changes to the protocol that fall outside of the SOP must be made, revalidation is required. Revalidation will vary from a single run showing that Cq values are unchanged for a range of reference samples (eg as might appropriate for a change to a buffer) to full revalidation (eg for use of an alternative sample preparation method or primer change). Appendix 6 is a guide to the level of revalidation that might be considered acceptable for a range of changes to the protocol.

Assays in use with validation records that do not comply with the requirements of this guidance document should be retrospectively provided with a compliant validation file wherever practical. Historical records should be reviewed and summarised to provide evidence that the assay is fit for purpose and a validation checklist should be used to cross-reference these documents. This also applies to non-commercial assays adopted from another laboratory.

Further validation work may not be required even if insufficient data are available to produce a compliant validation file. However, it may be necessary to implement appropriate data collection to provide supportive information.

A conclusion based on the information obtained from the historical data should be documented. This must include a formal declaration that the assay is suitable for use.

15 Assay 'Roll-Out'

The collection of diagnostic validation data will necessarily involve limited technical 'roll-out' within the developing laboratory. Full adoption of the new assay into clinical use will include consideration of the reagent supply chain and arrangements for quality assessment, including EQA wherever appropriate. If the test is likely to meet the needs of a range of end-users roll-out should be considered. A roll-out plan should include the timeline, equipment, personnel and risk management proposals. If the assay is for clinical diagnostic purposes, the requirements of the European Directive must be complied with and, if the assay is to be made available for use outside of the laboratory that designed it, CE marking affixed for kits or control materials.

Following roll-out an assay verification study and continuing quality assessment should be performed. This should be organised by the developing laboratory or a qualified third party (eg an EQA body). Verification should involve the testing of validation panels or be limited in scope to a few well characterised samples and standards. In some circumstances it may be desirable that validation materials are originated in laboratories other than that of the developer. Round-robins may be useful in ensuring continuing attention to quality.

16 Production of Reagents for Routine Use

Following development of a satisfactory assay it is essential that the procedures for production of reagents for routine use are clearly defined to ensure that the assay continues to be fit for purpose when further batches are made. An SOP(s) must be prepared which describes the production process in detail, including specific information about the reagents, equipment and conditions. Where appropriate, worksheets should be prepared to allow recording of all manufacturing information to allow a complete reconstruction of the manufacturing process for every batch.

The SOP(s) must also include details of how interim (where necessary) and final product(s) are quality controlled prior to batch release. A worksheet should be prepared to allow recording of quality control results and should include a final batch release sign off by a responsible person.

All staff involved in preparation of reagents must be suitably trained/have evidence of their competency documented.

17 Post-Implementation Surveillance and Verification

'Post-implementation' is equivalent to the term 'post-marketing' in the IVDD and other regulatory literature.

Once the assay has been validated and put into use, information on its performance should still be gathered. Details of any problems, eg, poor performance, false positives or negatives, should be documented. All problems should be investigated and appropriate corrective action implemented. Timely communication with other users of the assay or similar assays is very important, both to make them aware of any problems and as part of the root cause corrective action. If necessary, the risk assessment should be reviewed in light of the new information. All information must be included in the project dossier (ie that includes the assay development and validation file) which is not static but should be kept up to date with all current information. All project dossier information should be shared with laboratories performing the test.

The assay should be reviewed periodically to ensure continued improvement and to highlight performance compared to alternative assays, whether commercial or in-house. Periodic reviews should include review of additional nucleotide sequence data that may be available, and review of scientific literature for the identification of novel strains or variants. Dates for reviews must be included in the assays SOP.

18 Documentation

All documents relating to the assay development and performance monitoring must be filed in a retrievable and auditable manner. There should be a file where all related documents are either stored or cross-referenced.

Complete the worksheet (Appendix 5) for each assay and retain in the validation file. If key information is already documented, it is not necessary to transcribe it onto the form, cross-reference it. A list of appropriate documents is detailed in Appendix 5, Validation report summary for a kit or reagent.

The Project Leader and Project Manager must review the data, complete the validation checklist, and sign the validation section to authorise release of a new or modified kit or reagent or to assure that sufficient information has been provided to confirm that a kit or reagent already in use is fit for purpose.

A departmental report reference should be added to the header of the report form to ensure that all pages of the report form are identifiable. It is recommended that the report reference number is written NNN/YY, where the first three digits represent a unique number, incrementing by 1 each time, followed by a year code, eg, 001/06, 002/06 etc.

Appendix 1 – Design and Implementation of Internal Controls (ICs)

The problems of diagnostic PCR that are most intractable in terms of false negative results are extraction failure and reaction inhibition. Efficient extraction of target nucleic acid is dependent upon the nature of the micro-organism and extraction protocol used and is thus highly variable. The problem of reaction inhibition due to the presence of inhibitor in the specimen is unpredictable in that it may affect one or more sample within a batch even if no technical error has occurred. Extraction failure and reaction inhibition are far more common with clinical samples subject to primary diagnostic testing than with secondary samples.

The application of appropriate controls that facilitates the identification of false-negative results, especially those related to extraction failure and/or reaction inhibition is thus generally recommended and should always be included in PCRs that are intended for clinical diagnosis affecting patient management or significant public health decisions[5,10] and are implicitly or explicitly required in external (ISO, UKAS, CPA) and internal (UK SMI) standards.

There are various forms of PCR controls for detection of false-negative results, conferring different levels of assurance[11-14]. These are summarised in Table I. It must be emphasised that the scale is arbitrary and the overall degree of assurance also depends on which stages of the PCR are under control (extraction, reaction and inhibition).

Table I: Categories of PCR controls for detection of false-negative results

Method	General level of Assurance
Internal control using same amplimers (mimic controls)	Good
Internal control using different amplimers (non-mimic)	Good
Endogenous internal control (eg housekeeping gene)	Good
External control amplified in parallel paired reactions	Acceptable
Simple batch external control	Low

Different overall approaches are summarised in Table II. These approaches may be graded according to the perceived overall level of assurance they confer, based on whether they control for extraction and/or inhibition and the robustness of the used methods. The level of assurance is an arbitrary scale. In general, methods utilising robust controls (ie use of IC in multiplex assays) and addressing both extraction and inhibition were judged to confer a good level of assurance. Methods utilising less robust controls (ie external controls in combination with IC) for inhibition and extraction were judged to confer an acceptable level of assurance. Methods not utilising IC at all but only external controls were judged to confer low level of assurance. It should be noted that occasional assays were judged not to require controls for a false-negative result (typing assays having a base sequence as endpoint in which a "negative" necessarily implies a faulty assay and would be repeated).

Table II: Summary of control methods

Extraction control	Inhibition control	Degree of assurance
Control (phage coated RNA and/or DNA as appropriate) spiked into each sample	Extracted positive nucleic detected in same tube	Good
Human DNA indicative of good sample	Extract tested for intrinsic human template in the same tube	Good
Generic target (typing or gene variant assay)	Generic target detected in same tube	Good
Positive control cells extracted in separate tube	Positive target spiked into each sample detected in same tube	Good
No extraction control	Positive target spiked into each sample detected in same tube	Acceptable
Control DNA (eg plasmid) spiked into each sample	Extract tested for control template in a parallel tube	Acceptable
Human DNA indicative of good sample	Extract tested for intrinsic human template in a parallel tube	Acceptable
Extraction of nucleic acid tested by a parallel assay	Inhibition tested by a parallel assay	Low
Batch extraction control	Extract tested for spiked low copy template in a parallel tube	Low
No extraction control	Extract tested for spiked control template in a parallel tube	Low

	Internal controls		External controls

External controls

External controls use reference material in a parallel PCR reaction. In its simplest form, external control utilises one or few spiked vessels within a single run of a certain testing batch. Alternatively, external control can be achieved by running parallel pairs of reactions with one tube in each pair serving as a control. The use of external controls (parallel controls) has advantages in that the control PCR reaction cannot interfere with test accuracy and the control amplicon does not need to be distinguished from the test amplicon. For that reason external controls are very simple to design and implement.

However, two main disadvantages exist as compared to IC. First, the number of PCR reactions that must be set up is double. Second, and more importantly, since they are not ICs, the potential for false negatives due to sample inhibition, though greatly reduced, remains.

Internal controls

Unlike external controls which are run in a parallel reaction vessel, IC harness the possibility for multiplexing PCR reactions using a second target molecule that can be amplified with, but readily distinguished from, other products of the reaction in the same reaction vessel.

An ideal IC sequence should provide control for all procedures, reagents and samples included in the PCR test and therefore would be present in all correctly taken samples, would be

released from the same cellular compartment as the target sequence under the extraction conditions, would be non-competitively amplified using the same primers as the target sequence and also produce an amplicon of identical size and similar GC content, that can easily be distinguished from it. Although it is rarely possible to achieve this ideal, non-ideal IC are nevertheless valuable.

A simple IC for reaction inhibition might involve spiking the IC DNA into the PCR master mix at an appropriate concentration. IC targeting both extraction failure and reaction inhibition might involve spiking the IC sequence, or any form of target control nucleic acid that requires extraction, into the sample prior to the extraction phase. Spiking should involve a discrete quantity of exogenous nucleic acid into the sample at a known concentration.

Spiked IC can be engineered to be a molecular mimic of the target sequence. The mimic sequences generally have priming sites identical to those of the target and intra-primer sequences ('stuffer') differing only to the extent needed to allow straightforward discrimination from the target. The degree of difference required between authentic sequence and mimic depends upon the method used for analysis. With gel electrophoresis, the IC amplicon must be sufficiently different in size from the target amplicon to allow them to be reliably distinguished. However, when a real-time probe system (eg TaqMan) is available, even a single base difference between IC and target may be sufficient. The disadvantage of a molecular mimic IC is competitive binding to primers and feedback inhibition due to heterologous amplicon annealing (ie leading to suppression of the minor population). Where this is likely to be a significant problem the best alternative is to use an IC with similar characteristics to the target (ie length, guanine plus cytosine content, melting temperature) but that requires different primers and that is sufficiently dissimilar that no annealing of IC to target sequences can occur.

Delivery of the IC into the assay needs some consideration, especially when it must be added to a complex sample prior to extraction of the target nucleic acids. The addition of naked double stranded DNA that includes the IC sequence, in either linear, open circle or supercoiled form, to samples immediately prior to or post the addition of the 'lysis' buffer (ie any solution added to the sample to promote cell lysis or to strip proteins from nucleic acids or to prepare the nucleic acid for adsorption to a solid support) is not expected to be problematic and should result in non-preferential recovery of the IC sequence together with the target sequence. However, when it is necessary to add the IC as RNA (eg when the assay is designed to detect an RNA virus) it is usual to take additional steps to avoid ribonuclease action which may affect sequence integrity especially during longer term storage of the control material or to produce an RNA transcript that is resistant to RNase. One approach is to prepare the control RNA as 'armoured RNA' (ie RNA coated with bacteriophage MS2 coat protein a commercial product that may require a license for use in diagnostic systems) or to make use of products such as RNAlater™ which is a concentrated solution of ammonium sulphate.

A frequently used alternative that is appropriate for certain specimens is to use a host nucleic acid sequence as control target. In this case all appropriately taken samples will be positive for the control. For the diagnosis of human infectious disease an important consideration is that the sample is adequate and this may entail that human cells be present. Consequently, it may be advantageous (or at least acceptable) to use a human gene as the IC in assays designed to detect infectious agents. One constraint of this approach is that the quantity of IC present depends upon the sample quality, which may be highly variable. It is therefore important to ensure that when large quantities of IC are present amplification of the target is not overwhelmed, whilst maintaining an adequate signal from the IC when lower amounts are present in the sample. Generally, on the basis that human DNA is present in large quantities in properly taken samples of relevant tissues, the best solution to this problem is a compromise

in which the IC PCR is detuned (ie by reducing the concentration of one or both primers) to prevent the quantity of IC amplicon reaching the high levels that inhibit target amplification.

When a PCR assay is designed for detection of a particular type of a pathogen that is known to be present in a sample, suitable control may be inherent in the test. For example, in tests where an amplicon is sequenced the invariant bases, which are always present, provide a suitable IC. Alternatively, when the pathogen is known, conserved genes in its genome can be used for the purpose of IC in a reaction multiplexed with amplification of the test target sequence.

Risk assessment

A risk assessment of the use of controls in in-house PCR assays is summarised in table III. The risk score is derived from the arbitrary assessment of the impact and consequences of having a false-negative result and the probability for having such a result. The impact of a false-negative is considered highest for specimens affecting patient care, either through primary or confirmatory diagnostics and lowest for typing for surveillance purposes. The probability of a false-negative is considered highest for primary specimens owing to the varying target density, influence of sampling method and risk of inefficient extraction and reaction inhibition. The probability is lower for secondary specimens due to the high target density and predictable extraction and inhibition and lowest for typing assays which are applied on specimens known to be positive for the target micro-organism. The degree of assurance conferred by controls informs the mitigation required to manage the risk.

Table III: Risk assessment tool for PCR assays

PCR application	Impact of a false-negative*	Probability of a false-negative*	Risk score	Minimal risk mitigation needed via controls
Clinical diagnosis from primary specimen	3	3	9	Good assurance
Clinical diagnosis from secondary specimen	3	2	6	Good assurance
Confirmatory testing (reference) of primary specimen	3	3	9	Good assurance
Confirmatory testing (reference) of secondary specimen	3	2	6	Good assurance
Treatment critical typing from primary specimen	2	2	4	Acceptable assurance
Treatment critical typing from secondary specimen	2	1	2	Acceptable assurance
Typing for surveillance from primary specimen	1	1	1	Low assurance
Typing for surveillance from secondary specimen	1	1	1	Low assurance

* 1 – low; 2 – moderate; 3 - high

Conclusions

Complexity and diversity of specimens and targets for PCR detection in diagnostic assays means that a standardised approach is not always possible for the design and implementation of internal controls, but a hierarchy of assurance can be considered.

Real-time PCR diagnostic platforms, besides their other operational advantages, currently offer the most convenient means of implementing assays with IC. Internal controls are superior to external parallel controls.

The development and use of in house assays should clearly specify the target clinical specimen, and the intended use in a clinical care pathway. This information should be used to determine the assurance required from control strategy.

Validation data for in house assays must contain explicit discussion of control strategy and consideration of the impact of a false negative result.

Appendix 2 – Design Parameters for Validation Experiments

Parameter	Design
Analytical sensitivity	Four different assay runs with at least three replicates per dilution of the sample.
Diagnostic sensitivity	Testing of samples (that have been tested using the 'gold-standard' or appropriate alternative assay) from cases with the defined clinical profile(s). The minimum numbers of samples to be tested (to give a required level of reliability of the sensitivity measurement) will depend on the prevalence of disease and can be calculated from the minimum sensitivity levels and the 95% CI shown in table I.
Analytical specificity	Testing of DNA extracted from as many variants as possible of the target organism, genetically related organisms and organisms likely to be found in positive and negative cases (>100) with the defined clinical profile(s).
Diagnostic specificity	Testing of >50 samples that were positive and >50 samples that were negative using the gold-standard assay.
Efficiency (quantitative assays)	Test 10 fold dilutions of a positive sample or control in triplicate. Dilution range to give C_qs from <12 to >35 cycles.
Linearity (quantitative assays)	Test 10 fold dilutions of a positive sample or control in triplicate. Dilution range to give C_qs from <12 to >35 cycles.
Measurement range	Test 10 fold dilutions of a positive sample or control in triplicate. Range to extend from lowest practical dilution to ten fold beyond highest dilution giving a positive result.
Precision (quantitative assays)	Three samples (high, medium, and low positive) assayed at least four times or more in one run and over at least four different runs on different days.
Reproducibility (quantitative assays)	Three samples (high, medium, and low positive) assayed at least four times or more in one run and in at least four different runs on different days. These to be run in different laboratories or using different reagent batches or different instruments.
Analytical accuracy (quantitative assays)	Three analytical standards (high, medium, and low positive) assayed at least four times or more in one run and over at least four different runs on different days.
Clinical accuracy (quantitative assays)	Three clinical standards (high, medium, and low positive) assayed at least four times or more in one run and over at least four different runs on different days.
Reference intervals	Testing of >100 samples (that have been tested using the gold-standard assay) from cases with the defined clinical profile(s).
Clinical validation	Analysis of samples from cases with the defined clinical profile(s) with follow-up. This is on-going audit of assay performance.
Shelf-life	Samples from three batches stored at the designed storage temperature. Aliquots used to assay three samples (high, medium, and low positive) at least four times or more in one run and in at least two different runs on

	different days.

Appendix 3 – Methods for Determination of the Shelf-Life

Method 1

1. Take sufficient aliquots of sample for the entire study and store them at the required temperature(s).
2. At the agreed interval (eg, weekly or monthly) take the samples from storage and test them using the same control material throughout the study.
3. Continue the experiment until you have reached the agreed term or until the reagent has started to lose activity.
4. Plot a graph of results and assign a shelf-life based on the evidence obtained. This is usually the time when the reagent is first observed to lose activity minus a period to allow a margin of error.

Method 2

1. Take sufficient aliquots for the study of sample and store them at a very low temperature at which it is known they will not deteriorate (eg, -80°C).
2. At agreed intervals (eg, weekly or monthly) take some samples from storage at the low temperature and store them at the proposed storage temperature.
3. Continue to place samples at the agreed shelf-life study temperatures until the agreed length of the study period.
4. Remove all samples from storage and test simultaneously.
5. Plot a graph of results and assign a shelf-life based on the evidence obtained. This is usually the time when the reagent is first observed to lose activity minus a period to allow a margin of error.

Advantages and disadvantages of the two methods of determining shelf-life

	Advantages	Disadvantages
Method 1	If one test gives odd results it will not affect other results so could be seen as an outlier when later compared to all results in the series. Can keep testing samples until reagent starts to lose activity, assuming there are sufficient samples. If problems occur in testing at a later stage then have previous results to base evidence on.	Have to set up a test each time. Need to ensure that the control remains stable. Are only comparing each sample against the control rather than each other. May not be able to use same batches of ancillary reagents which could include a variable. If use same batches of ancillary reagents they may change over time, introducing another variable.
Method 2	Only have to perform assay once. Assay tests all samples at the same time so are comparing with a single control and against each other.	If something goes wrong with test then none of them is valid. Do not have an ongoing knowledge of how the reagent is performing so could retain samples after the reagents

	Only using one set of ancillary reagents, eg, buffers, enzymes, sera, so better control of process.	first brought out have expired or decided to test all samples while they still have a long shelf-life left.

Appendix 4 – Assay Development Process Checklist

Name of assay	
Specific targets (pathogens and genes) of the assay	
Project manager	
Project leader	
Staff performing the hands-on development work	

Section	Subsection	Contributor	Deadline	Completion
Planning & inception	Project drivers – gap analysis			
	Setting purpose and objective for new assay			
	Sign-off process			
	Setting up a project risk register			
	Health and Safety issues			
	Business analysis			
	Commercial analysis			
	Regulatory compliance			
	Human resources aspects			
	Project planning and initiation			
Assay development	Identification of targets			
	Setting benchmarks and performance measures			
	Options appraisal for technologies			
	Identification of gold standards			
	Identification of reference material and reagents			
	In silico design			
	Technical design			
	Optimisation			
	Design of post-amplification analysis			
	Development of controls			
	Assessment of analytical performance			
	Preparation of specific SOPs (MIQE compliant)			

Section	Subsection	Contributor	Deadline	Completion
Validation	Design of validation study			
	Assessment of analytical performance			
	Choice of study population			
	Choice of sample type			
	Sample size calculation			
	Performance of clinical validation			
	Data analysis and interpretation			
	Assessment of usability			
	Reagent stability studies			
	Further optimisation if required			
	Production of documentation			
Roll-out and verification	Review of needs and users			
	Planning of roll-out			
	Field verification studies			
	Provision of validation panels			
	Inter-laboratory comparison			
Implementation	Implementation in laboratory work-flows			
	Implementation in diagnostic pathways			
	Implementation in clinical pathways			
	Post-marketing assessment			
	Long-term data collection			
	Periodic QC/QA			
	Periodic proficiency testing			
	Production aspects of implementation			
	Periodic reassessment of fitness for purpose			

Appendix 5 – Validation Report Summary for a Kit or Reagent

(Add more lines in all boxes below, as required)

1) Brief description of the assay:

2) Project team			
Role, eg statistician, lab. worker	Name	Laboratory	Area of expertise eg, statistician, molecular scientist, HIV serology
Project Manager			
Project Leader			

3) Purpose of assay and background, including reason for introduction:

4) Brief details of assay validation plan:

5) Relevant SOPs	
Number	Title

6) Relevant COSHH and other risk assessments not listed on SOPs	
Number	**Title**

7) Cross-reference all other related documents associated with this study (list can be added to and deleted as appropriate)				
	YES	NO	N/A	Location of Documents
Comparisons with previously used or alternative test methods				
EQA data (several years, as available)				
IQC data				
IQA data				
In-house R&D records				
Work books (especially applying to in-house testing)				
Results of testing:				
known positives				
known negatives				
low positives				
high positives				
challenging samples				
	YES	NO	N/A	Location of Documents
Have the following assay characteristics been evaluated/reviewed:				

i. Sensitivity ii. Specificity iii. Reproducibility				
Published and unpublished papers and reports, including formal evaluations				
Work carried out with collaborating laboratories				
Previous test results				
Review meeting minutes				
Manufacturer's instructions				
Manufacturer's product specification				
Product safety data sheets				
Method descriptions and SOPs relevant to performing assay being validated				
Any other supporting information				
Has the test been validated by collaborating laboratories				
Has the assay been costed?				
Have customers been informed of significant changes in assay performance: eg, Sensitivity, specificity, turnaround times				
Has the User Manual been updated				

8) Diary (include dates of all important events, such as review meetings)	
Event	Date
Project start	

9) Conclusions (include brief summary)

Complete validation checklist before completing the authorisation section below

VALIDATION AUTHORISATION SECTION

This assay is suitable for use

Signed (Project Leader) Date_____

Introduction of assay authorised

Signed (Project Manager) Date_____

Appendix 6 - A Guide to the Level of Revalidation Considered Acceptable For a Range of Changes to the Protocol

	Protocol change					
Revalidation activity	primers/ probes	reaction mix components (ex – primers/ probes)	extraction method	control material	instrument or conditions	analysis
Analytical sensitivity	yes	yes	yes	no	yes	yes
Diagnostic sensitivity	yes	*	yes	no	*	*
Analytical specificity	yes	*	limited$	limited	limited	limited
Diagnostic specificity	yes	*	*	no	*	*
Efficiency (quantitative assays)	yes	yes	no	no	yes	yes
Linearity (quantitative assays)	yes	*	yes	no	*	*
Measurement range	yes	*	*	no	*	*
Precision (quantitative assays)	yes	*	yes	no	*	*
Reproducibility (quantitative assays)	yes	*	yes	no	*	*
Analytical accuracy (quantitative assays)	yes	*	yes	yes	*	*
Clinical accuracy (quantitative assays)	yes	*	yes	no	*	*
Reference intervals	yes	*	*	no	*	*
Clinical validation	yes	*	*	no	*	*
Shelf-life	partial&	partial	no	partial	no	no

* the level of revalidation required will depend upon professional judgement in the particular circumstances of the assay.

$ for limited revalidation of analytical specificity it is expected that testing of a reduced panel of strains would be sufficient.

& for partial revalidation of shelf-life re-testing of a single batch is expected.

Appendix 7 – Project Plan Sign-Off Form

This form should be completed and attached to the Project Plan

Title of project: *insert project title*

Names and roles of review team members: *insert list here*

The project plan has been reviewed and conforms to guidance on the development and validation of diagnostic tests that depend on nucleic acid amplification and detection

Chair of review team signature: *insert signature and printed name*

Validation plan sign-off form

This form should be completed and attached to the Validation Plan

Title of project: *insert project title*

Names and roles of review team members: *insert list here*

The validation plan has been reviewed and conforms to guidance on the development and validation of diagnostic tests that depend on nucleic acid amplification and detection

Chair of review team signature: *insert signature and printed name*

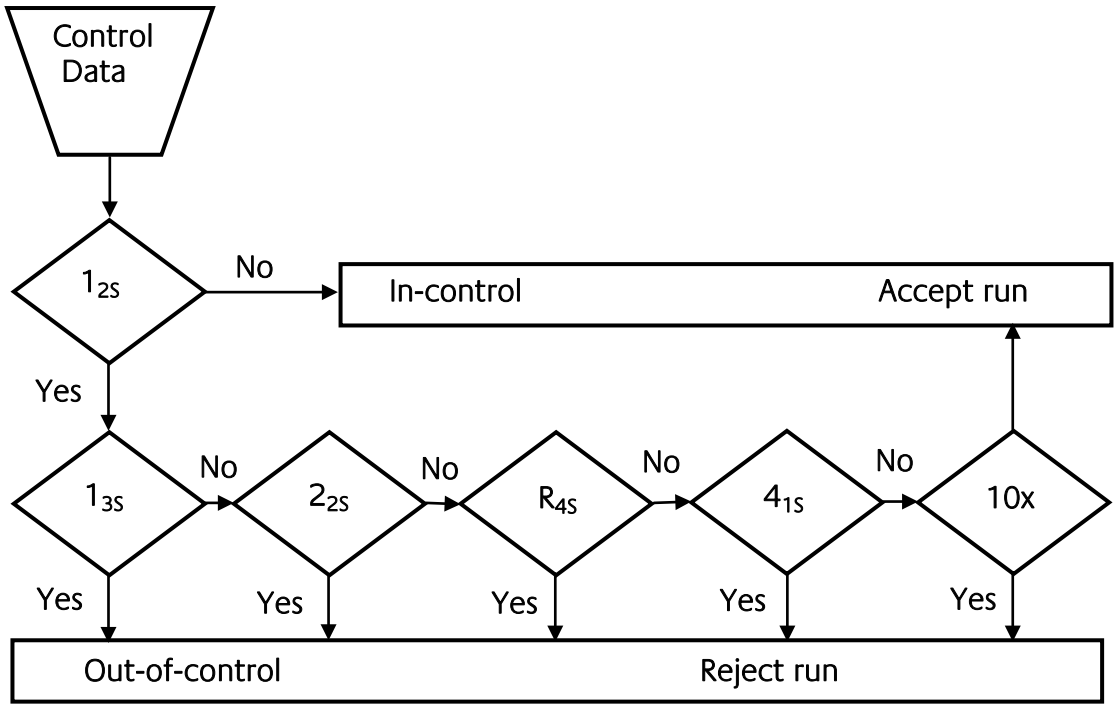
Appendix 8 - Westgard

Application of Westgard rules to run QC for nucleic acid amplification based diagnostic assays

Introduction

Westgard rules are designed to identify routine runs for rejection on the basis of the positive control results that fail one or more of a series of tests. The tests are selected to be stringent in their ability to identify poor performance whilst having a low unwarranted rejection rate. For example, test one is often whether the positive value exceeds the mean +/- 3SD (1_{3S} rule) and test two whether two consecutive values in the same run exceed the mean +/- 2SD (2_{2S} rule). A combination of these rules as illustrated in the Figure is a reasonable alternative to application of a single 1_{2S} rule which would result in a greater occurrence of false rejections.

Figure. A typical use of Westgard rules



In real-time PCR assays there are two types of control that might be considered for use as run quality indicators, internal and external (ie batch) positive controls. The potential use of internal control (IC) data for run QC is attractive but requires care. Internal controls are designed primarily to allow the identification of individual test failures. Depending upon the assay design, the IC may give a low value in an individual test when the target of the assay is present at high levels. Alternatively, in the absence of a strong test signal, a low IC value may indicate that an inhibitor is present in the sample or sub-optimal performance of other reagents in a single tube. Although, these features of ICs in real-time PCR are problematic for application to run QC, ICs may nevertheless be useful for run QC providing the data are screened according to pre-determined rules designed to remove datum points where the IC fails due to factors that do not effect the entire run.

A significant problem in applying Westgard rules to real-time PCR data is that it is necessary that the positive control results have a normal distribution. Examination of data obtained within Microbiology Services Division for a number of assays with different targets indicates that plots of crossing threshold (CT) versus number of results within each bin are not normally distributed. The distribution is skewed to the right (ie more spread on the side of the plot

representing high Cq values. The result of this skew is that a standard application of the Westgard rules relying on standard deviations will result in unwarranted rejection of runs where the Cq values are high but actually within the expected range (ie not indicating a problem with the run). To avoid this problem it is recommended that SD values are calculated separately for points above and below the mean. This correction is designed to ensure that the Westgard rules are equally 'tight' both above and below the mean. It is not designed to create an equal probability of runs failing due to both cases since theory dictates that reaction failures are most likely to give high Cq values.

Rules for screening IC data

Where IC data are to be used for run QC it is essential that the results are screened to remove datum points affected by either competition or inhibition since these factors are influenced by the sample independent of the run quality.

To avoid run rejection due to high positive samples out-competing their ICs, it is advised that ICs in samples that give positive results below a predetermined quantification cycle (Cq) be eliminated from the QC analysis. The level will vary depending upon the assay design but should set on the basis of empirical evidence. For example, the rule might be that IC data points are discarded when the positive sample Cq reaches the level required to result in a two cycle mean increase in the Cq of the IC.

To avoid run rejection due to sample inhibition the following steps are advised. Since ICs are normally spiked at close to the limit of detection they may fail (no Cq) when the sample contains relatively modest levels of inhibition. Where the IC fails the result should be discarded for run QC purposes. To further eliminate the possibility of unwarranted run rejection IC data points should be discarded when the Cq values are three cycles higher than the mean. This rule essentially eliminates the possibility of run rejection due to high Cq s of ICs in individual tests.

Selection of Westgard rules

The Westgard rules should be applied to all external controls and ICs screened as described above. Means and SDs should be estimated using control measurements obtained under normal run conditions (ie a validation series). For real-time PCR positive controls a series of >50 Cq values from >5 runs should be considered a reasonable minimum. Control measurements from invalid runs should not be included (ie when setting up series for a new assay each run must first be screened against the remaining runs prior to inclusion). When using the run QC monitoring approach described below the number of run control measurements should not exceed 10% of the number of control measurements in the validation series.

Example of an acceptable set of rules for real-time PCR run QC

This set of rules may not be useful for all assays. More work is required to determine generally applicable rules for use with real-time PCR positive control data. As explained above Cq values for a positive control in real-time PCR are often non-normal in their distribution. To overcome this issue SDs for Cq s above and below the mean should be evaluated and applied independently.

Rule one should be 1_{3S} (ie reject when one control measurement in a group exceeds the mean plus 3s). Application of this rule is expected to lead to a false rejection rate of approximately 1% for runs including 2-4 controls.

Rule two should be 2_{2S} (ie two consecutive violations of the 1_{2S} rule in a run).

Rule three should be R_{4S} (ie reject when one control measurement in a run exceeds the mean plus 2s and another exceeds the mean minus 2s).

Rule four should be 4_{1S} (ie reject when four consecutive control measurements in a run exceed the same mean plus 1s or the same mean minus 1s control limit).

Rule five should be Nx (ie reject when N consecutive control measurements fall on one side of the mean where N is the lesser of the number of control measurements or 8).

Appendix 9 – Troubleshooting Diagnostic Real-Time PCR

Problem	Technical issue(s)	Corrective action(s)	Clinical implication(s)
Low analytical sensitivity	Poor nucleic acid solubilisation (eg lysis)	Modify extraction method	False negatives
Low analytical sensitivity	Nucleic acid hydrolysed	Modify method	False negatives
Low analytical sensitivity	Nucleic acid cross-linked	Modify method to use shorter amplicon	False negatives
Low analytical sensitivity	Poor extraction efficiency	Modify method	False negatives
Low analytical sensitivity	Annealing temperature too high	Reduce temperature	False negatives
Low analytical sensitivity	Annealing time too short	Increase duration	False negatives
Low analytical sensitivity	Extension time too short	Increase duration	False negatives
Low analytical sensitivity	Denaturation time too short	Increase duration	False negatives
Low analytical sensitivity	Denaturation time too long	Decrease duration	False negatives
Low analytical sensitivity	Denaturation temperature too high	Reduce temperature	False negatives
Low analytical sensitivity	Poor nucleic acid purity – PCR inhibitors	Modify extraction method	False negatives
Low analytical sensitivity	Poor nucleic acid purity - macromolecules	Modify extraction method	False negatives
Low analytical sensitivity	Denaturation temperature too low	Increase temperature	False negatives
Low analytical sensitivity	dNTPs hydrolysed	Replace dNTPs and examine storage protocol	False negatives
Low analytical sensitivity	Polymerase activity too low	Examine protocol	False negatives
Low analytical sensitivity	Primers or probes contaminated, degraded or adsorbed to storage container	Replace reagent and examine storage protocol	False negatives
Low analytical sensitivity	Primers bind target inefficiently	Redesign primers or annealing conditions	False negatives
Low analytical sensitivity	Primer concentration too low	Increase concentration	False negatives
Low analytical sensitivity	Secondary structure within PCR amplicon	Redesign primers, modify annealing	False negatives

		conditions or increase primer concentration	
Low analytical sensitivity	Primers form dimers or other artefacts	Redesign primers or modify protocol (eg reduce primer concentration or use hot-start	False negatives or false positives
Low analytical sensitivity or reproducibility	Low pipetting accuracy	Check instrumentation or revise protocol (ie to eliminate low volume pipetting)	False negatives or inaccurate quantification leading to wrong diagnosis
Low analytical sensitivity or reproducibility	Poor real-time PCR instrument performance	Check instrumentation	False negatives, false positives or inaccurate quantification leading to wrong diagnosis
Low analytical sensitivity or reproducibility	Poor extraction instrument performance	Check instrumentation	False negatives or inaccurate quantification leading to wrong diagnosis
Low analytical sensitivity or specificity	Poorly chosen PCR cycle parameters	Redesign parameters	False negatives or false positives
Low analytical sensitivity or specificity	Poorly chosen PCR reaction mix	Redesign mix	False negatives or false positives
Low analytical specificity or specificity	Polymerase activity too high	Reduce polymerase concentration	False negatives or false positives
Low analytical sensitivity or quantitative accuracy	Probe system yields poor signal/noise ratio	Redesign probe	False negatives or inaccurate quantification leading to wrong diagnosis
Low analytical reproducibility	Polymerase from a supplier with erratic quality control (eg enzyme supplied with variable specific activity or concentration)	Change supplier	False negatives, false positives or inaccurate quantification leading to wrong diagnosis
Low analytical reproducibility	Poor PCR set-up robot performance	Check instrumentation	False negatives or inaccurate quantification leading to wrong diagnosis
Low analytical reproducibility	Low pipetting reproducibility	Check instrumentation	False negatives or inaccurate quantification leading to wrong diagnosis

Low analytical reproducibility	Poor nucleic acid purity - salts	Modify extraction method	False negatives, false positives or inaccurate quantification leading to wrong diagnosis
Low analytical specificity	Annealing temperature too low	Increase temperature	False positives
Low analytical specificity	Annealing time too long	Reduce duration	False positives
Low analytical specificity	Primer concentration too high	Reduce concentration	False positives
Low analytical specificity	Extraction instrument causes cross-contamination	Check instrumentation	False positives
Assay not quantitative	Low extraction matrix capacity	Modify extraction method	Inaccurate quantification leading to wrong diagnosis
Low diagnostic sensitivity	Primers do not match all intended target strains	Redesign primers	False negatives
Low diagnostic sensitivity	Probe(s) do not match all intended target strains	Redesign probe(s)	False negatives
Low diagnostic specificity	Primers match non-target strains	Redesign primers	False positives
Low diagnostic specificity	Probe(s) match non-target strains	Redesign probe(s)	False positives
Internal control fails	Multiple potential issues		None provided assay is interpreted according to appropriate SOP
Positive control fails	Multiple potential issues		None provided assay is interpreted according to appropriate SOP
Negative control fails	Cross-contamination		None provided run is rejected
Run fails multi-rule QC	Multiple potential issues		None provided run is interpreted according to appropriate SOP

References

1. European Commission. Annex 15 to the EU Guide to Good Manufacturing Practice. 2001. p. 1-11.

2. Banoo S, Bell D, Bossuyt P, Herring A, Mabey D, Poole F, et al. Evaluation of diagnostic tests for infectious diseases: general principles. Nat Rev Microbiol 2010;8:S17-S29.

3. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Toward complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Clin Chem 2003;49:1-6.

4. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009;55:611-22.

5. Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. Clin Microbiol Rev 2010;23:550-76.

6. Patient-Centered Outcomes Research Institute. Draft Methodology Report: "Our Questions, Our Decisions: Standards for Patient-centered Outcomes Research". p. 1-206.

7. Patient-Centered Outcomes Research Institute. Standards in the Design, Conduct and Evaluation of Diagnostic Testing For Use in Patient Centered Outcomes Research. p. 1-59.

8. Stability testing of in vitro diagnostic reagents. British Standards Institution. 2002.

9. Application of risk management to medical devices. British Standards Institution. 2009.

10. Molecular diagnostic methods for infectious diseases; approved guidelines, 2nd edition. CLSI MM3-A2 ed. 2006.

11. Lange M, Neubauer H, Seyboldt C. Development and validation of a multiplex real-time PCR for detection of Clostridium chauvoei and Clostridium septicum. Mol Cell Probes 2010;24:204-10.

12. Picard FJ, Gagnon M, Bernier MR, Parham NJ, Bastien M, Boissinot M, et al. Internal control for nucleic acid testing based on the use of purified Bacillus atrophaeus subsp. globigii spores. J Clin Microbiol 2009;47:751-7.

13. Brightwell G, Pearce M, Leslie D. Development of internal controls for PCR detection of Bacillus anthracis. Mol Cell Probes 1998;12:367-77.

14. Pugnale P, Latorre P, Rossi C, Crovatto K, Pazienza V, Gottardi AD, et al. Real-time multiplex PCR assay to quantify hepatitis C virus RNA in peripheral blood mononuclear cells. J Virol Methods 2006;133:195-204.