GUIDANCE FOR THE VALIDATION OF

PHARMACEUTICAL QUALITY CONTROL

ANALYTICAL METHODS

NHS PHARMACEUTICAL QUALITY ASSURANCE COMMITTEE

FIRST EDITION
MARCH 2005

GUIDANCE FOR THE VALIDATION OF PHARMACEUTICAL QUALITY CONTROL ANALYTICAL METHODS

Prepared on behalf of the NHS Pharmaceutical Quality Assurance Committee by:

Mr. R. Parkinson
Ms. J. Beaufoy

This is one of a series of guidance documents prepared by the NHS Pharmaceutical Quality Assurance Committee.

Contact:

Mr. R. Parkinson
Quality Control Laboratory
Nottingham City Hospital
Hucknall Road
NOTTINGHAM
NG5 1PB

Tel: 0115 9934975
Fax: 0115 9628013
e-mail: rparkins@ncht.trent.nhs.uk

GUIDANCE FOR THE VALIDATION OF PHARMACEUTICAL QUALITY CONTROL ANALYTICAL METHODS

CONTENTS

GUIDANCE FOR THE VALIDATION OF PHARMACEUTICAL QUALITY CONTROL ANALYTICAL METHODS

PURPOSE AND SCOPE

This document is intended to provide a framework for NHS Pharmaceutical Quality Control laboratories to enable a consistent and robust approach to analytical method validation. The document is written in the context of the quality control testing of unlicensed medicines manufactured under a 'Specials' licence and it should be recognised that validation of methods for the purpose of submissions for Marketing Authorisations may need to go into further detail, although the same principles will apply.

The document is intended to form a basis for the introduction of a national database of validated methods for use throughout the NHS. Thus although it is written in the form of a guidance document, the recommendations contained herein must be followed before a method can be accepted onto the national database, in order to assure other users that the method has been validated to an agreed standard. It is accepted that laboratories developing methods for their own local use might decide that it is not appropriate to follow all the steps outlined in the document, although it is hoped that the guidance will still be of value.

Laboratories accessing methods on the database need to recognise that some local verification might still be necessary, for example if there are slight differences in the product matrix or equipment used. A section on verification is therefore also included. System suitability testing is considered in this section.

As well as testing routine manufactured products or their ingredients, NHS Quality Control Laboratories are often called upon to perform one off or ad hoc investigations. In these cases, carrying out a full method validation would not be appropriate or a justified use of resources. It is necessary to carry out some validation in order to demonstrate confidence in the results. A section on abridged validations is also included.

For simplicity, throughout this document, drug products, in process samples or other prepared materials or mixes, for which the test method is being validated, will be referred to as 'products'. Pure raw materials, ingredients or other single substances will be referred to as 'substances'. The word 'analyte' is used generically to refer to the entity being tested for, for example, the active ingredient or an impurity.

In each section, suggested acceptance criteria are given. These are intended for guidance and when deciding on acceptance criteria, fitness for purpose must always be considered. Sometimes acceptance results tables are given. These give example comparisons of results with acceptance criteria and possible decisions. These do not necessarily show all possible combinations of results but offer pointers as to ways to interpret various results.

Validation of microbiological testing methods is outside the scope of this document.

Pharmacopoeial methods are generally accepted as validated within the context of the monograph in which they appear but will usually require some local verification when used to test non-monograph products or substances.

The main sections on validation procedure and abridged validation are divided into four sections, viz: identity tests, qualitative impurity testing, quantitative impurity testing and assay. Whilst this lengthens the document and inevitably leads to some duplication, it means that in use, only the section appropriate to the purpose of the test being validated needs to be considered.

INTRODUCTION

According to ISO 9000:2000 validation is confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled.

There is a temptation, when presented with the problem of introducing a new method, particularly when there is a pressure of time, to carry out very little, or even no formal validation. It is very easy to think up a method, try it on the sample and if the results are within the expected range, to consider the method has 'worked' and is 'fit for purpose'. This lack of validation may occur with the best of intention, thus a method might be used to test a 'one off' sample. This 'one off' then gradually increases in frequency until it becomes routine, but validation of the method somehow gets overlooked on the way.

There are various reasons for carrying out method validation. These include:

- Ethical – establish fitness for purpose on behalf of clients.
- Commercial – "Due Care" in product liability.
- Regulatory – GMP/GLP requirement. Also other accreditation standards would demand evidence of fitness for purpose.
- Transportability – Considerable resource goes into method development. Validation to an agreed standard allows methods to be adopted elsewhere with confidence.

Method validation is therefore a tool to demonstrate that a particular method is fit for its intended purpose. Prior to method validation, there will generally be a period of method development, which will be used as a preliminary indication that the method is likely to provide the information required of it. Some of the data generated during this stage might be usable when the full method validation process is carried out. For example, if the analyte in question is considered likely to exhibit suitable absorption of ultraviolet light, the first step might be to prepare a simple solution to check the response. Prior to carrying out a full validation study however, it would be advisable to check linearity. Additionally, if the method is intended to use a fixed A(1%, 1 cm) this might be determined during the linearity test. However, linearity would be one of the parameters to be checked during validation. It can be seen therefore that the distinction between development and validation can become blurred. Once a method is deemed to have potential, a validation protocol should be drawn up, with consideration given as to the sequence in which the steps are carried out. Thus in the above example, the linearity and determination of A(1%, 1 cm) would be carried out early in the validation and the A(1%, 1 cm) value incorporated into the method to test further parameters such as reproducibility.

VALIDATION PLAN

Validation of analytical methods is not accidental but must be a planned activity with the acceptance criteria for each performance parameter being defined objectively in advance on the basis of fitness for purpose. Prior to commencement, a validation plan needs to be drawn up for the method.

Although there are many types of analytical method, guidance is limited to four broad categories within which most methods can be considered or from which the principles outlined here can be applied. These are identification, qualitative impurity testing (limit tests), quantitative impurity testing and assay.

The main steps in drawing up a validation plan are:

1. Decide the analytical requirements

a. Define the analyte, its normal concentration range and the range the method needs to be capable of quantifying.
b. Define the matrix. For the purposes of this document the matrix is generally taken as being the simplest available vehicle for the product. This could for example be water for a simple aqueous solution or aqueous cream, where this is used in the preparation of a cream.
c. Define any potential interferents. These may be other excipients in the formulation, possible impurities or in the case of methods testing for impurities might be the active ingredient.

2. Define the performance requirements

a. Performance characteristics

The performance characteristics which need evaluating must be defined. These will depend on the type of test being validated and its intended purpose. As guidance, the following examples give an indication as to the likely parameters to consider.

i. Identification tests:

Specificity is likely to be the most important, if not the only, parameter to consider. The context in which the test is to be carried out is also important as supporting tests or data can be used to overcome lack of specificity. For example, finished products might require less specific tests than raw materials, particularly if the test is carried out for an in-house production facility where production records are available to the laboratory and there is limited scope for interfering ingredients to be used in error.

ii. Qualitative impurity testing (limit tests):

Specificity and Limit of Detection are most likely to be considered however, other parameters might also be applicable especially if the method is semi-quantitative. Again it is important that specificity is considered in light of potential interferents.

iii. Quantitative impurity testing:

It is likely that bias, precision, specificity, quantitation limit and linearity will need to be evaluated. In addition to this, the range for which the method has been validated should be defined.

iv. Assay:

It is probable that bias, precision, specificity and linearity will need to be evaluated. In addition to this, the range for which the method has been validated should be defined. Quantitation and detection limits are generally not required unless concentrations of analyte are very low or the method lacks sensitivity.

b. Performance acceptance criteria

Having decided which performance parameters need evaluating, it is then necessary define the acceptance criteria. Acceptance criteria may be quantitative or qualitative and depend on the purpose to which the analytical procedure being validated is to be put. Acceptance criteria are discussed in greater detail in the next section so they can be viewed in the context of the testing methodology.

3. Define the methods of evaluation

For each performance requirement, the tests to be performed must be defined. This will be a brief statement of what tests are carried out and on what samples. Where a test or series of tests can provide results for more than one performance characteristic cross references will be made. For example, results from tests used to determine precision may also be used to determine bias.

The tests required to evaluate each performance characteristic are described in the following section.

4. Describe how the data are to be evaluated

For each performance requirement, the method of evaluating the data must be defined. This will again be a brief statement on how to process the data and make a comparison with the acceptance

criteria. Where a test or series of tests can provide results for more than one performance characteristic cross references will be made, although the method of processing the data will probably be different. For example, where results from tests used to determine precision are also used to determine bias, precision will be assessed by looking at variances whereas bias will be assessed by looking at the means, although some consideration of precision will also be involved.

Detailed methods of processing the data from the tests required to evaluate each performance characteristic are described in the following section.

VALIDATION PROCEDURE

The methods for evaluating each performance characteristic in the context of the type of test where they are typically required are described below. Each of the four categories of test (identity, qualitative impurity, quantitative impurity and assay) is considered in turn. Under each category, the performance characteristics most likely to be evaluated are discussed. For each performance characteristic, as it applies to the category of test, a brief discussion of the factors likely to be considered in deciding its significance is given (principles) followed by methodologies and recommended acceptance criteria (testing). Where a characteristic needs to be evaluated in a different context, parallels can be drawn in deciding how to carry out the evaluation and what acceptance criteria might be appropriate.

1. Identification

Specificity

Principles:

Identification tests should be able to discriminate between compounds of closely related structures which are likely to be present. For the purpose of validation, in order to achieve the required specificity, identity tests will be validated as a set of tests and not as individual tests.

For substances, due to the potential for labelling errors, it is not generally possible to decide upon all potential interferents, although there might be some closely related substances, which need differentiating. It is therefore necessary, wherever possible to select methods generally accepted as specific and if necessary applying the results of a combination of tests. Thus for organic compounds IR spectroscopy is often the method of choice, coupled with tests such as optical rotation, to differentiate different optical isomers, or wet chemistry identification of ions to differentiate salts or free bases.

In the case of products, then, where the test is being applied for the purpose of testing of products manufactured in-house, the requirements for specificity might be less exacting than when the test is applied to products prepared externally. This is because other documented data (eg. batch production records) can be used to support the identity and the potential interferants are probably more readily known.

Testing:

Apply the set of identity tests being validated to a sample comprising the analyte (in the appropriate matrix if applicable), samples comprising the known or expected interferents (in the appropriate matrix if applicable) and, where applicable, a sample comprising the product matrix. If the matrix alone gives a positive reaction or the sample gives a negative reaction to any of the tests, that test must be removed from consideration.

If chromatography is used as an identity test, either in combination with other tests or in isolation, the same samples as above would be subjected to the test. A resolution factor of >=1.5 between the analyte and any other peak is generally deemed acceptable (ie. positive for the analyte and negative for the interferent).

Acceptance results table:

Sample	Results			
Analyte (in matrix)	All tests +ve	All tests +ve	All tests +ve	All tests +ve
Interferent A (in matrix)	All tests –ve	All tests –ve	Some tests +ve	All tests –ve
Interferent B (in matrix)	All tests –ve	Some tests +ve	Some tests +ve	All tests +ve
Matrix (if applicable)	All tests –ve	All tests –ve	All tests –ve	All tests –ve
Decision	Accept	Accept with caution	Accept with caution	Reject

2. Qualitative impurity testing (limit tests)

Impurities to be tested for will generally be either present through the manufacturing process of an ingredient, a related substance ineffectively removed from an ingredient or a degradation product of an ingredient.

Specificity

Principles:

Qualitative impurity tests should primarily be able to discriminate between the analyte (ie. impurity) and the active ingredient, which may be a closely related compound. It must be able to detect the analyte in the presence of (usually) significantly higher concentrations of the ingredient and without interference from other ingredients or the matrix.

Testing:

Where the analyte is known and available, this should be used for spiking samples (product or substance). Where the analyte is not known or is unavailable as will sometimes be the case with degradation impurities, samples should be stored under suitable stress conditions eg. light, heat, humidity, acid/base hydrolysis and oxidation. Apply the limit test being validated to a sample spiked with the limit concentration of the analyte or degraded as above, an unspiked sample consisting of the product or substance and the matrix, if applicable.

Acceptance results table:

Sample	Results			
Analyte (spiked in sample)	Test +ve	Test +ve	Test –ve	Test +ve
Unspiked sample	Test –ve	Test +ve	Test –ve	Test +ve
Matrix (if applicable)	Test –ve	Test –ve	Test –ve	Test +ve
Decision	Accept	Reject	Reject	Reject

If chromatography is used, the spiked sample should show a resolution factor of >=1.5 between the analyte and any other peak. The unspiked sample should show nothing which will interfere with detection of the analyte, however a small peak might be observed if some of the analyte is present as an impurity. Where this is the case, the spectrum and purity of the peak can be examined if possible to confirm it is due to the analyte. The matrix, where tested, should show no peak corresponding to the analyte peak.

Detection limit

Principles:

Most qualitative limit tests will involve the comparison of the test solution, appropriately treated, with a standard solution of the impurity similarly treated. The comparison might be made either visually (eg. a precipitate or colour) or instrumentally. Determination of an absolute value for the

detection limit is generally not required. A limit for the impurity is generally set on the basis of other criteria and validation needs to demonstrate that the method will give reliable results at that level. This is therefore not strictly the detection limit but seeks to confirm the test is being carried out above what would be the theoretical detection limit. If the impurity is not known or not available, this test cannot be done.

Some limit tests might be considered semi-quantitative. For example, sometimes HPLC might be used and the limit quoted in terms of a comparison of secondary peaks with the main peak. For the purpose of this document, such tests would be classed as qualitative.

Testing:

Where the test involves comparison of the test with a standard and the impurity is known and available, a standard containing the impurity at its limit value in a suitable medium should be prepared. Test samples consisting of the product or substance spiked with the analyte (impurity) at concentrations equivalent to 80%, 100% and 120% of the limit value should be prepared. These should be subjected to a blind limit test in duplicate on three different days, if possible using different members of staff (ie. six tests at each concentration in all). Where the test is judged visually, the results should be recorded as ‘pass’, ‘fail’ or ‘no difference’. If the result is measured instrumentally, test readings differing by greater than 10% from the standard are recorded as a ‘pass’ or ‘fail’ whereas readings differing by no more than 10% are recorded as ‘no difference’.

Acceptance criteria:

For acceptance, all tests at the 80% level should be reported as ‘Pass’ and all results at the 120% level as ‘Fail’. The results at 100% should be a mix of predominantly (ie. at least 50%) ‘no difference’ with some ‘pass’ and/or ‘fail’.

Where the test involves comparison of secondary peaks to main peaks, test a sample consisting of the product or substance spiked with the analyte (impurity) at a concentration equivalent to its limit value. This might need to be done by stressing samples and by trial and error. Carry out the limit test on the sample in triplicate and calculate the area of the secondary peak relative to the main peak. Calculate the percentage coefficient of variation (%CV).

Acceptance criteria:

It is generally acceptable for the %CV to be below 10%.

3. Quantitative Impurity Testing

Impurities to be tested for will generally be either present through the manufacturing process of an ingredient, a related substance ineffectively removed from an ingredient or a degradation product of an ingredient. Quantitative impurity testing would also be used during product development for stability trials.

Generally, validation of a quantitative impurity test method requires the impurities to be known and available, for the preparation of standards. If the impurity being quantified is not known or is not available, the only way a quantifiable method can be validated is by comparison with a second well-characterised procedure (eg. pharmacopoeial or other fully validated method), which does not require the use of standards. If this is not the case, a truly quantitative test cannot be carried out.

Tests such as HPLC where impurities are limited by comparing the areas of their peaks with that of the active ingredient (eg. ‘the area of any subsidiary peak is not greater than 1% of the area of the main peak’) would be considered qualitative.

Before carrying out the testing, nominal limit concentrations need to be decided upon, so that appropriate concentrations can be tested. Where the limit is known, this will be used as the nominal limit concentration. Where there is no limit value defined or it is not known, it is

recommended that the limit is set to 5% for stability testing and 1% for other impurity testing, both relative to the concentration of the active ingredient or the ingredient from which the impurity derives. For example, if a method were being used for the stability study of an aspirin suspension with the nominal concentration of 100 mg in 5 ml, the limit for salicylic acid would be set to 5 mg in 5 ml. When preparing test samples all ingredients other than the analyte would be present at their nominal concentration.

Specificity

Principles:

Quantitative impurity tests should primarily be able to discriminate between the analyte and the active ingredient or the ingredient from which the impurity derives, which may be a closely related entity. It must be able to quantify the impurity in the presence of (usually) considerably higher concentrations of the ingredient and without interference from the matrix.

For impurity testing, it is recommended that the matrix should affect results by no more than 2% and other interferents by no more than 5%.

Testing:

Where the impurities are known and available, perform the impurity test, in duplicate, on a sample comprising the nominal limit concentration of the analyte (impurity) in the appropriate matrix if applicable, a sample consisting of the product or substance spiked with the same nominal limit concentration of the analyte and a sample of unspiked product or substance. Calculate the results with reference to the analyte sample results. Care must be taken in preparing samples to ensure the impurity is not unintentionally present due to other ingredients, in particular the main active.

Acceptance results table:

Sample	Reported analyte content R (% analyte sample result)			
Analyte (in matrix)	100%	100%	100%	100%
Substance or product spiked with analyte	95%<R<105%	95%<R<105%	R > 105%	R < 95%
Unspiked substance or product	R < 2%	R > 2%	R < 2%	R < 2%
Decision	Accept	Reject – Consider using matrix as a blank	Reject – Positive interference from active	Reject – Negative interference from active

If chromatography is used, the spiked sample should show a resolution factor of >=1.5 between the analyte and any other peak. The unspiked sample should show nothing which will interfere with detection of the analyte, however a small peak might be observed if some of the analyte is present as an impurity. Where this is the case, the spectrum and purity of the peak can be examined if possible to confirm it is due to the analyte.

Where the impurities are unknown or are not available, a second well-characterised procedure (eg. pharmacopoeial or other fully validated procedure) is used as a reference method. Perform the impurity test by both methods, in triplicate, on a sample containing the nominal concentration of the active ingredient (in the appropriate matrix if applicable). As the impurity is not available, it might be necessary to stress the sample to generate sufficient impurity to quantify.

Results need to be calculated and compared to the reference method.

Acceptance criteria:

The impurity profiles produced by both methods should be similar. There should be no practically significant difference between the results given by the two methods (see section on statistical analysis of results). The results of this test may be used as a measure of bias (see later).

If chromatography is used as the impurity test, in addition, a resolution factor of >=1.5 between the ingredient and the impurity is acceptable. Peak purity tests, where available, confirm the peak for the analyte is not attributable to more than one compound.

Linearity

Principles:

A linear relationship should be demonstrable over the range of use for the impurity testing. The range of use will depend on the purpose to which the test is to be put. Linearity is often considered only in the context of instrumental analysis, however, it is necessary to demonstrate that other methods such as volumetric analysis react quantitatively over the range of interest. For the purpose of validation, consideration of linearity will be restricted to instrumental analysis. Linearity of other types of method will be accepted provided acceptable bias values are demonstrated at the centre and both extremes of the range (see section on bias).

Most instrumental methods will involve some sample preparation followed by a determination of concentration by instrumental means such as spectrophotometry or HPLC. Linearity of the instrument response and linearity of the whole method need to be demonstrated. If the instrument response is linear but the whole method does not show linearity, this must be attributable to some effect such as interferences or poor recovery. Whilst in certain circumstances it might be deemed appropriate to carry out a full linearity test for both the instrumental response and the whole method, it is generally acceptable to demonstrate whole method linearity by determining bias at the centre and both extremes of the range (see section on bias). Therefore consideration here is only given to instrument linearity.

Testing:

Where the method is to be used for routine Quality Control release testing, there will generally be a limit defined within the product specification. In this case, the purpose of the test is to ensure linearity around the concentration of the limit. It is also particularly important to safeguard against plateauing at concentrations above the limit, which could result in unsatisfactory results being reported as passes. Linearity should therefore be tested over the range of 50 to 200% of the limit concentration. Concentrations outside this range need not be quantified accurately for the purpose of release testing.

When the test is used for stability testing, although it may well be that an acceptable limit is defined, results will probably need to be quantified over a wider range. Typically it will be necessary to quantify concentrations significantly below the limit concentration, so that the course of degradation may be followed throughout the stability trial. Linearity should normally be tested over the range 10 to 200% of the proposed limit concentration.

Prepare a minimum of 6 concentrations (including a zero if required) of a standard, in triplicate. For impurity testing, constants of proportionality, eg. A(1%, 1 cm) are not generally used to calculate the concentration, it is more usual to perform the test by comparison with a standard. Therefore the exact concentration of the standards need not be known.

Where the impurity is known and available, prepare a solution of it using a pure chemical standard dissolved in the solvent to be used in the method. Using this solution, prepare a series of dilutions in triplicate. The concentrations of this series of standards should be reasonably evenly spaced over the range investigated, thus they should not be prepared by serial dilution.

If the impurity is not known or is not available, it is still necessary to demonstrate linearity. In this case, it might be necessary to generate degradation product by stressing samples or by some other means and adjusting by trial and error to obtain a concentration suitable for dilution to give a range of standards.

Where the method requires the use of an internal standard, this should be added at the same concentration to all the standards and the response calculated as the ratio of impurity to internal standard. Determine the response for each triplicate of each concentration. Where the method would normally call for replicate readings (eg. chromatography), carry out the normal number of replicates and calculate the mean for each.

Calculate the mean of the responses at each concentration and plot the mean response (y) against the concentration (x). For the purpose of demonstrating linearity, the value for the concentration used in the plot need not be the absolute concentration value (which might be difficult to plot). For example, if the series of standards is prepared by a series of dilutions from a more concentrated standard, the volumes used could be plotted. Thus if 5, 10, 15, 20, 25 and 30 ml of a standard are all diluted to (say) 100 ml for reading, then in place of concentration, 5, 10, 15, 20, 25 and 30 would be used on the plot. The plot should always be inspected for linearity and a subjective assessment made as to fitness for purpose. Even plots which show considerable tailing can give statistically good correlations but might not be fit for purpose. The regression line should be calculated by the method of least squares (see section statistical analysis of results). The correlation coefficient, y-intercept and slope of the line should be determined. Using the equation for a straight line graph (y=mx+c), calculate the expected value for y at each point and determine the residual y(observed)-y(calculated) for each replicate at each point. Plot the residuals against concentration. The residuals should be randomly distributed (about zero). If a pattern is observed, a degree of non-linearity is indicated.

Acceptance criteria:

Subjective inspection of the plot of response against concentration appears linear.

Inspection of the residuals plot should show a random distribution about zero. Where this is not the case, and a pattern is observed, this might indicate lack of linearity or difference in precision over the range. This need not lead to rejection of the method as the results might not be deemed significant in the context of the method as a whole. Bias and precision over the range need to be considered.

It is difficult to give a definitive value for an acceptable correlation coefficient, however a minimum value of 0.999 is recommended.

<u>Precision</u>

Principles:

Various levels of precision need to be demonstrable over the working range of the method. Repeatability is the least exacting measure of precision and in many respects is the least relevant because in normal use, repeatability conditions will not be achieved. It is however important to establish repeatability at an early stage as a method which is inherently unrepeatable will not be sufficiently robust to satisfy the requirements for intermediate precision and reproducibility. Intermediate precision gives the most realistic measure of precision for in-house use of a method, however does not necessarily confirm that a method is sufficiently robust to transfer to another laboratory. Where methods are to be used for in-house stability testing, it would generally be sufficient to carry out a determination of intermediate precision. Reproducibility determination seeks to confirm that the method is sufficiently robust to be adopted in other laboratories, without the need for further validation, although local verification might be needed. This should be determined for Quality Control product release limit testing.

Precision tests are carried out on single homogeneous samples. The tests can be combined with

bias testing, in which case the content of the analyte must be known. If the precision test is being carried out in isolation, the exact concentration need not be known, although it should be within the limits stated below.

For repeatability, the ICH guidelines recommend a minimum of 9 determinations over the specified range or a minimum of 6 determinations at the nominal concentration. For limit testing, in order to demonstrate precision over the range, as a minimum, samples should be tested between 50 and 200% of the limit concentration. For stability testing, precision should be determined between 10 and 200% of the proposed limit concentration. To demonstrate intermediate precision, further samples of the same homogenous mix of the nominal concentration are tested within the same laboratory, on a different day and if possible by a different analyst and using different equipment. To demonstrate reproducibility, further samples of the same homogenous mix of the nominal concentration are tested in a different laboratory.

Setting limits for precision will depend primarily on the purpose to which the method is to be put. As a general rule a percentage coefficient of variation (%CV) of 10% would be acceptable.

Testing:

For QC release methods, prepare homogeneous samples containing 50, 100 and 200% (+/- 10%, ie. 45-55%, 90-110% and 180-220%) of the limit concentration of the analyte and for stability study methods, prepare homogeneous samples containing 10, 100 and 200% (+/- 10%, ie. 9.0-11.0%, 90-110% and 180-220%) of the limit concentration of the analyte. The samples should be prepared in a formulation containing all other ingredients and excipients at their nominal concentration. If this is not possible, for example if the full formulation is not known or the separate ingredients are not available, the samples can be prepared by spiking the available formulation with the analyte. Where the testing is also to be used to determine bias, the content may need to be accurately weighed and recorded. Where the impurity is not known or is not available, it might be possible to generate degradation product by stressing samples or by some other means and adjusting by trial and error to obtain appropriate concentrations.

Carry out the determination at each concentration in triplicate under repeatability conditions, ie. same analyst, same equipment, same time. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP’s to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

To test for intermediate precision, carry out triplicate testing on the same homogenous mix of the nominal limit concentration (ie. 100%) sample as was used for the repeatability testing. The testing is done within the same laboratory, on a different day and if possible by a different analyst and using different equipment. If the test is used for stability studies and there is doubt over the stability of the product, this needs to be taken into account and it may be necessary to carry out intermediate precision tests on the same day.

To test for reproducibility, arrange for further triplicate testing on the same homogenous mix of the limit concentration sample as was used for the repeatability testing, to be tested by a different laboratory. Exact details of methodology must be provided and followed by the second laboratory. The testing is therefore done in a different laboratory, on a different day, by a different analyst and using different equipment. Again, if there is doubt over the stability of the product, this needs to be taken into account and it may be necessary to coordinate the reproducibility tests to carry them out on the same day and taking care to ensure consistent conditions of storage.

Calculate the standard deviation and percentage coefficient of variation (%CV) at each concentration, for the replicates carried out under repeatability conditions (see section on statistical evaluation).

Similarly, calculate the standard deviation and %CV for the intermediate precision and reproducibility tests using Analysis of Variance (ANOVA) calculations to give the appropriate between and within group variances.

Acceptance results table:

Tests	Comparison of %CV with specified limit			
Repeatability	Below Limit	Above limit at Low or High concentration	Below Limit	Below Limit
Intermediate Precision	Below Limit	Below Limit	Below Limit	Above Limit
Reproducibility	Below Limit	Below Limit	Above Limit	Above Limit
Decision	Accept	Reject - Consider adjusting range or changing sample size or dilutions	Reject – Consider possibility of systematic error in Lab 2	Reject – Consider ruggedness testing.

In deciding what further development work might be required, consideration can be given to other statistical data available. In particular consideration of the 'F' values from the ANOVA calculations can point to possible causes for failure to meet the set limits. If the ratio F/F_{crit} is greater than 1 this shows a significant difference between the sets of data produced on different days or in different laboratories. Values of greater than 1 can result from either within group variation (repeatability) or between group variation (intermediate precision or reproducibility). Between group variation is likely to have a more profound effect on the F ratio. A Reproducibility F Ratio of very much greater than 1, with an Intermediate Precision F Ratio of < 1 is likely to indicate a systematic error between the two laboratories maybe as a result of misinterpretation of the method or equipment calibration. Where both ratios are considerably greater than 1, this would indicate a high degree of variation between days, but not necessarily between laboratories. Where both ratios are only slightly above 1, this might indicate poor repeatability.

By looking in detail at the results, it should be possible to get an idea as to the likely problem. Thus by comparing the means, and standard deviations for each set of results, it should be possible to decide whether there is likely to be a systematic error occurring.

Where necessary, clarification of the method, calibration checks or ruggedness testing, to determine the critical factors which need tighter control can be carried out and the precision testing repeated.

Limit of detection/quantitation

Principles:

Generally, for quantitative impurity testing, Limit of Detection is not required and it is only necessary to determine the Limit of Quantitation. The Limit of Quantitation is the lowest amount of analyte in a sample, which can be quantitatively determined with suitable precision and accuracy. For the purpose of impurity testing, an acceptable level of precision would generally be represented by a percentage coefficient of variation (%CV) of 10%. Studies could therefore be carried out at a range of concentrations down to very low levels and the %CV determined to find the point at which it reaches the maximum acceptable value (eg. 10%). However, for most purposes, in considering 'fitness for purpose', this is not deemed necessary and it is adequate to demonstrate that the precision is acceptable at the lowest end of the range, which needs to be quantified. Consideration must therefore be given as to what the method is used for.

Where the method is to be used for QC release limit testing, it must be demonstrable that the method's limit of quantitation is well below the impurity limit to be applied to the product. It is

recommended that this is tested at 50% of the limit concentration.

Where the method is to be used for stability testing or other product characterisation testing, it will generally be necessary to be able to quantify levels much lower than the eventual 'limit'. As a rule, it is recommended that the limit of quantitation should be below 10% of the 'limit' value. Therefore, if the stability method needs to determine a degradation product up to a 'limit' value of 5% of the nominal concentration of the active ingredient, then the limit of quantitation should be below 0.5% of the nominal concentration of the active.

This information can be found from the tests for precision. An indication of the likely Limit of Quantitation can sometimes be gained from the linearity experiments.

Acceptance criteria:

The percentage coefficient of variation (%CV) obtained for the repeatability tests at the lowest end of the range complies with the acceptance criteria set for precision at that end of the range.

Bias

Principles:

Bias should be established over the range of the procedure. The most critical part of the range for impurity testing purposes is around the limit value, therefore most effort is concentrated on testing at this content. Bias at the extremes of the range is less critical, therefore wider tolerances can be applied and fewer samples tested. Data processing at the extremes can also be less exacting.

Bias can be established by application of the procedure to a sample containing the analyte at the required concentration in the product or substance. Where the impurity is known and available, this should be a pure reference substance (eg. BPCRS) or other certificated material. Where the impurity is not known and available, bias will have to be checked by comparison with another validated method.

Where the impurity is known and available, then, in the case of the testing of a substance, this should be added at the required concentration to the pure substance. For a product, the impurity is added at the required concentration, to the full product formulation.

Where the impurity is not known and available, bias is established by comparison of the results from the proposed method with those produced by an alternative validated method. A validated method would either be one taken from a current recognised pharmacopoeia (BP, EP or USP) or another well-characterised method, the accuracy of which is stated and/or defined. In this case it might be necessary to generate appropriate amounts of impurity by stressing the sample and dilution. A certain amount of trial and error might be required.

According to the ICH guidelines, accuracy can also be inferred once precision, linearity and specificity have been established, however, this is only likely to be the case for ingredients or simple preparations such as single ingredient solutions. As these are probably the easiest and least time consuming to test, this approach it is not recommended and a test for bias should be carried out.

The essence of bias testing, is the comparison of the reported value with the reference value to determine if there is a significant difference. It must be recognised that statistical and practical significance are not necessarily the same. Determination of statistical significance will depend on the number of test results and their standard deviation. Thus, if a large number of replicates are carried out and there is a high level of precision (ie. small standard deviation), then even a small bias will be statistically significant. Conversely, if there are only a small number of replicates and the standard deviation is high, then even a large bias would not be judged statistically significant. To prove statistically whether there was a difference equating to a bias of 1*SD would require 16 replicates and of 0.5*SD would require 55 samples. It can be seen from the above that, almost

paradoxically, the more precise a method, the more likely it is that it will show a statistically significant bias. Thus a straightforward test for a significant difference between the observed and the reference value is not, on its own, considered appropriate. The main risks are that, where the precision is low, a bias which is in practice unacceptable, would be deemed statistically insignificant or, where the precision is high, a bias which in practice is acceptable, is considered statistically significant. The first risk is minimised by taking sufficient replicates and setting limits on acceptable precision (see section on precision). The approach to be adopted is that an acceptable level of bias appropriate for the usage of the method is defined and testing carried out to determine both the statistical and practical significance of the bias in light of this limit. As a general principle, the following limits on bias are recommended.

	At the limit concentration	At extremes of range
Substance	+/- 2.0%	+/- 5.0%
Product	+/- 5.0%	+/- 10.0%

Testing:

Where the impurity is known and available, then, for QC release methods, prepare homogeneous samples containing 50, 100 and 200% (+/- 10%, ie. 45-55%, 90-110% and 180-220%) of the limit concentration of the analyte (impurity) in the substance or the product as appropriate. For stability study methods, prepare homogeneous samples containing 10, 100 and 200% (+/- 10%, ie. 9.0-11.0%, 90-110% and 180-220%) of the limit concentration in the substance or the product. The samples must be prepared quantitatively so that the concentration of the analyte is known. In the case of products, the samples should be prepared in a formulation containing all other ingredients and excipients at their nominal concentration. If this is not possible, for example if the full formulation is not known or the separate ingredients are not available, the samples can be prepared by spiking the available formulation with the analyte. Care must be taken to ensure there is no impurity present due to the active or other ingredients.

Where the impurity is not known and available, then the bias is determined by comparison with a validated reference method. The samples for testing will need to be prepared (possibly by stressing the ingredient or product) by trial and error, by assessing the concentration by the validated method and subsequent treatment or dilution, until the appropriate concentrations are obtained. It will not be possible to prepare the samples with the same level of accuracy as can be achieved by the use of certificated impurity.

Carry out the determinations at the extremes in triplicate under repeatability conditions, ie. same analyst, same equipment and same time. Carry out nine determinations at the limit concentration. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

Calculate the results in terms of percentage of the theoretical value.

For example, where the comparison is to be made against a known reference value, if the sample being tested is prepared by dissolving 'x' grams of reference standard with a certificated purity of 'p' percent, in sufficient solvent to give 250 ml, then the 'stated' content (as percent concentration) is given by the equation:

$$S = \frac{x * 100 * p}{250 * 100}$$

If the results for the measured content (again as percent concentration) are 'r' percent then the

result as a percent of stated is given by the equation:

$$C = \frac{r * 100}{S}$$

Where the impurity is not known or available and the comparison is between two methods, ‘x’ and ‘p’ will not be known. Provided the sample tested is homogeneous and used for both methods, the results are calculated in percentage terms based on the average result obtained by the reference method. Thus, if results from the reference method were, for example, within the range 4.15 to 4.32 with an average of 4.20, then individual results for the standard method for entry into the spreadsheet would range from 98.81% to 102.86% (r/4.20*100). The results for the method being validated would be calculated similarly as r/4.20*100.

In this way, the results to be processed on the spreadsheet will always be approximately 100, which simplifies the setting of cell attributes.

For the tests carried out at the extremes of the range, calculate the mean of the results for the method under test and compare this with the theoretical value taken either from the stated content and weight taken to prepare the sample, or from the mean of the results determined by the reference method.

For tests carried out at the nominal concentration, enter the results into the appropriate spreadsheet, for comparison with reference standard or comparison with reference method.

Acceptance Criteria:

At the extremes, the bias value is less than the bias limit required.

At the nominal content, interpret the results from the spreadsheet. There are a number of possible outcomes outlined below:

i. Bias is statistically insignificant. A bias which would be considered significant is less than required limit on bias

In this case we know that the bias is statistically insignificant and it must be within the limit we have set, therefore this would be acceptable.

ii. Bias is statistically insignificant. A bias which would be considered significant is greater than required limit on bias.

In this case, although we know that the bias is statistically insignificant, the results are such that a bias greater than that considered acceptable would also be statistically insignificant. It is possible therefore that although the bias is statistically insignificant, it could be practically significant (ie. outside the required limit). It is therefore necessary to look in detail at the actual results obtained. If the actual bias is within the limit set, this would indicate that the bias is unlikely to be practically significant. However this value is based on the difference between the mean values for the test and reference samples and the true value will be within a range around this value. For this reason, a check is also made on the maximum bias value. This is difference between the extreme possible values for the means based on 95% confidence limits. If this value is below the required bias limit, the method’s bias would be accepted as not practically significant. Additional replicates could be carried out to try to reduce the standard deviation.

iii. Bias is statistically significant. A bias which would be considered significant is less than required limit on bias.

Here, although the bias is statistically significant, it might not be practically significant because a statistically significant bias could be less than the required limit. If the actual bias is within the limit set, this would indicate that the bias is unlikely to be practically significant. However this

value is based on the difference between the mean values for the test and reference samples and the true value will be within a range around this value. For this reason, a check is also made on the maximum bias value. This is difference between the extreme possible values for the means based on 95% confidence limits. If this value is below the required bias limit, the method's bias would be accepted as not practically significant.

iv. Bias is statistically significant. A bias which would be considered significant is greater than required limit on bias.

In this instance, the bias would be deemed unacceptable. It might be worth considering testing additional replicates in order to reduce the standard deviation, which will have the effect of reducing the range of bias, which would be considered statistically significant.

4. Assay

Specificity

Principles:

The need for specificity in assays depends very much on the intended use. Ideally the procedure should be able to discriminate between compounds of closely related structures, which are likely to be present. It also needs to differentiate between the active ingredient and any excipients, which may be present. However, in many contexts, striving for a high level of specificity would lead to unnecessary use of time and effort. This is particularly true for assays designed for release purposes of in-house products. The level of specificity required needs to be clearly defined in light of the other tests being performed (eg. identity tests and limit tests). A good example of this is the BP assay for aspirin tablets. The assay depends on the hydrolysis of aspirin to salicylic and acetic acids by sodium hydroxide and the subsequent titration of the excess sodium hydroxide. Clearly the method will not differentiate between aspirin and its degradation products, which are controlled by a limit test for salicylic acid. Where an assay is being developed specifically for stability work, it is generally more appropriate to apply the principals described under quantitative limit testing.

It is recommended that both the matrix and the other interferents should affect results by no more than 2%.

Testing:

Where the interferents are known and available, perform the assay, in duplicate, on a sample containing the analyte (in the appropriate matrix if applicable) at its nominal concentration, the above sample spiked with appropriate levels of the known or expected interferents and, where applicable, a sample of the matrix. Calculate the results with reference to the pure analyte sample results.

Acceptance results table:

Sample	Results (% pure sample result)			
Analyte (in matrix)	100%	100%	100%	100%
Sample spiked with interferent A	98%<R<102%	98%<R<102%	R>102%	98%<R<102%
Sample spiked with excipient B	98%<R<102%	98%<R<102%	98%<R<102%	98%>R
Matrix (if applicable)	R < 2%	R > 2%	R < 2%	R < 2%
Decision	Accept	Reject - Consider using matrix as a blank	Reject – Positive interference from interferent A	Reject – Negative interference from excipient B

In addition, if chromatography is used, the spiked samples should show a resolution factor of >=1.5 between the analyte and any other peak. The matrix, where tested, should show no peak corresponding to the analyte peak.

Where the interferents are unknown or are not available and a second well-characterised procedure (eg. pharmacopoeial or other fully validated procedure), is available, perform the assay by both methods, in triplicate, on a sample containing the nominal concentration of the active ingredient (in the appropriate matrix if applicable), the product matrix and, where appropriate, the above active ingredient sample stored under stressed conditions to generate interferents.

Acceptance criterion:

There should be no practically significant difference between the results given by the method being validated and the reference method (see section on statistical analysis of results) for all samples.

Linearity

Principles:

A linear relationship should be demonstrable over the range of the assay. The range might be determined according to the purpose of the assay or a linearity test can be carried out over a wide range and the linear range determined. Linearity is often considered only in the context of instrumental analysis, however, it is necessary to demonstrate that other methods such as volumetric analysis react quantitatively over the range of interest. For the purpose of validation, consideration of linearity will be restricted to instrumental analysis. Linearity of other types of method will be accepted provided acceptable bias values are demonstrated at the centre and both extremes of the range (see section on bias).

Most instrumental methods will involve some sample preparation followed by a determination of concentration by instrumental means such as spectrophotometry or HPLC. Linearity of the instrument response and linearity of the whole method need to be demonstrated. If the instrument response is linear but the whole method does not show linearity, this must be attributable to some effect such as interferences or poor recovery. Whilst in certain circumstances it might be deemed appropriate to carry out a full linearity test for both the instrumental response and the whole method, it is generally acceptable to demonstrate whole method linearity by determining bias at the centre and both extremes of the range (see section on bias). Therefore consideration here is only given to instrument linearity.

Testing:

For assays of substances or drug products, the ICH guideline recommends a minimum of 80 to 120 percent of the nominal test concentration and for content uniformity it recommends 70 to 130 percent as a minimum. In order to allow wider use and transportability of methods and to safeguard against severe non-linearity, especially plateauing a wider range should be examined. As a minimum, instrument linearity should be tested between 50 and 150% of the nominal concentration.

Prepare a minimum of 6 concentrations of standard (including zero if required) in triplicate using a pure chemical standard dissolved in the solvent to be used in the assay method. Where the linearity testing is also being used to determine a proportionality constant, such as the A(1%,1cm), the weight of pure chemical standard must be accurately taken and recorded and three initial solutions prepared for dilution. Where the test is used solely for linearity testing, the weight need not be accurately known and a single initial solution is prepared for dilution.

The concentrations of the standards should be reasonably evenly spaced over the range investigated, thus they should not be prepared by serial dilution. Where the method requires the use of an internal standard, this should be added at the same concentration to all the standards.

Determine the response for each triplicate of each concentration. Where the method would normally call for replicate readings (eg. chromatograph), carry out the normal number of replicates and calculate the mean for each standard.

Calculate the mean of the responses at each concentration and plot the mean response (y) against the concentration (x). Where three initial solutions were prepared, it is necessary to compensate for differences in weights taken before calculating the mean. It is suggested that this is done by calculating the average weight (a) and determining what each response would have been if that weight had been taken, by the formula r*a/w, where r is the response and w is the actual weight taken. For the purpose of demonstrating linearity, the value for the concentration used in the plot need not be the absolute concentration value (which might be difficult to plot. For example, if the series of standards is prepared by a series of dilutions from a more concentrated standard, the volumes used could be plotted. Thus if 5, 10, 15, 20, 25 and 30 ml of a standard are all diluted to (say) 100 ml for reading, then in place of concentration, 5, 10, 15, 20, 25 and 30 would be used on the plot. The plot should always be inspected for linearity and a subjective assessment made as to fitness for purpose. Even plots which show considerable tailing can give statistically good correlations but might not be fit for purpose. The regression line should be calculated by the method of least squares (see section statistical analysis of results). The correlation coefficient, y-intercept and slope of the line should be determined. Using the equation for a straight line graph (y=mx+c), calculate the expected value for y at each point and determine the residual y(observed)-y(calculated) for each replicate at each point. Plot the residuals against concentration. The residuals should be randomly distributed (about zero). If a pattern is observed, a degree of non-linearity is indicated.

Acceptance criteria:

Subjective inspection of the plot of response against concentration appears linear.

Inspection of the residuals plot should show a random distribution about zero. Where this is not the case, and a pattern is observed, this might indicate lack of linearity or difference in precision over the range. This need not lead to rejection of the method as the results might not be deemed significant in the context of the method as a whole. Bias and precision over the range need to be considered.

It is difficult to give a definitive value for an acceptable correlation coefficient, however a minimum value of 0.999 is recommended.

Precision

Principles:

Various levels of precision need to be demonstrable over the working range of the method. Repeatability is the least exacting measure of precision and in many respects is the least relevant because in normal use, repeatability conditions will not be achieved. It is however important to establish repeatability at an early stage as a method which is inherently unrepeatable will not be sufficiently robust to satisfy the requirements for intermediate precision and reproducibility. Intermediate precision gives the most realistic measure of precision for in-house use of a method, however does not necessarily confirm that a method is sufficiently robust to transfer to another laboratory. Reproducibility determination seeks to confirm that the method is sufficiently robust to be adopted in other laboratories, without the need for further validation, although local verification might be needed.

Precision tests are carried out on single homogeneous samples. The tests can be combined with bias testing, in which case the content of the analyte must be known. If the precision test is being carried out in isolation, the exact concentration need not be known, although it should be within the limits stated below.

For repeatability, the ICH guidelines recommend a minimum of 9 determinations over the

specified range or a minimum of 6 determinations at the nominal concentration. In order to demonstrate precision over the range, as a minimum, repeatability should be tested between 50 and 150% of the nominal concentration. To demonstrate intermediate precision, further samples of the same homogenous mix of the nominal concentration are tested within the same laboratory, on a different day and if possible by a different analyst and using different equipment. To demonstrate reproducibility, further samples of the same homogenous mix of the nominal concentration are tested in a different laboratory.

Setting limits for precision will depend primarily on the purpose to which the method is to be put but must also be realistically achievable and are therefore also dependent on the methodology. It is therefore difficult to set hard and fast rules. Generally speaking, methods used for assaying ingredients need the highest level of precision. Limits tend to be tighter (typically 99.0 – 100.5%), any deviation in results for ingredients will reflect in the finished product and higher levels of precision are likely to be achievable due to lack of interference and matrix effects. Thus, assaying an ingredient by a method which only achieved a result which was +/- 5% would normally not be acceptable, because a result reported as 100% could actually relate to an ingredient which was only 95% pure. Even if this level of purity in itself did not present a risk, by the time additional random errors of manufacture are added, a product well outside specification could be prepared. As a general principle, the following limits for percentage coefficient of variation (%CV) are recommended:

Substance (by simple forward titration):	0.5%
Substance (by simple back titration):	1.0%
Substance (by reaction and titration):	2.0%
Substance (by instrumental method eg. uv, HPLC):	2.0%
Product (simple matrix followed by titration):	2.0%
Product (simple matrix followed by instrumental):	4.0%
Product (complex matrix requiring extraction):	7.5%

Testing:

For product methods, prepare homogeneous samples containing 50, 100 and 150% (+/- 10%, ie. 45-55%, 90-110% and 135-165%) of the nominal content in the full product formulation. For substance methods, use quantities equivalent to 50, 100 and 150% (+/- 10%, ie. 45-55%, 90-110% and 135-165%) of the quantity specified in the method. Where the testing is also to be used to determine bias, the content will need to be accurately weighed and recorded. Carry out the determination at each concentration in triplicate under repeatability conditions, ie. same analyst, same equipment and same day. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

To test for intermediate precision, carry out triplicate testing on the same homogenous mix of the nominal concentration as was used for the repeatability testing. The testing is done within the same laboratory, on a different day and if possible by a different analyst and using different equipment. If there is doubt over the stability of the analyte, this needs to be taken into account and it may be necessary to carry out the intermediate precision on the same day.

To test for reproducibility, arrange for further triplicate testing on the same homogenous mix of the nominal concentration as was used for the repeatability testing, to be tested by a different laboratory. Exact details of methodology must be provided and followed by the second laboratory. The testing is therefore done in a different laboratory, on a different day, by a different analyst and using different equipment. Again if there is doubt over the stability of the analyte, this needs to be taken into account and it may be necessary to coordinate the reproducibility tests to carry them out on the same day and taking care to ensure consistent conditions of storage.

Calculate the standard deviation and percentage coefficient of variation (%CV) at each concentration, for the replicates carried out under repeatability conditions.

Similarly, calculate the standard deviation and %CV for the intermediate precision and reproducibility tests using Analysis of Variance (ANOVA) calculations to give the appropriate between and within group variances.

Acceptance results table:

Tests	Comparison of %CV Results with specified limit			
Repeatability	Below Limit	Above limit at Low or High Concentration	Below Limit	Below Limit
Intermediate Precision	Below Limit	Below Limit	Below Limit	Above Limit
Reproducibility	Below Limit	Below Limit	Above Limit	Above Limit
Decision	Accept	Reject - Consider adjusting range or changing sample size or dilutions	Reject – Consider possibility of systematic error in Lab 2	Reject – Consider ruggedness testing.

In deciding what further development work might be required, consideration can be given to other statistical data available. In particular consideration of the 'F' values from the ANOVA calculations can point to possible causes for failure to meet the set limits. If the ratio F/F_{crit} is greater than 1 this shows a significant difference between the sets of data produced on different days or in different laboratories. Values of greater than 1 can result from either within group variation (repeatability) or between group variation (intermediate precision or reproducibility). Between group variation is likely to have a more profound effect on the F ratio. A Reproducibility F Ratio of very much greater than 1, with an Intermediate Precision F Ratio of < 1 is likely to indicate a systematic error between the two laboratories maybe as a result of misinterpretation of the method or equipment calibration. Where both ratios are considerably greater than 1, this would indicate a high degree of variation between days, but not necessarily between laboratories. Where both ratios are only slightly above 1 this might indicate poor repeatability.

By looking in detail at the results, it should be possible to get an idea as to the likely problem. Thus by comparing the means, and standard deviations for each set of results, it should be possible to decide whether there is likely to be a systematic error occurring.

Where necessary, clarification of the method, calibration checks or ruggedness testing, to determine the critical factors which need tighter control can be carried out and the precision testing repeated.

Bias

Principles:

Bias should be established over the range of the procedure. The most critical part of the range for Quality Control purposes is around the nominal value, therefore most effort is concentrated on testing at this content. Bias at the extremes of the range are less critical, therefore wider tolerances can be applied and fewer samples tested. Data processing at the extremes can also be less exacting.

Bias can be established by application of the procedure to a sample containing the analyte of known purity. For a substance this may be a pure reference substance (eg. BPCRS) or other certificated material. For a product, this should be a similar reference or certificated material added quantitatively to the full product formulation. If pure product matrix and any other ingredients cannot be obtained (or prepared) it may be acceptable to add a known quantity of the analyte to the drug product and analysing the difference. In this case, however, it might be difficult to test over

the whole range.

Alternatively, bias can be established by comparison of the results from the proposed method with those produced by an alternative validated method. A validated method would either be one taken from a current recognised pharmacopoeia (BP, EP or USP) or another well-characterised method, the accuracy of which is stated and/or defined. As a rule, comparison against a validated method is not the preferred method, due to increased sources of error.

According to the ICH guidelines, accuracy can also be inferred once precision, linearity and specificity have been established, however, this is only likely to be the case for ingredients or simple preparations such as single ingredient solutions. As these are probably the easiest and least time consuming to test, this approach it is not recommended and a test for bias should be carried out.

The essence of bias testing, is the comparison of the reported value with the reference value to determine if there is a significant difference. It must be recognised that statistical and practical significance are not necessarily the same. Determination of statistical significance will depend on the number of test results and their standard deviation. Thus, if a large number of replicates are carried out and there is a high level of precision (ie. small standard deviation), then even a small bias will be statistically significant. Conversely, if there are only a small number of replicates and the standard deviation is high, then even a large bias would not be judged statistically significant. To prove statistically whether there was a difference equating to a bias of 1*SD would require 16 replicates and of 0.5*SD would require 55 samples. It can be seen from the above that, almost paradoxically, the more precise a method, the more likely it is that it will show a statistically significant bias. Thus a straightforward test for a significant difference between the observed and the reference value is not, on its own, considered appropriate. The main risks are that, where the precision is low, a bias which is in practice unacceptable, would be deemed statistically insignificant or, where the precision is high, a bias which in practice is acceptable, is considered statistically significant. The approach to be adopted is that an acceptable level of bias appropriate for the usage of the method is defined and testing carried out to determine both the statistical and practical significance of the bias in light of this limit.

As a general principle, the following limits on bias are recommended.

	At nominal concentration	At extremes of range
Substance	+/- 0.5%	+/- 1.0%
Product (simple matrix and/or critical use)	+/- 1.0%	+/- 2.0%
Product (simple matrix and/or non-critical use)	+/- 2.0%	+/- 3.0%
Product (complex matrix and critical use)	+/- 3.0%	+/- 5.0%
Product (complex matrix and non-critical use)	+/- 5.0%	+/- 5.0%

nb. A simple matrix would be, for example, a solution, powder, tablets/capsules.
A complex matrix would be, for example, creams, ointments or where extractions are needed.
Critical use would be, for example, injections, potent/toxic actives, fixed dosage administration (eg. tablets, suppositories).
Non-critical use would be, non-potent ingredients where the quantity administered is not fixed (eg. lotions, creams, ointments).

Testing:

For products, prepare homogeneous samples containing 50, 100 and 150% (+/- 10%, ie. 45-55%, 90-110% and 135-165%) of the nominal content of analyte in the full product formulation. Where the bias is determined by comparison with a known reference or ‘theoretical’ value, the samples must be prepared quantitatively so that the concentration is known. Where the bias is determined by comparison with a validated reference method, the content need not be accurately weighed and

recorded.

For substances, carry out the required determinations on accurately weighed quantities of reference material, equivalent to 50, 100 and 150% (+/- 10%, ie. 45-55%, 90-110% and 135-165%) of the nominal content.

Carry out the determinations at the extremes in triplicate under repeatability conditions, ie. same analyst, same equipment same day. Carry out nine determinations at the nominal concentration. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

Calculate the results in terms of percentage of the theoretical value.

For example, where the comparison is to be made against a known reference value, if the sample being tested is prepared by dissolving 'x' grams of reference standard with a certificated purity of 'p' percent, in sufficient solvent to give 250 ml, then the 'stated' content (as percent concentration) is given by the equation:

$$S = \frac{x * 100 * p}{250 * 100}$$

If the results for the measured content (again as percent concentration) are 'r' percent then the result as a percent of stated is given by the equation:

$$C = \frac{r * 100}{S}$$

Where the comparison is between two methods, 'x' and 'p' need not be known accurately provided the sample is homogeneous and used for both methods. In this case a nominal value for x can be used as the absolute values need not be known. Thus if the sample being tested is prepared by dissolving about 4 grams of ingredient, in sufficient solvent to give 250 ml, then the 'stated' content (as percent concentration) is given by the equation:

$$S = \frac{4 * 100}{250} = 1.6\%$$

If the results for the measured content (again as percent concentration) are 'r' percent then the result as a percent of stated is given by the equation:

$$C = \frac{r * 100}{S}$$

In this way, the results to be processed on the spreadsheet will always be approximately 100 which simplifies the setting of cell attributes.

For the tests carried out at the extremes of the range, calculate the mean of the results for the method under test and compare this with the theoretical value taken either from the stated content and weight taken to prepare the sample, or from the mean of the results determined by the reference method. The bias is given by subtracting the reference value from the observed value, thus a 'low' result would give a negative bias.

For tests carried out at the nominal concentration, calculate the mean, standard deviation, variance and standard deviation of the mean for the method under test and, where applicable, for the

reference method. Determine whether there is a statistically significant difference between the means using a t-test. Also calculate the maximum bias which would be statistically significant, the range of biases possible (95% confidence) for the method and reference and thus the maximum theoretical bias. These calculations can be performed by entering the results into the appropriate spreadsheet, for comparison with reference standard or comparison with reference method.

Acceptance Criteria:

At the extremes, the bias value is less than the bias limit required.

At the nominal content, interpret the results from the calculations or spreadsheet. There are a number of possible outcomes outlined below:

i. Bias is statistically insignificant and a bias which would be considered significant is less than required limit on bias

In this case we know that the bias is statistically insignificant and it must be within the limit we have set, therefore this would be acceptable.

ii. Bias is statistically insignificant and bias which would be considered significant is greater than required limit on bias.

In this case, although we know that the bias is statistically insignificant, the results are such that a bias greater than that considered acceptable would also be statistically insignificant. It is possible therefore that although the bias is statistically insignificant, it could be practically significant (ie. outside the required limit). It is therefore necessary to look in detail at the actual results obtained. If the actual bias is within the limit set, this would indicate that the bias is unlikely to be practically significant. However this value is based on the difference between the mean values for the test and reference samples and the true value will be within a range around this value. For this reason, a check is also made on the maximum bias value. This is difference between the extreme possible values for the means based on 95% confidence limits. If this value is below the required bias limit, the method's bias would be accepted as not practically significant. Additional replicates could be carried out to try to reduce the standard deviation.

iii. Bias is statistically significant and a bias which would be considered significant is less than required limit on bias.

Here, although the bias is statistically significant, it might not be practically significant. If the actual bias is within the limit set, this would indicate that the bias is unlikely to be practically significant. However this value is based on the difference between the mean values for the test and reference samples and the true value will be within a range around this value. For this reason, a check is also made on the maximum bias value. This is difference between the extreme possible values for the means based on 95% confidence limits. If this value is below the required bias limit, the method's bias would be accepted as not practically significant.

iv. Bias is statistically significant. A bias which would be considered significant is greater than required limit on bias.

In this instance, the bias would be deemed unacceptable. It might be worth considering testing additional replicates in order to reduce the standard deviation, which will have the effect of reducing the bias range which would be considered statistically significant.

<u>Ruggedness testing</u>

Principles:

Ruggedness testing or rubustness testing is intended to determine which experimental factors are important to the reliability of the method so that these can be defined and potential variability

controlled within the methodology. Ruggedness testing is therefore a set of experiments designed to show which factors have a significant effect on the performance of the method.

The need for ruggedness testing can be decided upon by knowledge of the physico/chemical reactions occurring, from experience of similar methods or as a consequence of precision testing results. For example, poor precision, in particular with respect to reproducibility or intermediate precision could point to differences in interpretation of instructions or end point.

Factors which might be expected to influence results might include, for wet chemistry, volumes of reagents, times allowed for reactions to complete, temperature of reaction, the effect of light and extraction times. For instrumental methods, for example HPLC, other factors such as pH of mobile phase, flow rate, composition of flow rate and brand of column might need consideration. For example if an analyte needs to be dissolved in 50 ml of solvent before titration, consideration might be needed as to whether adding slightly less solvent might affect results. This would be particularly relevant if the 50 ml was just sufficient to overcome solubility and there were also insoluble ingredients present, which would mean that dissolution could not be observed. Similarly if heating were required for a reaction or dissolution, then, if a lesser time were used the analyte might not be completely dissolved or the reaction might not go to completion. Conversely, heating for longer might result in decomposition of the entity being measured. Ruggedness testing involves running a series of experiments with the factors set at ‘high’ and ‘low’ extreme values.

Ruggedness testing can also be used to determine the effect of using different brands of reagent or equipment. One example of this might be HPLC where the brand of column might affect the results. In this case, instead of ‘high’ and ‘low’ values for the factor the two columns would be run as the alternative ‘values’.

By cunning experiment design, the number of tests can be kept to a minimum.

Testing:

Carry out testing in duplicate or triplicate on a homogeneous sample of the product or substance, prepared at its nominal concentration. The exact concentration need not be known.

It is recommended that, in all cases where there is more than one factor, which can affect the results, a series of experiments is set up according to one of the tables below. These are based on Plackett-Burman designs. The number of experiments (n) must always be a multiple of four and the number of factors which can be examined is one less then the number of experiments (n-1). For example if there were three factors being examined, four experiments would be required. If there were six factors, then eight experiments would be required to examine all combinations, but this would allow up to seven factors to be examined. If there are fewer factors than (n-1) then dummy factors would need to be inserted in the experiment design.

In designing the experiments, firstly the factors to be examined must be decided upon and then the alternative values assigned. In the above examples, if 50 ml of solvent is specified in the method, it might be decided to examine the effect of using less. Therefore a low value of 40 ml and a high value of 50 ml might be used. If 30 minutes heating are called for and it is considered possible for too little heating to result in incomplete dissolution and too much to result in decomposition, then a low value of 20 minutes might be used with a high value of 40 minutes. In the case of HPLC columns, the ‘low’ value would be one brand and the ‘high’ value an alternative brand. The experiments are carried out in accordance with the tables below, with the ‘low’ value represented by ‘-‘ and the ‘high’ value represented by ‘+’.

Table 1: Design for 3 or 11 factors

	Factors										
Experiment		1	2	3	4	5	6	7	8	9	10	11
	1	-	-	-	-	-	-	-	-	-	-	-
	2	+	-	+	-	-	-	+	+	+	-	+
	3	+	+	-	+	-	-	-	+	+	+	-
	4	-	+	+	-	+	-	-	-	+	+	+
	5	+	-	+	+	-	+	-	-	-	+	+
	6	+	+	-	+	+	-	+	-	-	-	+
	7	+	+	+	-	+	+	-	+	-	-	-
	8	-	+	+	+	-	+	+	-	+	-	-
	9	-	-	+	+	+	-	+	+	-	+	-
	10	-	-	-	+	+	+	-	+	+	-	+
	11	+	-	-	-	+	+	+	-	+	+	-
	12	-	+	-	-	-	+	+	+	-	+	+

Table 2: Design for 7 factors

	Factors							
Experiment		1	2	3	4	5	6	7
	1	+	+	+	+	+	+	+
	2	+	+	-	+	-	-	-
	3	+	-	+	-	+	-	-
	4	+	-	-	-	-	+	+
	5	-	+	+	-	-	+	-
	6	-	+	-	-	+	-	+
	7	-	-	+	+	-	-	+
	8	-	-	-	+	+	+	-

Thus considering a method with three factors which are to be examined, the top left portion of table 1 will be looked at. Experiment 1 will have all three factors at their '-' or 'low' value. Experiment 2 will be a repeat with factors 1 and 3 at their '+' or 'high' value. Experiment 3 will have factors 1 and 2 at their '+' or 'high' values and experiment 4 will have factors 2 and 3 at their '+' or 'high' value. It can therefore be seen that by combining the results of the different experiments in different ways, each factor can be considered whilst cancelling out the effects of the other factors. Thus, if the results of experiments 1 and 4 are averaged, this will give an average result with factor 1 at its '-' level. As both factors 2 and 3 are present once at the '-' and once at the '+' level, any effect due to these factors is cancelled. Similarly averaging experiments 2 and 3 will give an average result with factor 1 at its '+' level, whilst the effects of the other factors cancel. The difference between these two averages is the effect Factor 1 has on the result and this can be tested for significance. To look at the effect of factor 2 experiments 1 and 2 are averaged for the 'high' result and experiments 3 and 4 are averaged for the 'low' result. Similarly, to look at the effect of factor 3 experiments 1 and 3 are averaged as are experiments 2 and 4.

Calculate the results for each experiment in terms of percentage of stated. For each factor, calculate the average results for the experiments where it is at the 'low' value and the average for the experiments where it is at the 'high' value. Determine the difference and examine to see if the difference is significant using a t-test. The standard deviation from precision testing is used in the calculations. The calculations can be performed by entering the results into the appropriate spreadsheet.

Acceptance criteria:

There is no significant difference between the results obtained with alternative values for each factor.

ABRIDGED VALIDATION

Where a method needs to be developed on a one-off or ad hoc basis, it is generally only appropriate to carry out minimal validation. This is due to pressure of time and limited resources. Despite this, it is still necessary to work to a logically reasoned and documented plan appropriate to the needs of the method.

A distinction needs to be drawn between the true ad hoc request and a method to be used in the short term. It is recognised that a lot of ad hoc requests do not justify a lengthy validation exercise.

Methods used in the short term in one laboratory might include stability trials, tests on products used for a short period (eg. for a clinical trial) or test required for a specific project. These methods should be subjected to a full validation procedure to ensure fitness for purpose. This is partly to ensure that the method can be registered as valid for future use or use elsewhere and partly to underpin the study being carried out. In drawing up the validation plan however, due consideration can be given to likelihood of further use of the method and this might influence, for example, the need for some of the precision studies.

This section is therefore intended to give guidance on the truly ad hoc methods. These tests will often be in response to complaints or queries such as ‘what is this white powder?’ or ‘do these tablets contain paracetamol?’. Although these are ad hoc requests, the consequences of reporting erroneous results are potentially serious. It is important therefore to be able to demonstrate fitness for purpose.

Ad hoc methods would not generally be validated by comparison with a reference method as, where this is available, it would generally be unnecessary to develop an alternative. Therefore this is not discussed. If a reference method is to be used, the section on full validation should be referred to.

It is not always possible to obtain details of or be able to source all excipients or matrix components. Where this is the case, it may only be possible to use an approximation of the matrix and additional methods of demonstrating fitness for purpose might need to be devised.

1. Identification

Specificity and Sensitivity

Testing:

Apply the set of identity tests to a sample comprising the analyte of interest (in a suitable matrix if available and applicable), and, where available and applicable, a sample comprising the matrix containing, where possible, any other known ingredients or excipients or a sample consisting of the potential interferents.

Acceptance results table:

Sample	Results			
Matrix (if applicable) with interferents	All tests -ve	Some tests +ve	All tests –ve	All tests +ve
Analyte (in matrix)	All tests +ve	All tests +ve	Some tests +ve	All tests +ve
Decision	Accept	Accept with caution	Accept with caution	Reject

In addition, if chromatography is used as an identity test, a sample containing the analyte of interest and any other possible compounds present should be run. Where appropriate, a resolution factor of >=1.5 between the analyte of interest and other compounds is acceptable (ie. positive for the analyte of interest). If possible, a peak purity test can also be carried out on the analyte peak.

2. Qualitative impurity testing (limit tests)

Specificity

Testing:

Where the impurity is known and available, this should be used for spiking samples. Where the impurity is not known or is unavailable as will sometimes be the case with degradation impurities, samples should be stored under suitable stress conditions eg. light, heat, humidity, acid/base hydrolysis and oxidation. Apply the limit test being validated to a sample spiked with the limit concentration of the analyte or degraded as above, an unspiked sample comprising the active ingredient (in the appropriate matrix if applicable), and, where applicable and if possible, a sample comprising the matrix.

Acceptance criteria table:

Sample	Results			
Analyte (in sample)	Test +ve	Test +ve	Test –ve	Test +ve
Unspiked sample	Test –ve	Test +ve	Test –ve	Test +ve
Matrix (if applicable)	Test –ve	Test –ve	Test –ve	Test +ve
Decision	Accept	Reject	Reject	Reject

If chromatography is used, the spiked sample should show a resolution factor of >=1.5 between the analyte and any other peak. The unspiked sample should show nothing which will interfere with detection of the analyte, however a small peak might be observed if some of the analyte is present as an impurity. Where this is the case, the spectrum and purity of the peak can be examined if possible to confirm it is due to the analyte. The matrix, where tested, should show no peak corresponding to the analyte peak.

Detection limit

Testing:

Where the test involves comparison of the test with a standard and the impurity is known and available, a standard containing the impurity at its limit value in a suitable medium should be prepared. Test samples consisting of the product in its matrix or the substance spiked with the analyte at concentrations equivalent to 50%, 100% and 150% of the limit value should be prepared. Where the matrix is not available, a similar one should be chosen so that any extraction steps can be reproduced. The test samples should be subjected to a blind limit test in duplicate. Where the test is judged visually, the results should be recorded as ‘pass’, ‘fail’ or ‘no difference’. If the result is measured instrumentally, a test reading greater than 10% different from the standard is recorded as a ‘pass’ or ‘fail’ whereas readings differing by no more than 10% are recorded as ‘no difference’.

Acceptance criteria:

For acceptance, all tests at the 50% level should be reported as ‘Pass’ and all results at the 150% level as ‘Fail’. The results at 100% should be a mix of predominantly (ie. at least 50%) ‘no difference’ with some ‘pass’ and/or ‘fail’.

Where the test involves comparison of secondary peaks to main peaks, test a sample consisting of the product (in its matrix) or the substance spiked with the analyte (impurity) at a concentration equivalent to its limit value. This might need to be done by stressing samples and by trial and error. Carry out the limit test on the sample in triplicate and calculate the area of the secondary peak relative to the main peak. Calculate the %CV.

Acceptance criteria:

It is generally acceptable for the %CV to be below 10%.

3. Quantitative Impurity Testing

Specificity

Testing:

Perform the impurity test, in duplicate, on a sample comprising the nominal limit concentration of the analyte (impurity) in the appropriate matrix if applicable, a sample consisting of the product or substance spiked with the same nominal limit concentration of the analyte and, where applicable and available, a sample of unspiked product or substance. Calculate the results relative to the analyte sample. It is recommended that the matrix should affect results by no more than 2% and other interferents by no more than 5%.

Acceptance results table:

Sample	Reported Impurity content R (% impurity sample result)			
Analyte (in matrix)	100%	100%	100%	100%
Sample spiked with analyte	95%<R<105%	95%<R<105%	R > 105%	R < 95%
Unspiked substance or product	R < 2%	R > 2%	R < 2%	R < 2%
Decision	Accept	Reject – Consider using matrix as a blank	Reject – Positive interference from active	Reject – Negative interference from active

In addition, if chromatography is used, the spiked sample should show a resolution factor of >=1.5 between the analyte and any other peak. The unspiked sample should show nothing which will interfere with detection of the analyte, however a small peak might be observed if some of the analyte is present as an impurity. Where this is the case, the spectrum and purity of the peak can be examined if possible to confirm it is due to the analyte.

Precision

Testing:

Prepare homogeneous samples containing 50, 100 and 150% (+/- 10%, ie. 45-55%, 90-110% and 135-165%) of the limit concentration in the product matrix or substance. Where the testing is also to be used to determine bias, the content will need to be accurately weighed and recorded. Carry out the determination at each concentration in triplicate under repeatability conditions, ie. same analyst, same equipment, same day. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

Calculate the standard deviation and percentage coefficient of variation (%CV) at each concentration, for the replicates carried out under repeatability conditions (see section on statistical evaluation). For ad hoc testing, a %CV of 10% is generally acceptable.

Acceptance criteria table:

Tests	Comparison of %CV with specified limit	
Repeatability	Below Limit	Above limit at Low or High concentration
Decision	Accept	Reject – Consider adjusting range or changing sample size or dilutions

Bias and Linearity

Testing:

Prepare homogeneous samples containing 50, 100 and 150% (+/- 10%, ie. 72-88%, 90-110% and 118-132%) of the limit concentration of the analyte in the product or substance as appropriate. The samples must be prepared quantitatively so that the concentration of analyte is known.

Carry out the determinations in triplicate under repeatability conditions, ie. same analyst, same equipment, same day. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

Calculate the results in terms of percentage of stated.

For example, if the sample being tested is prepared by dissolving 'x' grams of reference standard with a certificated purity of 'p' percent, in sufficient solvent to give 250 ml, then the 'stated' content (as percent concentration) is given by the equation:

$$S = \frac{x * 100 * p}{250 * 100}$$

If the results for the measured content (again as percent concentration) are 'r' percent then the result as a percent of stated is given by the equation:

$$C = \frac{r * 100}{S}$$

Calculate the mean of the results for the method under test at each concentration. Determine the bias at each concentration from the mean of the results and the theoretical value, taken from the stated content and weight taken to prepare the sample, by subtraction.

Acceptance Criteria:

The bias value is less than the bias limit required at all concentrations. For ad hoc testing, a bias limit of 10% is recommended.

4. Assay

Specificity

Principles:

Specificity can be particularly difficult to demonstrate unequivocally with ad hoc methods due to difficulties in obtaining possible interferents. It is therefore important to apply a systematic

approach to build up as much data as is practicable within the normal constraints of time and resources. A number of mechanisms of interference might be occurring and it is necessary to consider the potential for these. Typically, interferents might have an additive effect by reacting similarly (eg. absorb UV at an overlapping wavelength), they might remove the analyte from solution (eg. by sorption), thereby reducing the measured effect or they might in some other way inhibit or enhance the measured effect of the analyte.

Testing:

Where the interferents are known and available, perform the assay, in duplicate, on a sample comprising the analyte (in the appropriate matrix if applicable), samples comprising analyte (in the appropriate matrix if applicable) spiked with an appropriate level of known or expected interferents and, where applicable, a sample comprising the product matrix. Calculate the results with reference to the pure analyte sample results. It is recommended that the matrix should affect results by no more than 2% and other interferents by no more than 5%.

Acceptance results table:

Sample	Results (% pure sample result)			
Analyte (in matrix)	100%	100%	100%	100%
Sample spiked with interferants	95%<R<105%	95%<R<105%	R>105%	R<95%
Matrix (if applicable)	R < 2%	R > 2%	R < 2%	R < 2%
Decision	Accept	Reject - Consider using matrix as a blank	Reject – Positive interference	Reject – Negative interference

In addition, if chromatography is used, the spiked sample should show a resolution factor of >=1.5 between the analyte and any other peak. The matrix, where tested, should show no peak corresponding to the analyte peak.

Where the interferents are unknown or are not available it will be necessary to use a combination of tests to build up evidence of specificity. Some possible tests are detailed below, however these are only examples and a protocol needs to be devised appropriate to the method.

If some of the excipients are available, test as above with whichever excipients are available and use knowledge of the other compounds concerned to predict whether interference will occur, for example does the excipient contain a chromophore, which might interfere with a spectrophotometric method?

If the method is chromatographic, then in addition to resolution it might be possible to perform a peak purity test. Where necessary, perform a peak purity test on a solution of pure standard to assist in setting the test thresholds.

Where it is possible that the interferents could be removing the analyte from solution, a further extraction/dissolution can be carried out on any residues to check for retention of analyte.

An additional check for interference can be carried out by adding a known quantity of analyte to the full product formulation at its nominal concentration. The increase in response can then be compared with that expected for the known addition. Care must be taken in performing the calculations, particularly where the samples are solids and mixed by weight.

Precision

Testing:

Prepare homogeneous samples containing 50, 100 and 150% (+/- 5%, ie. 45-55%, 90-110% and 135-165%) of the nominal concentration in the full product formulation or matrix. Where the full product formulation or matrix cannot be replicated, use a formulation as close as possible to the product matrix. Where the testing is also to be used to determine bias, the content may need to be accurately weighed and recorded. Carry out the determination at each concentration in triplicate under repeatability conditions, ie. same analyst, same equipment same time. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed in the normal way to provide the results for each of the triplicate tests.

Calculate the standard deviation and percentage coefficient of variation (%CV) at each concentration, for the replicates carried out under repeatability conditions (see section on statistical evaluation). A limit for the %CV of 5% is recommended for ad hoc assay methods.

Acceptance criteria table:

Tests	Comparison of %CV with specified limit	
Repeatability	Below Limit	Above limit at Low or High concentration
Decision	Accept	Reject – Consider adjusting range or changing sample size or dilutions

Bias and Linearity

Principles:

Ideally testing should be carried out on samples containing the analyte in the full product formulation and across the range. Where this is not possible, an additional indication of possible bias can be achieved by testing more than one batch of sample where these are available. Thus it might be reasonable to assume that batches would be randomly distributed about the mean and that if testing a number of batches gave results which were all either low or high, this might be indicative of bias.

Testing:

Prepare homogeneous samples containing 50, 100 and 150% (+/- 10%, ie. 45-55%, 90-110% and 135-165%) of the nominal content of the analyte in the full product formulation or matrix if available. Where the full product formulation or matrix cannot be replicated, use a formulation as close as possible to the product matrix. For tests on substances take 50, 100 and 150% of the quantity specified in the method. The samples must be prepared quantitatively so that the quantity is known.

Carry out the determinations in triplicate under repeatability conditions, ie. same analyst, same equipment, same time. It should be noted that the whole process should be replicated, thus if a method involves weighing, extracting, evaporating, dissolving, diluting and then carrying out HPLC, the whole process from weighing to chromatography must be carried out in triplicate. If the method itself includes replicates these should be carried out in accordance with the laboratories SOP's to provide a single result for each replicate. For example, if HPLC methods are normally carried out by performing quadruple injections of each sample, this should be done and processed

in the normal way to provide the results for each of the triplicate tests.

Calculate the results in terms of percentage of stated.

For example, if the sample being tested is prepared by dissolving ‘x’ grams of reference standard with a certificated purity of ‘p’ percent, in sufficient solvent to give 250 ml, then the ‘stated’ content (as percent concentration) is given by the equation:

$$S = \frac{x * 100 * p}{250 * 100}$$

If the results for the measured content (again as percent concentration) are ‘r’ percent then the result as a percent of stated is given by the equation:

$$C = \frac{r * 100}{S}$$

Calculate the mean of the results for the method under test at each concentration. Determine the bias at each concentration from the mean of the results and the theoretical value, taken from the stated content and weight taken to prepare the sample, by subtraction.

Acceptance Criteria:

The bias value is less than the bias limit required at all concentrations. For ad hoc testing, a bias limit of 5% is recommended.

Where the full product formulation or matrix cannot be replicated, if available, test at least two batches of material in the full formulation. Calculate the results as a percentage of the nominal content.

Acceptance Criteria:

If all results are either less than 95% or greater than 105% it is presumed that there is a possibility of negative or positive bias respectively and this needs further investigation.

VERIFICATION

Where a method, which has been validated in one laboratory, needs to be put into use in a different laboratory, a certain degree of local verification may be necessary. Although verification will not usually be as detailed as validation, the same process and principles apply. The level of verification required will depend on the degree of difference between the original application of the method and its application in the second laboratory.

A verification plan should be drawn up along the same lines as a validation plan, in order to document the process involved in deciding what verification tests to perform. Each performance characteristic should be considered and the likelihood of the method being influenced by local factors assessed. Where it is considered that local factors might affect the performance characteristic, similar steps should be taken as for method validation.

Reference to the section on validation should be referred to for details of sample preparation, testing and data processing.

Specificity

Where the method is to be applied to a substance or product, which is identical and in an identical matrix to that where the method was validated, specificity need not generally be verified. An exception might be where there is potential locally for a different interferent to be present through cross contamination. In this case, depending on the use to which the method is being put, it might be

necessary to verify the specificity with that interferent.

Where the product is different and contains different or additional potential interferents, for example the matrix contains different preservatives, it may be necessary to repeat the specificity part of the validation. In taking this decision, consideration of the degree of change is taken into account. For example, a change of preservative from methyl hydroxybenzoate to propyl hydroxybenzoate would probably not be considered significant whereas a change from a hydroxybenzoate to bronopol probably would be significant.

Linearity

Instrument linearity generally need not be verified locally provided the range is the same and where appropriate instrument parameters are comparable. In deciding this, it is important to ensure that it is the range of the instrumental measurement, which needs to be the same rather than that of the substance or product, therefore by appropriate sample preparation, products of different strength can be analysed without verification.

With methods such as chromatography, it is important to ensure that the range quoted includes consideration of the injection volume. Thus it is important to ensure that the quantity of analyte presented to the chromatograph is within the same range as the validated method.

Precision

Where the method is to be applied to a substance or product, which is identical and in an identical matrix to that where the method was validated, precision need not generally be verified.

Where there are differences, some precision testing needs to be done and the amount needs to be decided upon. An assessment needs to be made as to whether the application is likely to be more or less challenging to the method. Consideration needs to be given to such factors as product matrix and sample preparation. Where the matrix is simple, eg. a solution and sample preparation only requires dilution, precision is unlikely to be adversely affected and minimal verification is needed. Where sample preparation is more complex, for example where extractions are required or samples need to be heated or reacted, precision is more likely to be affected and more intensive verification is required.

Where minimal verification is required, this should be done by testing a single homogeneous sample containing the analyte at its limit or nominal concentration (dependant on the application of the method) in triplicate under repeatability conditions.

Where intensive verification is required, the intermediate precision should be determined by testing a single homogeneous sample containing the analyte at its limit or nominal concentration (dependant on the application of the method) in triplicate on two different days with different analysts and equipment (where possible).

The results can be processed in the same way as under the appropriate validation section and similar acceptance criteria and limits applied.

Bias

Where the method is to be applied to a substance or product, which is identical and in an identical matrix to that where the method was validated, bias need not generally be verified. An exception might be where the method requires equipment calibration. An example of this would be absorption spectroscopy. Where the method involves comparison with a standard, bias testing would not normally be done, however, if the method involves the use of an A(1%, 1 cm) a minimal check on bias should be carried out.

Where there are differences in the product, especially if the matrix differs or there are different excipients, a more intensive bias testing needs to be done.

Where minimal bias verification is required, this should be done by testing a single homogeneous sample containing the analyte at its limit or nominal concentration (dependant on the application of the method) in triplicate under repeatability conditions.

Where more intensive bias verification is required, this should be done by testing homogeneous samples containing the analyte at 50%, 100% and 150% of its nominal concentration for assays and 10%, 100% and 200% of its limit concentration for stability impurity methods and 50%, 100% and 200% of its limit concentration for limit impurity testing. All tests are carried out under repeatability conditions.

The results can be processed in the same way as under the appropriate validation section and similar acceptance criteria and limits applied.

System suitability testing

Some techniques can be more dependent on the specific equipment used than others. An example of this is chromatography. Differences in equipment, manufacturer of column methods of detection can all affect the efficiency of the process and separation. Columns also deteriorate with usage and repeat testing might be needed periodically. Where system suitability testing is required, this should be specified with the method. Where this is the case, system suitability should be carried out as part of local verification.

STATISTICAL EVALUATION OF RESULTS

Attempts have been made to keep the statistical analysis of results to a minimum and limit calculations to those needed to demonstrate fitness for purpose. In addition, electronic versions of the document contain a number of spreadsheets, which can be used to carry out the necessary calculations. These are supplied in excel file format. This section describes either the standard statistical calculations which might be applied or the statistics used in these spreadsheets. Other statistical tests might be appropriate in certain circumstances.

In applying statistical analysis to the results of analytical testing, a certain degree of caution must be exercised. Thus statistical significance and practical significance might not be the same. A confidence of 95% is used throughout.

Precision

Spreadsheets are provided for estimation of repeatability, intermediate precision and reproducibility.

The spreadsheet for repeatability can be used to process up to five replicates at each of three concentrations, representing the extremes of the usable range and the nominal concentration. It calculates the mean, standard deviation, relative standard deviation and % coefficient of variation by standard statistical methods.

The spreadsheet for intermediate precision and reproducibility can be used to process three replicates of the same sample, tested on three occasions. This relates to two sets of testing in the originating laboratory to provide data for intermediate precision and a further set of testing by an independent laboratory for a measure of reproducibility. The data are assessed by means of standard Analysis of Variance (ANOVA) calculations. The ANOVA calculations perform an F test to determine whether there is a significant difference between the sets of results produced on different days of by different laboratories. Intermediate precision is determined by consideration of the results from Laboratory 1 only and reproducibility is determined by consideration of all three sets of results. If the ANOVA shows there is no significant difference between the sets of results, the repeatability standard deviation is the standard deviation of all results and the intermediate precision or reproducibility is equal to the repeatability standard deviation. If there is a significant difference, the repeatability standard deviation is given by the square root of the within group mean square and the intermediate precision or reproducibility is given by the square root of the sum of the repeatability standard deviation squared and the between group standard deviation squared.

Linearity

A spreadsheet is provided for linearity testing of quantitative methods. Four columns are provided for the entry of results. The first column allows for the entry of a concentration term. Where a series of standards are prepared from a single concentrated standard solution, it is generally easier to enter the volume used. The remaining columns allow for the entry of triplicate series of results.

The spreadsheet then calculates the Average, Standard Deviation and % Coefficient of Variation (%CV) of the replicates at each concentration. From these results, the intercept, slope and correlation coefficient are calculated by the method of least squares regression. The 'theoretical' response at each 'concentration' is then calculated. From these data, the residuals for each replicate are calculated as y(obs)-y(calc).

A graph of Response against the concentration term is plotted together with a plot based on the calculated best straight line. This is checked visually for linearity.

A scatter graph of the residuals against the concentration term is plotted. Typically this plot will show the points scattered almost randomly around the zero. Where this is not the case, it is often possible to draw useful conclusions from the distribution. For example, a curved response might be represented by residuals being predominantly below zero at low concentration, above zero at mid concentrations and below zero at higher concentrations. Residuals might show a spread, which gradually increases with increasing (or reducing) concentration. This would be indicative of a reduced precision at one end of the range and might necessitate an alteration to the acceptable range for the method. An incorrect zero intercept might be represented by residuals gradually decreasing (positive to negative) as the concentration term increases.

A further plot, of the %CV against the concentration term is given. This might also give an indication of the acceptable range of the method.

It must be remembered that the linearity tests discussed here and processed by this spreadsheet, relate only to instrument linearity. For methods other than those used to analyse either starting materials or drug products with a simple product matrix (eg. single component solutions) the variance for the whole method is inevitably going to be greater than that determined during linearity tests. Therefore if the variance from this test is used as an indication of the limit of quantitation, this should be taken into account. It is recommended that, where this is the case, the acceptable range is presumed to be between the points at which the variance determined during linearity tests is half the maximum specified. The range is confirmed by the precision tests.

Bias

Bias testing is carried out by comparing the results from the method with either a known reference value or with the results from a second reference method.

At the extremes of the range, the mean of the test is calculated and where comparison is with a reference method, the mean of the results from the reference method is also calculated. Where the bias is determined by comparison with a known value, the reference value is calculated from the weight and purity of the reference material used. The mean result is then calculated in terms of a percentage of the reference value. The bias is the difference between this value and 100%.

Two spreadsheets are supplied for processing the data for determining the bias at the nominal, or, in the case of impurity testing, the limit concentration.

The spreadsheet for determining bias by analysing a sample of known concentration prepared from a reference material, allows for the entry of up to 15 test results and the reference value together with its uncertainty. Again the test results are calculated and expressed in terms of percentage of the reference value. Therefore the reference value entered will generally be 100.0. The number of entries for the reference is one and the mean is the same as the entered value. From the results, the mean, standard deviation and variance are calculated by normal statistical methods. The standard deviation of the mean

is also calculated as the standard deviation divided by the square root of the number of observations. From the reference value and its uncertainty, the extreme values for the reference are calculated. A t-test is used to determine if there is a statistical difference between the reported mean and the theoretical value. The t statistic is calculated as the difference between the observed and theoretical means divided by the standard deviation of the mean for the test results. The critical value of t is for a two tailed test. If t_{stat} is less than t_{crit} there is no significant difference between the results and the reference value. If t_{stat} is greater than t_{crit} there is a significant difference between the results and the reference value.

The spreadsheet for determining bias by comparing results from the method being validated with those from a reference method, allows for the entry of up to 15 results for both methods. The results should all be expressed in terms of the percentage of the reference method mean. Therefore the reference method mean should calculate out as about 100.0. From the results, the mean, standard deviation and variance are calculated for both methods by normal statistical methods. The standard deviations of the means are also calculated as the standard deviation divided by the square root of the number of observations. An F-test is carried out to determine if the variances of the two methods are significantly different. The F statistic is calculated as the ratio of the variances, using the larger as the numerator to ensure a result of greater than 1. The critical value of F is for a two tailed test. If F_{stat} is less than F_{crit} there is no significant difference between the variances for the two methods. If F_{stat} is greater than F_{crit} there is a significant difference between the variances for the two methods. If there is no significant difference between the variances of the two methods, the t statistic for comparing the means is calculated from the pooled data. If there is a significant difference between the variances, pooled data cannot be used for calculation of the t statistic. These calculations are performed by the spreadsheet, and the appropriate values for t_{stat} and t_{crit} automatically entered. If t_{stat} is greater than t_{crit} there is a significant difference between the results obtained by the two methods.

To assist in deciding whether any bias is practically significant, additional values are calculated. The maximum value of bias, which would result in a significant difference is given. The reported bias is given. The range of the actual means give the minimum and maximum values possible taking into account the reported means and standard deviation or uncertainty as appropriate. The maximum bias is calculated as the difference between the maximum reference value and minimum test value or the minimum reference value and maximum test value as appropriate.

Ruggedness

A number of experiments are run to look at the effect of factors which might affect the analytical result. The experiments are arranged so that by selecting half the experiments, one factor is at its ‘low’ value and in the other half the factor is at its ‘high’ value. In each half, all other factors are represented equally as ‘low’ and ‘high’. Therefore, the difference between the averages from each half is calculated and represents the effect of that factor. By considering different combinations of experiments this can be repeated for each factor in turn. A t-test is carried out to determine if there is a significant difference between the ‘low’ and ‘high’ results for each factor. The standard deviation (in terms of the %CV) from precision testing is used for the t-test.

A spreadsheet is provided for processing the data for ruggedness testing. The %CV from precision testing needs to be entered together with the number of results it was based on (generally 3, 6 or 9). The number of experiments entered must be 4, 8 or 12 and duplicate or triplicate results for all experiments must be entered as a percentage of stated. The spreadsheet calculates the averages for each factor at its ‘low’ and ‘high’ value and finds the difference. These differences are tested for significance by the t-test.

DEFINITIONS

A number of terms are used in the context of method validation. Some of these have a generally understood meaning, which might differ from their strict meaning in the context of chemical analysis. In addition, sometimes the reference sources carry slightly different definitions for a term. The definitions given here relate to the terms as they are used in this document. Where appropriate the source is referenced although it may not be lifted verbatim. The definitions given here have not been taken from a single source but have been chosen to aid the understanding of the subject and interpretation of this guidance.

Bias

Bias is a measure of 'trueness' which is defined in ISO 3534 as the closeness of agreement between the average value obtained from a large set of test results and an accepted reference value. The measure of trueness is normally expressed in terms of bias.

Under current ISO definitions, accuracy is a property of a result and comprises bias and precision.

Fitness for purpose

Fitness for purpose is a difficult concept to define, however in the field of Pharmaceutical Quality Control, a method may be considered fit for purpose if its results can be relied upon in reaching a decision as to whether or not to pass a product for patient use.

This might seem a simplistic definition, however implicit in it are all the necessary concepts of method validation including accuracy, precision, reproducibility and linearity of response as well as the avoidance of false positive or negative results.

Limit of detection

Detection Limit is defined in the ICH guidelines on validation of analytical procedures as 'the lowest amount of analyte in a sample that can be detected but not necessarily quantitated as an exact value'.

Limit of Quantitation

The Limit of Quantitation is defined by the ICH Guideline as 'the lowest amount of analyte in a sample which can be quantitatively determined with suitable precision and accuracy'. The quantitation limit is particularly relevant in very low level assays, such as those for impurities or degradation products.

Linearity

According to the Peer-Verified Methods Committee (PVMC) of the Association of Official Analytical Chemists (AOAC) linearity 'Defines the ability of the method to obtain test results proportional to the concentration of the analyte'.

Measurement Uncertainty

According to ISO 1993, 2nd ed 1995 Measurement Uncertainty is a parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand. The crucial factor is that the uncertainty value combines all uncertainties, whether arising from random variation or systematic effects. Some of these might be laboratory or product specific (eg. equipment, ingredient or reagent uncertainties), therefore determination of uncertainty for the purposes of this guidance is inappropriate.

Precision

ISO 3534-1:1993 defines precision as the closeness of agreement between independent test results obtained under stipulated conditions.

There are three commonly used terms relating to precision estimates:

Repeatability – this refers to tests performed under conditions which are as constant as possible, eg. short time interval, same laboratory, same analyst, same equipment. Although this is of limited value, since when a method is put into use, these conditions are unlikely to be met, it is often carried out during the early stages of method development. A method which does not perform satisfactorily under repeatability conditions will be unlikely to perform satisfactorily in other precision tests.

Intermediate precision – This refers to within lab variation, eg. same laboratory but different days, analysts and equipment (if available). This may be referred to as in-lab reproducibility. This measure is the most useful measure for methods for in-house use.

Reproducibility – This refers to tests carried out over a period of time, using different analysts in different laboratories with different equipment. This measure is of value where methods need to be transferable between laboratories.

Range

The range is defined in the ICH guidelines as 'the interval between the upper and lower concentration (amounts) of analyte in the sample (including these concentrations) for which it has been demonstrated that the analytical procedure has a suitable level of precision, bias and linearity.

Robustness

Robustness is defined in the ICH guidelines on validation of analytical procedures as 'a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters'. It provides an indication of the method's reliability during normal usage. Ruggedness testing is used to identify the experimental parameters which have a significant effect on performance of the method. These parameters need to be controlled if the method is to perform satisfactorily in routine use and is to be transportable between laboratories.

Specificity

Specificity is defined in the ICH guidelines as 'the ability to assess unequivocally the analyte in the presence of components which may be expected to be present. Typically these might include impurities, degradants, matrix etc'. Assurance that a method is specific ensures that the analyte can be identified, that an accurate assessment of impurity content (related substances) can be made and that the content or potency of the analyte in a sample can be accurately measured.

Validation

According to ISO 9000:2000 validation is confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled.

SUMMARY OF PERFORMANCE CHARACTARISTICS REQUIRED TO BE EVALUATED

PERFORMANCE CHARACTERISTIC	IDENTITY	QUALITATIVE IMPURITY	QUANTITATIVE IMPURITY	ASSAY
Specificity	Test for no interference from related substances, matrix and other ingredients	Test for no interference from the active ingredient, formulation and matrix.	Test for no interference from the active ingredient, formulation and matrix.	Unless controlled by other means eg. impurity testing, test for no interference from interferants, excipients and matrix.
Limit of Detection	N	Carry out replicate tests (6) at 80%, 100% and 120% of the limit concentration.	See tests for precision	N
Limit of Quantitation	N	N	See tests for precision	N
Linearity	N	N - But implicit in Limit of Detection test.	For instrumental methods determine linearity of response on 6 concentrations prepared in triplicate. For non-instrumental methods refer to bias testing at extremes of range.	For instrumental methods determine linearity of response on 6 concentrations prepared in triplicate. For non-instrumental methods refer to bias testing at extremes of range.
Precision 1 –Repeatability	N	Included in Limit of Detection	Test in triplicate at low, nominal and high concentration under repeatability conditions.	Test in triplicate at low, nominal and high concentration under repeatability conditions.
Precision 2 – Intermediate Precision	N	Included in Limit of Detection	Test in triplicate, the nominal concentration from above, using different analyst, day and instrument if possible.	Test in triplicate, the nominal concentration from above, using different analyst, day and instrument if possible.
Precision 3 – Reproducibility	N	N	Not normally required for in-house stability methods. If required, test in triplicate, the nominal concentration from above, using different laboratory.	Test in triplicate, the nominal concentration from above, using different laboratory.
Bias	N	N	Carry out replicate testing; (3) at the extremes of the range and (9) at the nominal concentration.	Carry out replicate testing; (3) at the extremes of the range and (9) at the nominal concentration.
Ruggedness testing	N	N	Dependent on results for Precision 3	Dependent on results for Precision 3

N = Not required

VALIDATION PLAN

ASSAY OF SODIUM CHLORIDE IN COMPOUND SODIUM CHLORIDE MOUTHWASH

Written by ……. Date……..

Approved by…….. Version…….

PRODUCT FORMULATION	
Sodium Bicarbonate	10 g
Sodium Chloride	15 g
Concentrated Peppermint Emulsion	25 ml
Double-strength Chloroform Water	500 ml
Water Sufficient to produce	1000 ml

INTRODUCTION	
Method:	Back titration of excess silver nitrate with ammonium thiocyanate using ammonium iron III sulphate as indicator.
Background:	The B.P. method for assay of Sodium Chloride in Compound Sodium Chloride Mouthwash is by the potentiometric method. This method has proved unreliable and time consuming in practice and the back-titration method used in the B.P. for assay of Sodium Chloride in tablets is to be introduced subject to validation.
Methodology:	Accurately pipette 10mls of sample into a 250ml conical flask. Add 15ml of 2M Nitric Acid (measuring cylinder) and 5ml of dibutyl phthalate. Add by pipette 50mls of 0.1M Silver Nitrate vs and shake vigorously for 1 minute. Add 5mls of Ammonium Iron III sulphate solution R2, and titrate with 0.1M ammonium thiosulphate vs until a reddish brown colour is obtained, which after shaking does not fade within 5 minutes. Each ml of silver nitrate (0.1M) ≡ 5.844mg of Sodium Chloride
Other considerations:	Consider ruggedness testing, if intermediate precision or reproducibility tests are outside the acceptance criteria. Parameters which could affect the result and should be considered are: time of initial shaking and formation of precipitate (1 minute), volume of Ammonium Iron III sulphate solution added (5 ml), volume of Nitric Acid added (15 ml) and time allowed to look for fading of colour (5 minutes).

WORKED EXAMPLE 1 - ASSAY

ANALYTICAL REQUIREMENT		
Analyte:	Sodium Chloride – Determined as Chloride. Nominal concentration is 1.5% sodium chloride. The method is to be tested at 50 and 150% of the nominal concentration to allow a usable range of 60 to 140% of nominal.	
Matrix:	Water	
Potential Interferants:	Sodium Bicarbonate	Considered as interferent A
	Peppermint Emulsion Double Strength Chloroform Water	Considered jointly as excipients B

PERFORMANCE REQUIREMENTS	
Specificity:	The matrix must affect results by less than 2% and interferents must affect the results by less than 2% at the nominal concentration.
Linearity:	The method shows acceptable accuracy across the stated range.
Precision:	Repeatability must be 2% or better across the stated range. Intermediate precision and reproducibility must be 2% or better at the nominal concentration. Intermediate precision and reproducibility F ratios must be less than 1 at the nominal concentration.
Bias:	The bias is no more than +/-3% at the extremes of the stated range. The bias is no more than +/- 2% at the nominal concentration.
Ruggedness:	Any experimental factors, which are found to affect the results are controlled within defined limits in the methodology.

WORKED EXAMPLE 1 - ASSAY

METHODS OF EVALUATION		
Parameter	Tests	Acceptance Criteria
Specificity:	Sample of Sodium Chloride at nominal concentration in water. Given a value of 100%	Not applicable assigned a value of 100% and all other results calculated with reference to this result.
	Sample of water	<2.0%
	Sample of Sodium Chloride in water with interferent A	98.0%<R<102.0%
	Sample of Sodium Chloride in water with excipients B	98.0%<R<102.0%
Linearity:	Full Formulation – see tests for bias.	Bias meets acceptance criteria at extremes of range and at nominal concentration.
Repeatability	Test full formulation in triplicate at 50%, 100% and 150% of nominal concentration.	%Coefficient of Variation <2%
Intermediate Precision	Test the same batch of the full formulation on two separate occasions in triplicate at the nominal concentration. Repeatability test results can be used provided the same batch is used.	Intermediate %Coefficient of Variation <2% Intermediate F Ratio <1
Reproducibility	Test the same batch of the full formulation on two separate occasions in house and once by external laboratory, all in triplicate at nominal concentration. Repeatability and intermediate precision test results can be used provided the same batch is used.	Reproducibility %Coefficient of Variation <2% Reproducibility F Ratio <1
Bias:	Test full formulation in triplicate at 50% and 150% of nominal concentration. Repeatability test results can be used.	97.0%<R<103.0%
	Carry out nine determinations on the same batch of the full formulation at nominal concentration (reproducibility test results can be used).	Bias insignificant and significant bias <2% OR Bias insignificant and significant bias >2% and maximum bias <2% OR Bias significant and significant bias <2% and maximum bias <2%
Ruggedness testing if necessary:	Test a batch of the full formulation in duplicate at the nominal concentration with: (a) ppt. time 30 sec and 60 sec; (b) volume ammonium iron III sulphate 3 ml and 7 ml; (c) fade time 2 min and 5 min; (d) volume of nitric acid 13 ml and 17 ml (e) volume of dibutyl phthalate 3 ml or 7 ml.	No significant difference in results for factor (a), (b), (c), (d) or (e).

<table>
<tr><th colspan="4">EVALUATION OF RESULTS - SPECIFICITY</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An evaluation of specificity is required by carrying out the method on:
1. a solution containing the analyte (sodium chloride) at its nominal concentration in the matrix;
2. a sample of the matrix (water);
3. a sample containing the analyte and interferent A at their nominal concentrations in the matrix;
4. a sample containing the analyte and excipients B at their nominal concentrations in the matrix.
Results from the matrix must be less than 2% and the potential interferents must affect results by no more than 2%.</td></tr>
<tr><td>Experiments</td><td colspan="3">All samples must quantitatively contain the same amount of the analyte although the exact concentration need not be known. This is achieved by preparing a solution containing four times the nominal concentration of sodium chloride (i.e. 6%) and adding it by volume.
Samples are tested under repeatability conditions, in duplicate.
Sample 1 is prepared by quantitatively diluting 25 ml of the concentrated solution of sodium chloride to 100 ml with water.
Sample 2 is water.
Sample 3 is prepared by dissolving 1.0 g of Sodium Bicarbonate in 50 ml water, adding by pipette 25 ml of the Sodium Chloride solution and diluting to 100 ml with water.
Sample 4 is prepared by adding 25 ml of the Sodium Chloride solution to 50 ml of Double Strength Chloroform Water, adding 2.5 ml Concentrated Peppermint Emulsion and diluting to 100 ml with water.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">As the method is a back titration, calculate the quantity of Silver Nitrate used in the reaction (V) from the expression 50 – T, where T is the titre.
Calculate the average for each sample (V_1, V_2, V_3 and V_4).
To demonstrate specificity the exact concentration need not be known as the acceptability depends on the percentage effect of the contributors. Therefore the values V_{1-4} can be used to represent concentration.
Calculate the concentration (C_x) for Samples 2 to 4 relative to Sample 1, by the general formula (V_x/V_1) * 100.
Confirm that the values for C_2 to C_4 meet the acceptance criteria</td></tr>
<tr><td>Results</td><td>Results:

C_2 = 1.17%
C_3 = 98.47%
C_4 = 99.80%</td><td>Acceptance criteria:

< 2%
98.0%<R<102.0%
98.0%<R<102.0%</td><td>Pass/Fail:

Pass
Pass
Pass</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to improve specificity is required.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently specific for purpose.</td></tr>
</table>

EVALUATION OF RESULTS - LINEARITY			
Performance Criteria	An evaluation of linearity is required by carrying out the method on the formulation at the extremes of the stated range and at the nominal concentration. The results from bias testing will be used and the acceptance criteria for bias must be met.		
Experiments	See bias testing.		
Evaluation of data	See bias testing.		
Results	Results - see bias testing. Bias – meets criteria	Acceptance criteria: Meets bias criteria at extremes and nominal.	Pass/Fail: Pass
Discussion	Acceptance criteria are met therefore no further work to improve linearity is required.		
Conclusion	The method is considered sufficiently linear for purpose.		

<table>
<tr><th colspan="4">EVALUATION OF RESULTS - PRECISION</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An estimate of precision is required by carrying out replicate determinations on single batches of the full formulation.
Repeatability is estimated by carrying out the method in triplicate at 50%, 100% and 150% of the nominal concentration.
Intermediate precision is estimated by carrying out the method in triplicate on two occasions at the nominal concentration.
Reproducibility is estimated by carrying out the method in triplicate on three occasions and in two laboratories at the nominal concentration.
The repeatability must be 2% or better across the stated range and the intermediate precision and reproducibility must be 2% or better and have an F ratio of less than 1 at the nominal concentration.</td></tr>
<tr><td>Experiments</td><td colspan="3">The exact concentration of Sodium Chloride need not be known if the experiments are to be used solely for determining precision. If the results are also going to be used for bias testing, then the weight of sodium chloride used must be recorded accurately and solutions made up volumetrically. Also the purity of the sodium chloride and its uncertainty value need to be known (from a C. of A.). Sufficient must be prepared for all experiments, quantities quoted are for 100 ml.
Samples 1, 2 and 3 are prepared by dissolving 0.75, 1.5 and 2.25 g respectively of Sodium Chloride and 1.0 g of Sodium Bicarbonate in 50 ml double strength chloroform water, adding 2.5 ml concentrated peppermint emulsion and diluting to 100 ml with water.
Samples 1 and 3 are tested on one occasion in triplicate. Each set of triplicate tests must be carried out by the same analyst on the same day and using the same equipment and reagents.
Sample 2 is tested in triplicate on three occasions. Two sets of triplicate tests must be carried out in house, by different analysts, on different days and using, as far as possible, different equipment and reagents and the third set must be carried out by another laboratory.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Calculate the concentration for each test as a percentage of the theoretical value, using the formula (50 – T)*0.5844*10/W, where T is the titre and W is the quantity of sodium chloride used to prepare 100 ml of the sample.
Enter the replicate results into the repeatability and intermediate precision/ reproducibility spreadsheets to determine the %CV and F ratios. The F ratios are calculated from ANOVA (analysis of variance) and values above 1 indicate a statistically significant difference between analysts or laboratories.
Confirm that the values for %CV and F ratios meet the acceptance criteria</td></tr>
<tr><td>Results</td><td>Results:
At 50%: Repeatability %CV = 0.77
At 150%: Repeatability %CV = 0.30
At 100%: Repeatability %CV = 0.49
Intermediate Precision %CV = 0.42
Intermediate Precision F ratio = 0.44
Reproducibility %CV = 0.38
Reproducibility F ratio = 0.77</td><td>Acceptance criteria:
=< 2%
=< 2%
=< 2%
=< 2%
< 1
=< 2%
< 1</td><td>Pass/Fail:
Pass
Pass
Pass
Pass
Pass
Pass
Pass</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to improve precision is required. The repeatability at 50% is less good although still well within the limit. Care should be taken applying the method to products with nominal concentrations at this end of the range. If necessary the volume of sample taken should be adjusted accordingly.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently precise for purpose.</td></tr>
</table>

WORKED EXAMPLE 1 - ASSAY

<table>
<tr><th colspan="4">EVALUATION OF RESULTS – BIAS</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An estimation of bias is required across the range by carrying out determinations at 50%, 100% and 150% of the nominal concentration using solution with an accurately known concentration.
The bias must be no more than +/-3% at the extremes of the range and no more than +/-2% at the nominal concentration.</td></tr>
<tr><td>Experiments</td><td colspan="3">If the samples used for precision testing were prepared volumetrically, with accurately recorded weights, the results can be used for the estimations of bias. If samples for bias testing need to be prepared, the weight of sodium chloride used must be recorded accurately and solutions made up volumetrically. Also the purity of the sodium chloride and its uncertainty value need to be known (from a C. of A.). Sufficient must be prepared for all experiments, quantities quoted are for 100 ml.
Samples 1, 2 and 3 are prepared by dissolving 0.75, 1.5 and 2.25 g respectively of Sodium Chloride and 1.0 g of Sodium Bicarbonate in 50 ml double strength chloroform water, adding 2.5 ml concentrated peppermint emulsion and diluting to 100 ml with water.
Samples 1 and 3 are tested on one occasion in triplicate. Each set of triplicate tests must be carried out by the same analyst on the same day and using the same equipment and reagents.
Sample 2 is tested nine times. Precision testing results can be used provided the samples were prepared accurately and the precision was acceptable.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Calculate the concentration for each test as a percentage of the theoretical value, using the formula (50 – T)*0.5844*1000/W*P, where T is the titre, W is the weight of sodium chloride of purity P%, used to prepare 100 ml of the sample.
At the extremes of the range, calculate the mean of the triplicate results and confirm that the bias values obtained by subtracting 100 from the mean of the results, meet the acceptance criteria.
At the nominal concentration, enter the replicate results into the bias testing – comparison with standard spreadsheet to determine the size and significance of any bias. Confirm that any bias is within the acceptance criteria.</td></tr>
<tr><td>Results</td><td>Results:
At 50%: Bias = 1.29%
At 150%: Bias = -0.37%
At 100%: Bias = -0.15%
Bias significant? = No
Significant bias = 0.29%
Maximum bias = 0.94%</td><td>Acceptance criteria:
< +/-3%
< +/-3%
Bias insignificant and significant bias <2% OR
Bias insignificant and significant bias >2% and maximum bias <2% OR
Bias significant and significant bias <2% and maximum bias <2%</td><td>Pass/Fail:
Pass
Pass
Pass

N/A

N/A</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to reduce bias is required.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently free from bias for purpose.</td></tr>
</table>

EVALUATION OF RESULTS - RUGGEDNESS

Performance Criteria

An evaluation of ruggedness is required by carrying out the method on the full formulation at its nominal concentration and varying any experimental factors which are considered likely to influence the results.
Factors to be tested and their 'low' and 'high' values are:

(a) precipitation time 30 sec and 60 sec;
(b) volume ammonium iron III sulphate 3 ml and 7 ml;
(c) fade time 2 min and 5 min;
(d) volume of nitric acid 13 ml and 17 ml
(e) volume of dibutyl phthalate 3 ml or 7 ml.

There should be no significant difference between the results obtained with each factor at its low and high value.

Experiments

The sample to be tested in each experiment must quantitatively contain the same amount of the analyte although the exact concentration need not be known. Therefore sufficient for all experiments must be prepared as one batch.
Eight experiments are carried out in duplicate, according to the table below which is based on the Plackett-Burman design. As only five factors are being examined, two dummy factors have to be inserted into the table. These are factors which cannot have an effect, those chosen were:

(f) burette marked with 'X' and 'Y' (same burette used);
(g) experiment carried out AM and PM.

	Factors							
Experiment		a	b	c	d	e	f	g
	1	60	7	5	17	3	Y	PM
	2	60	7	2	17	7	X	AM
	3	60	3	5	13	3	X	AM
	4	60	3	2	13	7	Y	PM
	5	30	7	5	13	7	Y	AM
	6	30	7	2	13	3	X	PM
	7	30	3	5	17	7	X	PM
	8	30	3	2	17	3	Y	AM

Evaluation of data

Calculate the concentration for each test as a percentage of the theoretical value, using the formula (50 – T)*0.5844*1000/W*P, where T is the titre, W is the weight of sodium chloride of purity P%, used to prepare 100 ml of the sample.
Enter the results into the ruggedness testing spreadsheet. The %CV from the precision testing needs to be entered. This should be the %CV for the intermediate precision.
Confirm that there is no significant difference between the results for each factor.
Check the difference quoted for each factor and decide which have the greatest effect.
Decide whether any of these need to be specified more tightly in the method.

Results

Results (in difference order):	Acceptance criteria:	Pass/Fail:
F_c = 0.32		Pass
F_d = 0.22	No significant difference	Pass
F_e = 0.16		Pass
F_b = 0.13		Pass
F_a = 0.08		Pass

Discussion

The greatest difference was observed with factor 'c', the time allowed for the colour to fade. Even this however showed no statistically significant difference between the two extremes.

Conclusion

The method is sufficiently rugged and it is not considered necessary to specify any parameters more tightly.

WORKED EXAMPLE 1 - ASSAY

SUMMARY OF RESULTS

Specificity

Sample 1: Average assay result = 1.499%
Sample 2: Average assay result = 0.0175% = 1.17% relative to Sample 1
Sample 3: Average assay result = 1.476% = 98.47% relative to Sample 1
Sample 4: Average assay result = 1.496% = 99.80% relative to Sample 1

Linearity

See results for Bias

Precision

Sample 1: Repeatability results =	0.7597% =	101.29% of stated
	0.7539% =	100.51% of stated
	0.7656% =	102.07% of stated
Sample 2: Repeatability results =	1.496% =	99.71% of stated
	1.505% =	100.29% of stated
	1.511% =	100.68% of stated
Sample 3: Repeatability results =	2.232% =	99.28% of stated
	2.244% =	99.80% of stated
	2.244% =	99.80% of stated
Sample 2: Intermediate results =	1.496% =	99.71% of stated
	1.496% =	99.71% of stated
	1.496% =	99.71% of stated
Sample 2: Reproducibility results =	1.496% =	99.71% of stated
	1.493% =	99.51% of stated
	1.495% =	99.61% of stated

Bias

Sample 1: Results =	0.7597% =	101.29% of stated	
	0.7539% =	100.51% of stated	Mean = 101.29%
	0.7656% =	102.07% of stated	
Sample 3: Results =	2.232% =	99.28% of stated	
	2.244% =	99.80% of stated	Mean = 99.63%
	2.244% =	99.80% of stated	
Sample 2: Results =	1.496% =	99.71% of stated	
	1.505% =	100.29% of stated	
	1.511% =	100.68% of stated	
	1.496% =	99.71% of stated	Mean = 99.85%
	1.496% =	99.71% of stated	
	1.496% =	99.71% of stated	
	1.496% =	99.71% of stated	
	1.493% =	99.51% of stated	
	1.495% =	99.61% of stated	

Ruggedness

Experiment results (% of stated)							
1	2	3	4	5	6	7	8
99.69	100.47	99.88	100.66	100.27	100.47	99.45	100.08
99.26	99.45	99.26	99.45	99.26	99.45	99.45	99.07

Repeatability Spreadsheet:

Repeatability Testing			
Enter triplicate results for each concentration (minimum)			
	Low Concentration	Nominal Concentration	High Concentration
Result 1 (required)	101.29	99.71	99.28
Result 2 (required)	100.51	100.29	99.80
Result 3 (required)	102.07	100.68	99.80
Result 4 (optional)			
Result 5 (optional)			
Number or results	3.00	3.00	3.00
Mean	101.29	100.23	99.63
Standard Deviation	0.78	0.49	0.30
RSD	0.0077	0.0049	0.0030
% Coefficient of variation	**0.77**	**0.49**	**0.30**

Intermediate Precision and Reproducibility Spreadsheet

Intermediate Precision/Reproducibility Testing			
Enter triplicate results for each set of tests			
	Laboratory 1	Laboratory 1	Laboratory 2
	Test 1	Test 2	Test 3
Replicate 1	99.71	99.71	99.71
Replicate 2	100.29	99.71	99.51
Replicate 3	100.68	99.71	99.61
Mean	100.227	99.710	99.610
SD	0.488	0.000	0.100
No. of results	3	3	3
Test Degrees Freedom	2	2	2
Total Degrees Freedom		5	8
Mean of all		99.97	99.85
SD of all		0.419	0.380
Intermediate Precision			
F Statistic	**3.362**		
F Critical	**7.709**		
F Ratio	**0.436**		
Intermediate Standard Deviation		**0.418779974**	
Intermediate % Coefficient of Variation		**0.41891263**	
Reproducibility			
F Statistic	**3.971**		
F Critical	**5.143**		
F Ratio	**0.772**		
Reproducibility Standard Deviation		**0.379751381**	
Reproducibility % Coefficient of Variation		**0.380326096**	

WORKED EXAMPLE 1 - ASSAY

Bias testing spreadsheet

Bias Testing - Comparison with Standard							
Enter nine results (minimum) for the method (as % of stated)							
		Test			Enter Reference		
Enter results on 3 x 3 grid		99.71	99.71	99.71	Value (normally 100)		100
		100.29	99.71	99.51			
		100.68	99.71	99.61	Uncertainty (+/- %)		0.5
Enter additional optional results on 3 x 2 grid							
		Number	9		Number		1
		Mean	99.84888889		Mean		100
		St Dev	0.379751381				
		Variance	0.144211111		Minimum		99.5
		St Dev of Mean	0.126583794		Maximum		100.5
t Statistic	1.194	t Crit	2.306	**Bias is not statistically significant**			
Maximum value of bias considered significant		**0.29190**					
Reported bias		**0.15111**					
Range of Actual Means	Test	99.557	100.141				
	Reference	99.500	100.500				
Maximum theoretical bias		0.943					

WORKED EXAMPLE 1 - ASSAY

Ruggedness testing spreadsheet

Ruggedness Testing												
Enter duplicate or Triplicate results for each experiment (as % of sataed)												
Enter the % Coefficient of Variation from Precision testing	0.42											
Enter the number of results this was based on (3, 6 or 9)	6											
Enter the number of experiments (4, 8 or 12)	8											
	Experiment Number											
	1	2	3	4	5	6	7	8	9	10	11	12
Enter duplicate or Triplicate results for each experiment	99.69	100.47	99.88	100.66	100.27	100.47	99.45	100.08				
	99.26	99.45	99.26	99.45	99.26	99.45	99.45	99.07				
Average:	99.48	99.96	99.57	100.06	99.77	99.96	99.45	99.58				
	Factor Number											
	a	b	c	d	e	f	g	h	I	j	k	
Low level result averages:	99.69	99.66	99.89	99.84	99.81	99.74	99.72					
High level result averages:	99.77	99.79	99.57	99.62	99.65	99.72	99.74					
Difference	**0.08**	**0.13**	**0.32**	**0.22**	**0.16**	**0.02**	**0.02**					
t value	**0.26**	**0.43**	**1.09**	**0.75**	**0.55**	**0.06**	**0.06**					
t critical 2.57	**Not Signif**	**Not Signif**	**Not Signif**	**Not Signif**	**Not Signif**	**Not Signif**	**Not Signif**					

VALIDATION PLAN

DETERMINATION OF 4-AMINOPHENOL IN PARACETAMOL TABLETS BP

Written by ……. Date……..

Approved by…….. Version…….

PRODUCT FORMULATION	
Paracetamol	500 mg
Inert ingredients and excipients to	0.562 g (taken from average weight of tablets)

INTRODUCTION	
Method:	HPLC method using hypersil ODS column and isopropyl alcohol/ water/ formic acid mobile phase.
Background:	A fully quantifiable method is needed for stability trials on repacked paracetamol tablets. The method needs to quantify 4-aminophenol to a nominal limit concentration of 5% of the paracetamol content.
Methodology:	Solution A - Accurately weigh about 1.0 g paracetamol, dissolve in 20 ml ethanol and dilute to 100 ml with ethanol. Solution B - Accurately weigh about 0.1 g 4-aminophenol, dissolve in 20 ml ethanol and dilute to 50 ml with ethanol. Dilute 5 ml to 20 ml volumetrically with ethanol. Standard – Pipette 10 ml of Soln. A and 10 ml Soln. B into 100 ml volumetric flask and dilute to volume with the mobile phase. Purity check – Pipette 10 ml of Soln. A into a 100 ml volumetric flask and dilute to volume with mobile phase, check for absence of a 4-aminophenol peak. Test – Accurately weigh an amount of mixed powder equivalent to 1.0 g paracetamol into a 100 ml flask, sonicate for 10 minutes with 30 ml ethanol, filter through filter paper into a 100 ml volumetric flask and rinse the flask through the same filter paper. Wash the filter paper with 2x5 ml portions of ethanol and dilute to volume with ethanol. Pipette 10 ml into a 100 ml volumetric flask and dilute to volume with mobile phase. HPLC parameters: Column: Hypersil ODS Mobile Phase: 1000 ml Water, 176 ml Isopropyl Alcohol, 1 ml Formic Acid Wavelengths: 218 and 236 nm Flow rate: 1.0 ml/min Injection volume: 20 mcl
Other considerations:	The powdered paracetamol tablets are used to prepare spiked samples should be tested for 4-aminophenol content. Reproducibility not required as method only to be used for in-house stability. Consider ruggedness testing, if intermediate precision tests are outside the acceptance criteria. The matrix and excipients are not available for specificity testing.

WORKED EXAMPLE 2 – QUANTITATIVE IMPURITY TEST

ANALYTICAL REQUIREMENT		
Analyte:	4-Aminophenol Limit concentration is 5% of the Paracetamol content. The method is to be tested between 10 and 200% of a concentration equivalent 5% of the Paracetamol content.	
Matrix:	Unknown powder.	
Potential Interferants:	Paracetamol Excipients	Considered together in inspiked tablets.

PERFORMANCE REQUIREMENTS	
Specificity:	As the matrix cannot be replicated, a sample of the analyte in the matrix cannot be prepared. Spiked and unspiked samples are used for testing. The resolution factor between the 4-Aminophenol peak and any other peak is at least 1.5 across the whole range. Peak purity for the 4-Aminophenol peak in the spiked sample should be similar to that obtained with the standards. Unspiked sample should show nothing which will interfere with detection of the analyte peak.
Linearity:	Inspection of the plot of response against concentration appears linear. Inspection of the residuals plot shows random distribution. Linear regression correlation coefficient is greater than 0.999. The method shows acceptable accuracy across the stated range.
Precision:	Repeatability must be 10% or better across the stated range. Intermediate precision must be 10% or better at the nominal limit concentration. Intermediate precision F ratio must be less than 1 at the nominal limit concentration.
Limit of Detection/ Quantitation:	The method shows acceptable precision at the bottom of the stated range.
Bias:	The bias is no more than +/-10% at the extremes of the stated range. The bias is no more than +/- 5% at the nominal limit concentration.
Ruggedness:	Any experimental factors, which are found to affect the results are controlled within defined limits in the methodology.

METHODS OF EVALUATION		
Parameter	Tests	Acceptance Criteria
Specificity:	Sample of powdered paracetamol tablets spiked with 4-aminophenol at 5.0% of the paracetamol content.	Resolution Factor between 4-aminophenol peak and any other peak is <= 1.5.
	Sample of powdered paracetamol tablets	Any peak corresponding to the 4-aminophenol has profile and peak purity similar to standards.
Linearity:	Standards of 4-aminophenol in mobile phase. Range 0.5 to 2 mg/100ml. Sample of powdered paracetamol tablets spiked with 4-aminophenol.	Plot of response against concentration linear. Residuals plot shows random distribution. Linear regression correlation coefficient is greater than 0.999. Bias meets acceptance criteria at extremes of range and at nominal concentration.
Repeatability	Test, in triplicate, samples of powdered paracetamol tablets spiked with 4-aminophenol at 0.5%, 5% and 10% of the paracetamol content.	%Coefficient of Variation <10%
Intermediate Precision	Test, in triplicate, on two occasions, the same sample of powdered paracetamol tablets spiked with 4-aminophenol at 5% of the paracetamol content. Repeatability test results can be used provided the same sample is used.	Intermediate %Coefficient of Variation <10% Intermediate F Ratio <1
Bias:	Test, in triplicate, sample of powdered paracetamol tablets spiked with 4-aminophenol at 0.5% and 10% of the paracetamol content. Repeatability test results can be used.	90.0%<R<110.0%
	Carry out nine determinations on sample of powdered paracetamol tablets spiked with 4-aminophenol at 5% of the paracetamol content. Repeatability and intermediate precision test results can be used.	Bias insignificant and significant bias <5% OR Bias insignificant and significant bias >5% and maximum bias <5% OR Bias significant and significant bias <5% and maximum bias <5%
Ruggedness testing if necessary:	Test a sample of powdered paracetamol tablets spiked with 4-aminophenol acid at 5% of the paracetamol content with: (a) flow rate 0.8 and 1.2 ml/min; (b) mobile phase 160 and 190 ml isopropyl alcohol; (c) mobile phase 0.8 and 1.2 ml formic acid (d) sonicate 8 and 12 min (e) filter paper type 1 and type 2 (f) ethanol volume 30 ml and 50 ml (g) wash 2x5 ml and 4x5 ml.	No significant difference in results for factor (a), (b), (c), (d), (e), (f) or (g).

<table>
<tr><th colspan="4">EVALUATION OF RESULTS – SPECIFICITY</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An evaluation of specificity is required by carrying out the method on a sample of powdered paracetamol tablets spiked with 4-aminophenol at 5% of the paracetamol content. The resolution factor must be greater than or equal to 1.5.
Interference from other ingredients is also checked by testing a sample of unspiked samples. There is nothing in the chromatogram to interfere with detection of the analyte. When peak purity thresholds are set with the chromatograms of the standards the unspiked sample passes the peak purity test. Any peak present in the unspiked sample has a similar profile to the 4-aminophenol peak.</td></tr>
<tr><td>Experiments</td><td colspan="3">The chromatograms obtained from the repeatability tests on powdered tablets spiked with 4-aminophenol at 5% of the paracetamol content may be used.
Test unspiked samples by the same method.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Determine the resolution between the 4-aminophenol peak and any other peak by the peak width at half height method of the BP.
Confirm that the peak resolution values meet the acceptance criteria
Carry out peak purity test using standard chromatography software and setting thresholds from the chromatograms of the standards.
Check the UV profile of any peak in unspiked sample which corresponds to the 4-aminophenol peak.</td></tr>
<tr><td>Results</td><td>Results:

R_1 = 3.15
R_2 = 1.89
R_3 = 1.90

Peak purity threshold set by standard = 990
Peak purity of unspiked sample = 999

UV profile of the peak in the unspiked sample matches that due to 4-aminophenol peak in standards.</td><td>Acceptance criteria:

=> 1.5
=> 1.5
=> 1.5

Unspiked sample peak purity > set threshold.

Matches</td><td>Pass/Fail:

Pass
Pass
Pass

Pass

Pass</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to improve specificity is required.
The matching of the profile of the peak in the unspiked sample with that of the 4-aminophenol peak in the standard together with the peak purity test indicates it is solely due to 4-aminophenol. Therefore nothing else present to interfere.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently specific for purpose.</td></tr>
</table>

<table>
<tr><th colspan="4">EVALUATION OF RESULTS - LINEARITY</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An evaluation of instrumental linearity is required by carrying out the chromatography on a range of standard solutions. The linear regression correlation coefficient must be greater than 0.999 and the residuals plot must appear randomly distributed.
An evaluation of method linearity is required by carrying out the method on the formulation spiked with salicylic acid at the extremes of the stated range and at the nominal limit concentration. The results from bias testing will be used and the acceptance criteria for bias must be met.</td></tr>
<tr><td>Experiments</td><td colspan="3">A solution is prepared by dissolving 0.25 g 4-aminophenol in sufficient ethanol to produce 100 ml. A range of intermediate standards is then prepared by diluting 1 ml, 2 ml, 5 ml, 10 ml, 15 ml and 20 ml to 50 ml volumetrically with ethanol. 5 ml of each intermediate standard is then pipetted into 50 ml volumetric flasks and each is diluted to volume with mobile phase. These standards are then injected in triplicate onto the chromatograph.

For method linearity, see bias testing.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">For instrumental linearity, enter the three results as peak area for each standard and the concentration as the volume used in preparing the intermediate standards.
The spreadsheet will calculate the least squares regression line and plot this together with the residuals and the variance.
Confirm that the regression plot appears linear and that the residuals plot appears randomly distributed about zero.
Confirm that the correlation coefficient meets the acceptance criterion.

For method linearity, see bias testing.</td></tr>
<tr><td>Results</td><td>Results:
Regression plot linear? Yes
Residuals random? No
Correlation coefficient = 0.9997

See bias testing.

Bias – meets criteria at extremes but not at nominal limit concentration.</td><td>Acceptance criteria:
Linear
Random
>=0.999

Meets bias criteria at extremes and nominal.</td><td>Pass/Fail:
Pass
Fail
Pass

Fails at nominal</td></tr>
<tr><td>Discussion</td><td colspan="3">The residuals were not random, but followed a curve, the shape of which indicates a non-linearity. Looking closely at the linear regression plot, it can be seen that there is a slight levelling off at the highest concentration. Removing the top point improves the distribution of the residuals so that they appear more random.
The bias at the top of the range is within the acceptance criteria. It is thought that most of the bias is due to extraction rather than non-linearity as there is also a bias at the nominal concentration.
Acceptance criteria are not fully met, however the slight instrument non-linearity at the top of the range is not considered significant therefore no further work to improve linearity is required.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The instrumental linearity method is considered sufficient for purpose.
Whole method linearity will be addressed with further bias testing.</td></tr>
</table>

<table>
<tr><th colspan="4">EVALUATION OF RESULTS - PRECISION</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An estimate of precision is required by carrying out replicate determinations on single batches of powdered tablets spiked with 4-aminophenol.
Repeatability is estimated by carrying out the method in triplicate at 10%, 100% and 200% of the nominal limit concentration.
Intermediate precision is estimated by carrying out the method in triplicate on two occasions at the nominal limit concentration.
The repeatability must be 10% or better across the stated range and the intermediate precision must be 10% or better and have an F ratio of less than 1 at the nominal limit concentration.</td></tr>
<tr><td>Experiments</td><td colspan="3">The exact concentration of 4-aminophenol need not be known if the experiments are to be used solely for determining precision. If the results are also going to be used for bias testing, then the weights of 4-aminophenol and of powdered tablets used must be recorded accurately and solutions made up volumetrically. Also the purity of the 4-aminophenol and its uncertainty value need to be known (from a C. of A.). Sufficient must be prepared for all experiments.
Samples 1, 2 and 3 are prepared by triturating 0.05, 0.5 and 1.0 g respectively of 4-aminophenol with 20 weighed and powdered tablets.
Samples 1 and 3 are tested on one occasion in triplicate. Each set of triplicate tests must be carried out by the same analyst on the same day and using the same equipment and reagents.
Sample 2 is tested in triplicate on two occasions. Two sets of triplicate tests must be carried out in house, by different analysts, on different days if practicable and using, as far as possible, different equipment and reagents.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Calculate the concentration for each test as a percentage of the theoretical value, using the formula $(A_t * W_s * W_t * 100)/(A_s * 2 * W_a * W)$, where A_t and A_s are the areas of the test and standard 4-aminophenol peaks respectively, W_s is the weight of 4-aminophenol used to prepare the standard, W_t is the weight of the 20 tablets used to prepare the spiked sample, W_a is the weight of 4-aminophenol used to prepare the spiked sample and W is the weight of spiked sample used.
Enter the replicate results into the repeatability and intermediate precision/ reproducibility spreadsheets to determine the %CV and F ratios. The F ratios are calculated from ANOVA (analysis of variance) and values above 1 indicate a statistically significant difference between analysts or laboratories.
Confirm that the values for %CV and F ratios meet the acceptance criteria</td></tr>
<tr><td>Results</td><td>Results:
At 50%: Repeatability %CV = 4.40
At 150%: Repeatability %CV = 0.48
At 100%: Repeatability %CV = 6.02
Intermediate Precision %CV = 9.05
Intermediate Precision F ratio = 1.60</td><td>Acceptance criteria:
=< 10%
=< 10%
=< 10%
=< 10%
< 1</td><td>Pass/Fail:
Pass
Pass
Pass
Pass
Fail</td></tr>
<tr><td>Discussion</td><td colspan="3">Although the %CV’s meet their acceptance criteria at all concentrations, the values are high and care must be exercise in using the method.
The intermediate precision test shows a significant between analyst variation (F ratio >1). On discussing with the technicians, this might be due to the filter papers used during the extraction.
Variations in replicate injections and in the standards throughout the sequence were not as great (%CV = 1.77), which points to variability in the extraction.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method should be subjected to ruggedness testing with particular emphasis on the extraction phase.</td></tr>
</table>

<table>
<tr><th colspan="4">EVALUATION OF RESULTS – BIAS</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An estimation of bias is required across the range by carrying out determinations at 10%, 100% and 200% of the nominal limit concentration using spiked samples with an accurately known concentrations. The bias must be no more than +/-10% at the extremes of the range and no more than +/-5% at the nominal concentration.</td></tr>
<tr><td>Experiments</td><td colspan="3">If the samples used for precision testing were prepared quantitatively, with accurately recorded weights, the results can be used for the estimations of bias. If samples for bias testing need to be prepared, the weight of 4-aminophenol used must be recorded accurately and solutions made up volumetrically. Also the purity of the 4-aminophenol and its uncertainty value need to be known (from a C. of A.). Sufficient must be prepared for all experiments, quantities quoted are for 100 ml.
Samples 1, 2 and 3 are prepared by triturating 0.05, 0.5 and 1.0 g respectively of 4-aminophenol with 20 weighed and powdered tablets.
Samples 1 and 3 are tested on one occasion in triplicate. Each set of triplicate tests must be carried out by the same analyst on the same day and using the same equipment and reagents.
Sample 2 is tested nine times by the same analyst on the same day. Precision testing results can be used provided the samples were prepared accurately.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Calculate the concentration for each test as a percentage of the theoretical value, using the formula $(A_t * W_s * W_t * 100)/(A_s * 2 * W_a * W)$, where A_t and A_s are the areas of the test and standard 4-aminophenol peaks respectively, W_s is the weight of 4-aminophenol used to prepare the standard, W_t is the weight of the 20 tablets used to prepare the spiked sample, W_a is the weight of 4-aminophenol used to prepare the spiked sample and W is the weight of spiked sample used.
At the extremes of the range, calculate the mean of the triplicate results and confirm that the bias values obtained by subtracting 100 from the mean of the results, meet the acceptance criteria.
At the nominal concentration, enter the replicate results into the ‘bias testing – comparison with standard’ spreadsheet to determine the size and significance of any bias. Confirm that any bias is within the acceptance criteria.</td></tr>
<tr><td>Results</td><td>Results:
At 50%: Bias = 7.15%
At 150%: Bias = -9.61%
At 100%: Bias = -8.11%
Bias significant? Yes
Significant bias = 4.68%
Maximum bias = -13.28%</td><td>Acceptance criteria:
< +/-10%
< +/-10%
Bias insignificant and significant bias <5% OR
Bias insignificant and significant bias >5% and maximum bias <5% OR
Bias significant and significant bias <5% and maximum bias <5%</td><td>Pass/Fail:
Pass
Pass
Fail

Fail

Fail</td></tr>
<tr><td>Discussion</td><td colspan="3">The bias at the low concentration is within the acceptance criteria but there is a high positive bias. This may be due to the contribution of a small amount of 4-aminophenol present in the tablets used.
At the high concentration, although again within acceptance criteria, there is a high negative bias, this might be due to incomplete extraction.
There is a high negative bias at the nominal limit concentration and failure to meet acceptance criteria. This might be due to the factor causing between analyst variation as one analyst obtained unbiased results.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method shows some bias and needs further testing. In particular variability in extraction needs to be examined through ruggedness testing.</td></tr>
</table>

SUMMARY OF RESULTS

Specificity

Resolution at 0.5% of paracetamol content = 3.15
Resolution at 5.0% of paracetamol content = 1.89
Resolution at 0.5% of paracetamol content = 1.90

Linearity

Volume	Replicate 1	Replicate 2	Replicate 3
1	122	128	134
2	241	239	237
5	628	621	631
10	1249	1243	1242
15	1843	1846	1843
20	2387	2398	2400

See also results for Precision and Bias

Precision

Sample 1: Repeatability results = 109.57% of stated
110.29% of stated
101.60% of stated
Sample 2: Repeatability results = 84.74% of stated
84.81% of stated
93.93% of stated
Sample 3: Repeatability results = 90.83% of stated
89.96% of stated
90.38% of stated
Sample 2: Intermediate results = 99.18% of stated
100.18% of stated
97.37% of stated

Bias

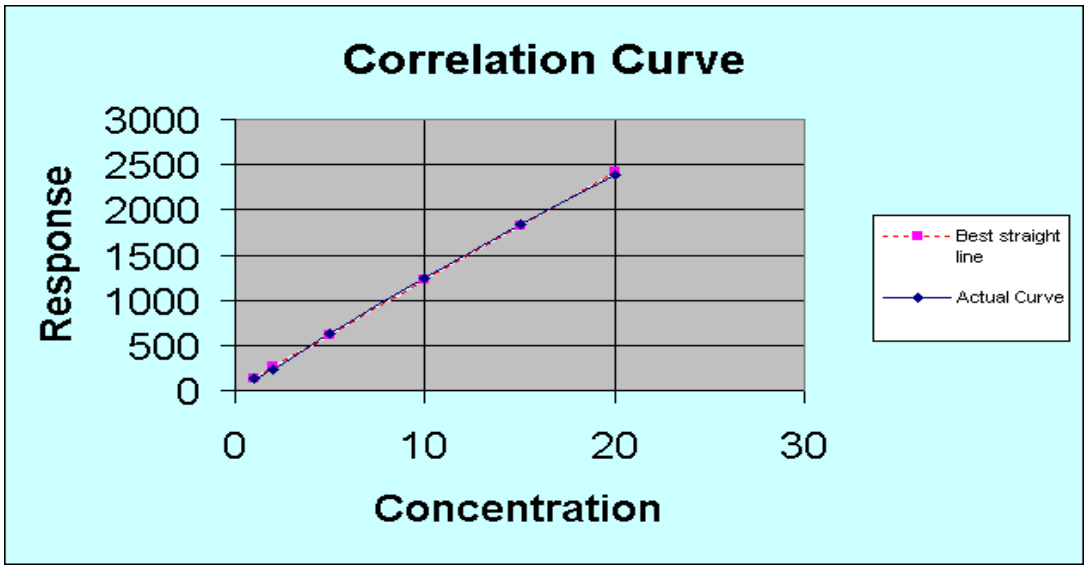
Sample 1: Results = 109.57% of stated
110.29% of stated
101.60% of stated
Mean = 107.15%
Sample 3: Results = 90.83% of stated
89.96% of stated
90.38% of stated
Mean = 90.39%
Sample 2: Results = 84.74% of stated
84.81% of stated
93.93% of stated
99.18% of stated
100.18% of stated
97.37% of stated
88.81% of stated
86.50% of stated
91.51% of stated
Mean = 91.89%

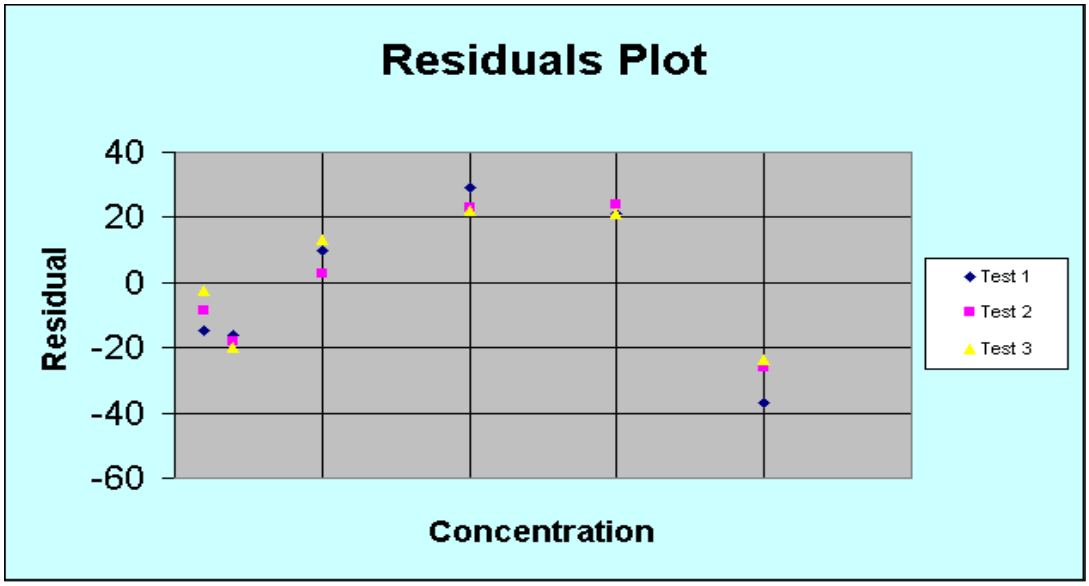
Linearity spreadsheet

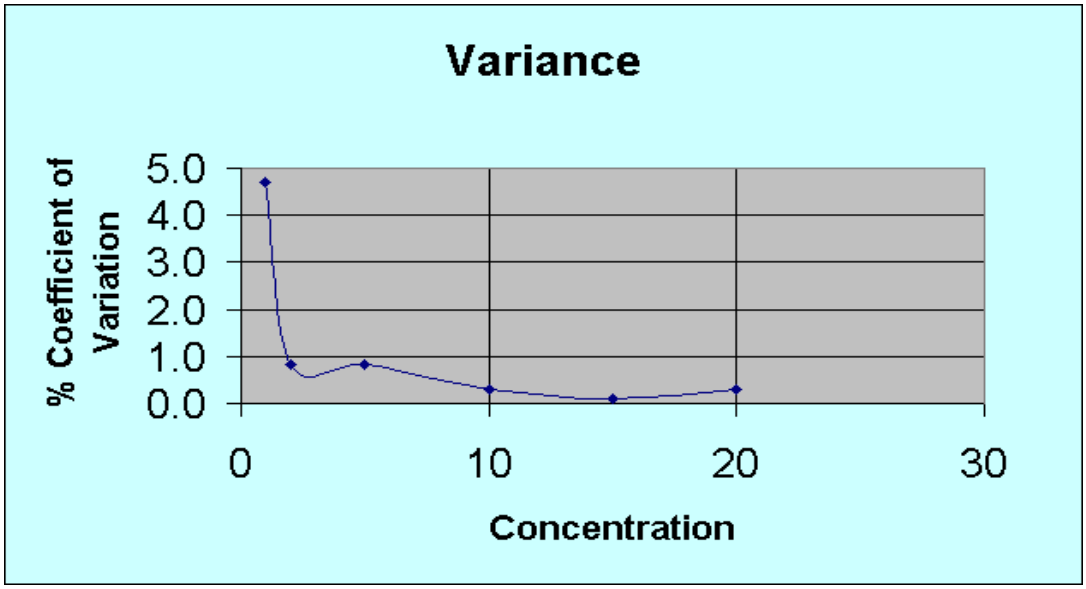
Linearity Testing								
Enter triplicate results for each concentration								
	Concentration	Replicate 1	Replicate 2	Replicate 3		Average	Stand Dev	%RSD (%CV)
Point 1	1	122	128	134		128.0	6.000	4.7
Point 2	2	241	239	237		239.0	2.000	0.8
Point 3	5	628	621	631		626.7	5.132	0.8
Point 4	10	1249	1243	1242		1244.7	3.786	0.3
Point 5	15	1843	1846	1843		1844.0	1.732	0.1
Point 6	20	2387	2398	2400		2395.0	7.000	0.3
Intercept	16.1817							
Slope	120.382							
Correlation	0.99971							

WORKED EXAMPLE 2 – QUANTITATIVE IMPURITY TEST

Linearity plots







Repeatability spreadsheet

Repeatability Testing			
Enter triplicate results for each concentration (minimum)			
	Low Concentration	Nominal Concentration	High Concentration
Result 1 (required)	109.57	84.74	90.83
Result 2 (required)	110.29	84.81	89.96
Result 3 (required)	101.60	93.93	90.38
Result 4 (optional)			
Result 5 (optional)			
Number or results	3.00	3.00	3.00
Mean	107.15	87.83	90.39
Standard Deviation	4.82	5.29	0.44
RSD	0.0450	0.0602	0.0048
% Coefficient of variation	**4.50**	**6.02**	**0.48**

Intermediate precision spreadsheet

Intermediate Precision/Reproducibility Testing			
Enter triplicate results for each set of tests			
	Laboratory 1 Test 1	Laboratory 1 Test 2	Laboratory 2 Test 3
Replicate 1	84.74	99.18	
Replicate 2	84.81	100.18	
Replicate 3	93.93	97.37	
Mean	87.827	98.910	
SD	5.286	1.424	
No. of results	3	3	
Test Degrees Freedom	2	2	
Total Degrees Freedom		5	
Mean of all		93.36	
SD of all		7.000	
Intermediate Precision			
F Statistic	**12.297**		
F Critical	**7.709**		
F Ratio	**1.595**		
Intermediate Standard Deviation		**8.450411233**	
Intermediate % Coefficient of Variation		**9.050618053**	

Bias spreadsheet

Bias Testing - Comparison with Standard								
Enter nine results (minimum) for the method (as % of stated)								
		Test			Enter Reference			
Enter results on 3 x 3 grid		84.74	99.18	88.81	Value (normally 100)			100
		84.81	100.18	86.50				
		93.93	97.37	91.51	Uncertainty (+/- %)			0.5
Enter additional optional results on 3 x 2 grid								
		Number	9		Number			1
		Mean	91.89222222		Mean			100
		St Dev	6.082679874					
		Variance	36.99899444		Minimum			99.5
		SD of Mean	2.027559958		Maximum			100.5
t Statistic	3.999	t Crit	2.306	**Bias is statistically significant**				
Maximum value of bias considered significant		**4.67556**						
Reported bias		**8.10778**						
Range of Actual Means	Test	87.201	96.568					
	Reference	99.500	100.500					
Maximum theoretical bias		**13.283**						

VALIDATION PLAN – ABRIDGED VALIDATION

ASSAY OF DOXAZOSIN MESYLATE IN DOXAZOSIN TABLETS

Written by ……. Date……..

Approved by…….. Version…….

PRODUCT FORMULATION	
Doxazosin	4 mg
Inert ingredients and excipients to	0.4857 g (taken from average weight of tablets)

INTRODUCTION	
Method:	UV absorbance of ethanolic solution at the maximum at about 245 nm.
Background:	A method is required to assay a generic brand of Doxazosin Tablets for the purpose of awarding on contract. Due to the short time available and the ad hoc nature of the request, an abridged validation is deemed appropriate.
Methodology:	Standard: Transfer about 0.1 g doxazosin mesylate accurately weighed (W_s) to a 100 ml volumetric flask and dissolve in sufficient ethanol 96% to produce 100 ml. Dilute 1 ml to 100 ml with ethanol 96%. Test: Transfer a quantity of powdered tablets W_t equivalent to 5 mg to a 100 ml volumetric flask Add approximately 80 ml of ethanol 96% and dissolve as completely as possible by sonicating for 30 minutes. Dilute to volume with ethanol 96%. Filter, discarding the first few mls of filtrate and dilute 10 ml to 50 ml with ethanol 96%. Read the uv absorbance of the test (A_t) and the standard (A_s) at the maximum at about 245 nm, using ethanol 96% as a blank. Content per tablet (mg as mesylate) = $\frac{A_t * W_s * P * AW}{A_s * W_t * 2}$
Other considerations:	All the excipients were not available for specificity testing, therefore tests were carried out using those available. Significant interference is considered unlikely due to absence of chromophores in unavailable excipients. All the excipients were not available to prepare samples therefore two batches were tested to look for any potential bias.

ANALYTICAL REQUIREMENT		
Analyte:	Doxazosin Nominal concentration is 4 mg doxazosin as the mesylate per tablet. The method is to be tested at 50 and 150% of the nominal concentration to ensure adequate linearity to detect out of specification results.	
Matrix:	Microcrystalline Cellulose Lactose Monohydrate	
Potential Interferants:	Magnesium Stearate Sodium Lauryl Sulphate Sodium Starch Glycollate Colloidal Anhydrous Silica	Considered jointly as excipients A

PERFORMANCE REQUIREMENTS	
Specificity:	The matrix must affect results by less than 2% and interferents must affect the results by less than 5% at the nominal concentration.
Precision:	Repeatability must be 5% or better across the stated range.
Linearity:	The method shows acceptable accuracy across the stated range.
Bias:	The bias is no more than +/-5% across the stated range. Testing of two random batches shows no potential bias.

METHODS OF EVALUATION		
Parameter	Tests	Acceptance Criteria
Specificity:	Sample of Doxazosin at nominal concentration in lactose and microcrystalline cellulose. Assigned a value of 100%	Not applicable assigned a value of 100% and all other results calculated with reference to this result.
	Sample of lactose and microcrystalline cellulose	<2.0%
	Sample of Doxazisin in matrix plus available excipients (sodium lauryl sulphate and magnesium stearate).	95.0%<R<105.0%
Linearity:	Full Formulation – see tests for bias.	Bias meets acceptance criteria at extremes of range and at nominal concentration.
Precision: (repeatability)	Test full formulation (doxazosin in matrix plus available excipients) in triplicate at 50%, 100% and 150% of nominal concentration.	%Coefficient of Variation <5%
Bias:	Test above formulation in triplicate at 50%, 100% and 150% of nominal concentration. Repeatability test results can be used. Test two random batches of finished product.	95.0%<R<105.0% Further investigation is required if both batches give results less than 95.0% or greater than 105.0%. The mean of the results for both batches is between 95.0% and 105.0% of the nominal value.

<table>
<tr><th colspan="4">EVALUATION OF RESULTS - SPECIFICITY</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An evaluation of specificity is required by carrying out the method on:
1. a sample containing the analyte (doxazosin) at its nominal concentration in the matrix;
2. a sample of the matrix (microcrystalline cellulose and lactose);
3. a sample containing the analyte and available excipients at their nominal concentrations in the matrix;
Results from the matrix must be less than 2% and the potential interferents must affect results by no more than 5%.</td></tr>
<tr><td>Experiments</td><td colspan="3">A standard, used to prepare other samples by dilution, consisting of the analyte at double its nominal concentration in the matrix is required (i.e. 8 mg in 0.5 g). This is prepared by triturating 194 mg of doxazosin mesylate with 6.5 g of microcrystalline cellulose and 3.5 g lactose monohydrate (all weights recorded quantitatively). A matrix trituration is made by triturating 15.5 g of microcrystalline cellulose and 8 g lactose monohydrate. A double strength excipient trituration is prepared by triturating 20 mg sodium lauryl sulphate and 100 mg Magnesium Stearate with 10 g of the matrix trituration.
Samples are tested under repeatability conditions, in duplicate. All samples are made up quantitatively.
Sample 1 is prepared by mixing 2 g of the double strength doxazosin trituration with 2 g of the matrix trituration.
Sample 2 is the matrix trituration.
Sample 3 is prepared by mixing 2 g of the double strength doxazosin trituration with 2 g of the DS excipient trituration</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Measure the absorbance (A) at the wavelength of the maximum. For all samples calculate the absorbance per gram taken, by the formula $C = A/W_s$, where W_s is the weight of sample used for the test. Calculate the average for each sample (C_1 to C_3). Correct the value for C_3 to allow for differences in weights used for the preparation of Samples 1 and 3 by the formula $C_3*(W_{d1} /(W_{d1} + W_{m1}))/ (W_{d3} /(W_{d3} + W_{e3}))$, where ‘d’ represents DS doxazosin trituration, ‘m’ represents the matrix trituration and ‘e’ represents the excipient trituration and 1 and 3 represent samples 1 and 3.
To demonstrate specificity the exact concentration need not be known as the acceptability depends on the percentage effect of the contributors. Therefore the values C_{1-3} can be used to represent concentration.
Calculate the concentrations for Samples 2 and 3 relative to Sample 1, by the general formula $(C_x/C_1) * 100$.
Confirm that the values for concentration for Solutions 2 and 3 meet the acceptance criteria.</td></tr>
<tr><td>Results</td><td>Results:

C_2 = 0.48%
C_3 = 103.3%</td><td>Acceptance criteria:

< 2%
95.0%<R<105.0%</td><td>Pass/Fail:

Pass
Pass</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to improve specificity is required.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently specific for purpose.</td></tr>
</table>

EVALUATION OF RESULTS - LINEARITY			
Performance Criteria	An evaluation of linearity is required by carrying out the method on the formulation at the extremes of the stated range and at the nominal concentration. The results from bias testing will be used and the acceptance criteria for bias must be met.		
Experiments	See bias testing.		
Evaluation of data	See bias testing.		
Results	Results - see bias testing. Bias – meets criteria	Acceptance criteria: Meets bias criteria at extremes and nominal.	Pass/Fail: Pass
Discussion	Acceptance criteria are met therefore no further work to improve linearity is required.		
Conclusion	The method is considered sufficiently linear for purpose.		

<table>
<tr><th colspan="4">EVALUATION OF RESULTS - PRECISION</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An estimate of precision is required by carrying out replicate determinations on single batches of the full formulation.
Repeatability is estimated by carrying out the method in triplicate at 50%, 100% and 150% of the nominal concentration.
The repeatability must be 5% or better across the stated range.</td></tr>
<tr><td>Experiments</td><td colspan="3">The exact concentration of Doxazosin need not be known if the experiments are to be used solely for determining precision. If the results are also going to be used for bias testing, then the weight of doxazosin used must be recorded accurately and triturations made up quantitatively. Also the purity of the doxazosin and its uncertainty value need to be known (from a C. of A.). Samples 1, 2 and 3 are prepared by triturating 0.5, 2.0 and 1.5 g respectively of the double strength (DS) doxazosin trituration with 1.5, 2.0 and 0.5 g respectively of the matrix trituration. Sample 1 from the specificity test can be used as Sample 2 for precision.
All samples are tested on one occasion in triplicate. Each set of triplicate tests must be carried out by the same analyst on the same day and using the same equipment and reagents.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">To simplify the calculation, nominal values have been substituted for variables which do not affect the precision calculation (eg. weight of doxazosin used to prepare the standard). Calculate the nominal concentration for each test as a percentage of the theoretical value, using the formula:
$C = \frac{A_t \times (W_d + W_m) \times 100}{A_s \times 2 \times W_t \times 2 \times W_d}$
Where A_t and A_s are the absorbances of the test and standard, W_t is the weight taken for the test and W_d and W_m are the nominal weights of doxazosin DS and the matrix trituration used to prepare the samples (eg. for Sample 1, these are 0.5 and 1.5 respectively).
Enter the replicate results into the repeatability spreadsheet to determine the %CV.
Confirm that the values for %CV meet the acceptance criteria</td></tr>
<tr><td>Results</td><td>Results:

At 50%: Repeatability %CV = 2.34
At 100%: Repeatability %CV = 1.10
At 150%: Repeatability %CV = 1.61</td><td>Acceptance criteria:

=< 5%
=< 5%
=< 5%</td><td>Pass/Fail:

Pass
Pass
Pass</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to improve precision is required.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently precise for purpose.</td></tr>
</table>

<table>
<tr><th colspan="4">EVALUATION OF RESULTS – BIAS</th></tr>
<tr><td>Performance Criteria</td><td colspan="3">An estimation of bias is required across the range by carrying out determinations at 50%, 100% and 150% of the nominal concentration using solution with an accurately known concentration.
The bias must be no more than +/-5% across the concentration range.</td></tr>
<tr><td>Experiments</td><td colspan="3">If the samples used for precision testing were prepared quantitatively, with accurately recorded weights, the results can be used for the estimations of bias. If samples for bias testing need to be prepared, then the weight of doxazosin used must be recorded accurately and triturations made up quantitatively. Also the purity of the doxazosin and its uncertainty value need to be known (from a C. of A.). Sufficient must be prepared for all experiments, quantities quoted are for 10 g.
Samples 1, 2 and 3 are prepared by triturating 0.5, 2.0 and 1.5 g respectively of the double strength (DS) doxazosin trituration with 1.5, 2.0 and 0.5 g respectively of the matrix trituration. Sample 1 from the specificity test can be used as Sample 2 for bias. All samples are tested on one occasion in triplicate. Each set of triplicate tests must be carried out by the same analyst on the same day and using the same equipment and reagents.
Test two batches of tablets according to the full method.</td></tr>
<tr><td>Evaluation of data</td><td colspan="3">Calculate the content of doxazosin mesylate per gram of powder C using the method’s calculation but omitting AW.
Calculate the theoretical content C_t from the formula:
$C_t = \frac{D * P}{(D + W_c + W_l) * 100} * \frac{W_d}{(W_m + W_d)}$
Where D, W_c and W_l are the weights of Doxazosin, microcrystalline cellulose and lactose used in preparing the doxazosin DS; W_m and W_d are the weights of the matrix trituration and the doxazosin DS used to make up the samples and P is the % purity of the doxazosin.
Calculate the mean of the triplicate results and confirm that the bias values obtained by subtracting 100 from the mean of the results, meet the acceptance criteria.
Calculate the content of the batches of tablet as % of stated.</td></tr>
<tr><td rowspan="5">Results</td><td>Results:</td><td>Acceptance criteria:</td><td>Pass/Fail:</td></tr>
<tr><td>At 50%: Bias = -3.49%</td><td>< +/-5%</td><td>Pass</td></tr>
<tr><td>At 150%: Bias = -0.81%</td><td>< +/-5%</td><td>Pass</td></tr>
<tr><td>At 100%: Bias = -1.10%</td><td>< +/-5%</td><td>Pass</td></tr>
<tr><td>Tablets Batch 1 = 106.4% of stated
Tablets Batch 2 = 102.4% of stated</td><td>Both batches are not < 95.0 or both batches are not >105.0</td><td>Pass</td></tr>
<tr><td>Mean of 1 and 2 = 104.4% of stated</td><td>95.0%<R<105.0%</td><td>Pass</td></tr>
<tr><td>Discussion</td><td colspan="3">Acceptance criteria are met therefore no further work to reduce bias is required.</td></tr>
<tr><td>Conclusion</td><td colspan="3">The method is considered sufficiently free from bias for purpose.</td></tr>
</table>

SUMMARY OF RESULTS	
Specificity	Sample 1: Average absorbance per gram = 2.254 Sample 2: Average absorbance per gram = 0.0108 = 0.48% relative to Sample 1 Sample 3: Average absorbance per gram = 2.328 = 103.3% relative to Sample 1
Linearity	See results for Bias
Precision	Sample 1: Repeatability results = 94.16% of nominal stated 92.07% of nominal stated 89.86% of nominal stated Sample 2: Repeatability results = 97.08% of nominal stated 95.36% of nominal stated 97.31% of nominal stated Sample 3: Repeatability results = 95.92% of nominal stated 93.53% of nominal stated 93.10% of nominal stated
Bias	Sample 1: Results = 4.9527 mg = 98.74% of stated 4.8428 mg = 96.55% of stated Mean = 96.51% 4.7267 mg = 94.24% of stated Sample 2: Results = 10.2133 mg = 99.70% of stated 10.0319 mg = 97.93% of stated Mean = 99.19% 10.2368 mg = 99.93% of stated Sample 3: Results = 15.1368 mg = 100.73% of stated 14.7598 mg = 98.22% of stated Mean = 98.90% 14.6912 mg = 97.76% of stated Tablets Batch 1 = 4.25 mg = 106.3% of stated 4.25 mg = 106.3% of stated Mean = 106.4% 4.26 mg = 106.5% of stated Tablets Batch 2 = 4.10 mg = 102.5% of stated Mean = 102.4% 4.09 mg = 102.3% of stated

Repeatability Testing			
Enter triplicate results for each concentration (min)			
	Low Concentration	Nominal Concentration	High Concentration
Result 1 (required)	94.16	97.08	95.92
Result 2 (required)	92.07	95.36	93.53
Result 3 (required)	89.86	97.31	93.10
Result 4 (optional)			
Result 5 (optional)			
Number or results	3.00	3.00	3.00
Mean	92.03	96.58	94.19
Standard Deviation	2.15	1.07	1.52
RSD	0.0234	0.0110	0.0161
% Coefficient of variation	**2.34**	**1.10**	**1.61**

SUBMISSION FORM

ANALYTICAL METHOD VALIDATION

REPORT FORM FOR SUBMISSION TO NHS NATIONAL DATABASE

This form is to be used for the submission of analytical methods, which have been validated in accordance with the guidance issued by the NHS QA Committee, for inclusion in the national database of validated methods for use in the NHS.

On completion please submit forms to the national QA website www.nelm.nhs.uk/QA by logging on and selecting 'Contact Us'. Attach the completed forms to the email template.

Submitter details

Full address of submitting organisation: Licence number (if applicable): Date of submission:

Contact details: Name: Position: Department: Tel: email:

SUBMISSION FORM

Product details

Analyte (enter the molecule being analysed): Product/Substance (delete one)
Full formulation (for product):
Full method details (enter full method with calculation): Part of molecule being analysed if appropriate (eg. Chloride):

SUBMISSION FORM

Validation details

Method type (tick one box)

Method purpose	Full	Abridged
Identity		
Qualitative impurity test		
Quantitative impurity test		
Assay		

Characteristics examined (shading indicates characteristics not normally required for each method purpose).

Please tick all characteristics which were evaluated even if results were inferred from another test eg. if linearity is inferred from precision and bias testing over the range, tick linearity as well as precision and bias.

PERFORMANCE CHARACTERISTIC	IDENTITY	QUALITATIVE IMPURITY	QUANTITATIVE IMPURITY	ASSAY
Specificity				
Limit of Detection				
Limit of Quantitation				
Linearity				
Precision 1 – Repeatability				
Precision 2 – Intermediate				
Precision 3 – Reproducibility				
Bias				
Ruggedness testing				

Validation results

PERFORMANCE CHARACTERISTIC	ACCEPTANCE CRITERIA	RESULT
Specificity		
Limit of Detection		
Limit of Quantitation		
Linearity		
Precision 1 – Repeatability		
Precision 2 – Intermediate		
Precision 3 – Reproducibility		
Bias		
Ruggedness testing		