National Public Health Service for Wales
Gwasanaeth Iechyd Cyhoeddus Cenedlaethol Cymru

NHS

Health Protection Agency

NATIONAL STANDARD METHOD

GUIDANCE NOTE

UNCERTAINTY OF MEASUREMENT IN TESTING

QSOP 4

Issued by Standards Unit, Evaluations and Standards Laboratory
Centre for Infections

STATUS OF NATIONAL STANDARD METHODS

National Standard Methods, which include standard operating procedures (SOPs), algorithms and guidance notes, promote high quality practices and help to assure the comparability of diagnostic information obtained in different laboratories. This in turn facilitates standardisation of surveillance underpinned by research, development and audit and promotes public health and patient confidence in their healthcare services. The methods are well referenced and represent a good minimum standard for clinical and public health microbiology. However, in using National Standard Methods, laboratories should take account of local requirements and may need to undertake additional investigations. The methods also provide a reference point for method development.

National Standard Methods are developed, reviewed and updated through an open and wide consultation process where the views of all participants are considered and the resulting documents reflect the majority agreement of contributors.

Representatives of several professional organisations, including those whose logos appear on the front cover, are members of the working groups which develop National Standard Methods. Inclusion of an organisation's logo on the front cover implies support for the objectives and process of preparing standard methods. The representatives participate in the development of the National Standard Methods but their views are not necessarily those of the entire organisation of which they are a member. The current list of participating organisations can be obtained by emailing standards@hpa.org.uk.

The performance of standard methods depends on the quality of reagents, equipment, commercial and in-house test procedures. Laboratories should ensure that these have been validated and shown to be fit for purpose. Internal and external quality assurance procedures should also be in place.

Whereas every care has been taken in the preparation of this publication, the Health Protection Agency or any supporting organisation cannot be responsible for the accuracy of any statement or representation made or the consequences arising from the use of or alteration to any information contained in it. These procedures are intended solely as a general resource for practising professionals in the field, operating in the UK, and specialist advice should be obtained where necessary. If you make any changes to this publication, it must be made clear where changes have been made to the original document. The Health Protection Agency (HPA) should at all times be acknowledged.

The HPA is an independent organisation dedicated to protecting people's health. It brings together the expertise formerly in a number of official organisations. More information about the HPA can be found at www.hpa.org.uk.

The HPA aims to be a fully Caldicott compliant organisation. It seeks to take every possible precaution to prevent unauthorised disclosure of patient details and to ensure that patient-related records are kept under secure conditions[1].

More details can be found on the website at www.evaluations-standards.org.uk. Contributions to the development of the documents can be made by contacting standards@hpa.org.uk.

Please note the references are now formatted using Reference Manager software. If you alter or delete text without Reference Manager installed on your computer, the references will not be updated automatically.

Suggested citation for this document:
Health Protection Agency (2005). *Uncertainty of Measurement in testing*. National Standard Method QSOP 4 Issue 5. http://www.hpa-standardmethods.org.uk/pdf_sops.asp.

INDEX

STATUS OF NATIONAL STANDARD METHODS 2
INDEX 3
AMENDMENT PROCEDURE 4
1 INTRODUCTION 5
2 DEFINITIONS 5
3 REASONS FOR EVALUATING UNCERTAINTY OF MEASUREMENT 6
4 GENERAL PRINCIPLES 6
5 GENERAL SOURCES OF UNCERTAINTY 7
6 SPECIFIC CONTRIBUTORY FACTORS TO UNCERTAINTY IN MICROBIOLOGY 7
7 SOURCES OF INFORMATION ON UNCERTAINTY 9
8 SAMPLE MATRICES 9
9 EVALUATION OF UNCERTAINTY OF MEASUREMENT FOR MICROBIOLOGICAL TESTING 10
10 LABORATORY BIAS 12
11 CONCORDANCE 12
12 EXPRESSION OF RESULTS 13
13 ASSESSMENT OF COMPLIANCE WITH SPECIFICATION 13
14 CONTACT 14
15 TABLE 1 STUDENT *t* DISTRIBUTION FOR THE 0.975 QUANTILE 15
16 APPENDIX 1: THE CALCULATION OF RELATIVE STANDARD DEVIATIONS OF REPRODUCIBILITY (RSD_R) TO ASSESS THE UNCERTAINTY OF MEASUREMENT FOR COUNTS 16
REFERENCES 21

AMENDMENT PROCEDURE

Controlled document reference	**QSOP 4**
Controlled document title	**Uncertainty of Measurement in Testing**

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

On issue of revised or new pages each controlled document should be updated by the copyholder in the laboratory.

Amendment Number/ Date	Issue no. Discarded	Insert Issue no.	Page	Section(s) involved	Amendment
5/ 30.08.05	4.1	5	1	**Front page**	Redesigned
			2	**Status of document**	Reworded
			4	**Amendment page**	Redesigned
			All	**All**	Whole document updated

GUIDANCE NOTE ON THE UNCERTAINTY OF MEASUREMENT IN TESTING

1 INTRODUCTION

1.1 This Guidance Note has been produced to help laboratories understand and comply with the requirements of BS EN ISO/IEC 17025[2] for estimation of uncertainty of measurement. This standard states that testing laboratories shall have and apply procedures for estimating uncertainty of measurement.

1.2 ISO 17025 recognises that the nature of the test method may preclude rigorous, metrologically and statistically valid, calculation of uncertainty, but in these cases the laboratory should at least attempt to identify the components of uncertainty and make a reasonable estimation. Further clarification on the requirements for microbiology laboratories is given in document EA-4/10 Accreditation for Microbiological Laboratories[3] and EA 4/16 EA Guidelines on the Expression of Uncertainty in Quantitative Testing[4].

1.3 General guidance on estimation of uncertainty is available in M3003[5], The Eurachem/CITAC Guide[6] and LAB 12[7]. More specific guidance for food microbiological testing has been published by Niemela[8], CCFRA[9] and by the British Standards Institution[10] for water microbiological testing.

1.4 Laboratories need to produce data on uncertainty of measurement (UM) for the tests that they perform and may be required to report relevant information in certain circumstances such as:

- when a client's instructions require a statement of uncertainty
- when it is required by the test specification
- when the uncertainty is relevant to the validity or application of the result; eg when the uncertainty affects compliance with a specification

1.5 Laboratories need to have a defined policy covering the provision of estimates of the uncertainties of the tests performed. When a statement of uncertainty is required by the client or specification, the laboratory should use documented procedures for the estimation, treatment and reporting of the uncertainty.

2 DEFINITIONS

The standard uncertainty of a test method is the uncertainty of the result expressed as a standard deviation.

The combined standard uncertainty is the result of the combination of uncertainty components.

The expanded uncertainty is obtained by multiplying the combined standard uncertainty by a coverage factor.

Coverage factor – a numerical factor used as a multiplier of the combined standard uncertainty in order to obtain the expanded uncertainty (typically a value of 2 – 3).

Sensitivity - the percentage of samples correctly found to be positive

Specificity – the percentage of samples correctly found to be negative

Bias – the difference between the expectation of the test results and an accepted reference value.

Repeatability – the closeness of the agreement between successive and independent test results obtained by the same method on portions of the same test sample under the same conditions (apparatus, operator, laboratory and within a short period of time)

Reproducibility – the closeness of the agreement between test results obtained by the same method on portions of the same test sample under different conditions (operator, equipment, laboratory)

3 REASONS FOR EVALUATING UNCERTAINTY OF MEASUREMENT

3.1 Uncertainty of measurement is a quantitative indication of the analytical variability of a result. It demonstrates how well the result represents the value of the quantity being measured in the test portion. It also allows an assessment of the reliability of the result, for example in comparing the results from different sources or with reference values or criteria given in microbiological standards, guidelines or specifications. This may be particularly important when interpreting the results of official control and formal samples (ie: those with potential legal implications).

3.2 The uncertainty may need to be taken into account when interpreting data. For example, observed differences in results may be accounted for by the uncertainty associated with the results rather than real differences in properties or performance. Thus if two competent laboratories examine different subsamples from the same sample source by the same method and obtain numerically different results, these results may not be different when UM is taken into account.

3.3 An evaluation, or at least a full consideration, of all the identifiable components that contribute to the uncertainty of a test result will allow valid results to be obtained and will indicate the aspects of the test that require attention to improve procedures.

3.4 Systematic assessment of the factors influencing the result and of the uncertainty forms a key part of method validation.

4 GENERAL PRINCIPLES

4.1 No attempt is made in this guidance note to assess sample variation, although laboratories are expected to have an understanding of the distributions of organisms within matrices they test and to take this into account when subsampling[3].

4.2 In general, no measurement or test is perfect and the imperfections (such as pipettor variation) may give rise to error of measurement in the result. Thus in an enumeration test the bacterial count obtained is only an approximation of the actual count and so the result is incomplete unless accompanied by a statement of the uncertainty associated with the count.

4.3 Uncertainty of measurement is a parameter associated with the result of a measurement that characterises the dispersion of values that could reasonably be attributed to the measurand (such as a count/g).

4.4 As uncertainty of measurement is a parameter associated with the result of a measurement it is only applicable to enumeration methods. However, factors that contribute to the uncertainty of the result of presence/absence tests should be identified.

4.5 In microbiological testing the greatest sources of uncertainty arise from sampling and the non-homogeneous distribution of micro-organisms in the sample.

4.6 There are two types of evaluations that may be used to calculate uncertainty:

Type A evaluations of uncertainty are done by calculations from a series of repeated observations, using statistical methods.

Type B evaluations of uncertainty are derived from other sources eg from calibration data.

When carrying out a microbiological test, the factors contributing to a type B evaluation usually form part of a type A evaluation and so may not need to be considered separately. They also usually represent such a small contribution to the combined standard uncertainty that they do not make a significant contribution. For most microbiological testing purposes, the combined standard uncertainty in a type A evaluation is not significantly different from the standard uncertainty; the type B components can therefore be ignored in the calculation but identified and shown to be under control.

4.7 To simplify the evaluation of uncertainty it is necessary to exclude "less than" results and to use counts over 100 /g because logs of such results are usually distributed in an approximately normal fashion.

5 GENERAL SOURCES OF UNCERTAINTY

5.1 Sources of measurement uncertainty include:

- incomplete definition of a test - the requirement is not clearly defined, eg a temperature range may be given for incubation
- the effects of environmental conditions on the measurement process, or imperfect measurement of environmental conditions
- personal bias in reading instruments
- instrument resolution or threshold of discrimination
- errors in the graduation of a scale
- values assigned to measurement standards (both reference and working) and reference materials
- changes in the characteristics or performance of a piece of equipment since the last calibration
- calibration factors not being applied
- values of constants and other parameters used in data analysis
- approximations and assumptions incorporated in the test method and procedure
- technical competence of the individual performing the test

5.2 All possible sources of UM are not necessarily independent. In reality they may overlap and inclusion of components that would contribute to a type B evaluation might result in an overestimation of UM.

Unrecognised systematic effects may also exist that cannot be taken into account but contribute to error. Such effects may be deduced from examination of results of an inter-laboratory comparison programme (proficiency testing).

6 SPECIFIC CONTRIBUTORY FACTORS TO UNCERTAINTY IN MICROBIOLOGY

6.1 Technical competence, bias and experience

- all stages of processing a sample, operation of equipment and qualitative or quantitative reading of tests
- variation between and within members of staff

6.2 Sample

- homogeneity of original sample source
- sample sent to a laboratory

- test portion used in the subsample analysis
- precision and accuracy of balance or volumetric equipment
- non-uniform distribution of micro-organisms between subsamples or test portions
- time, transport and storage conditions between sampling and testing

6.3 Homogenisation of sample

- degree of heterogeneity of suspensions made from the sample
- clumping of micro-organisms
- uneven distribution of micro-organisms
- insufficient mixing

6.4 Dilutions

- accuracy of pre- measured volumes or weights of dilution fluids
- volume of dilution fluid used
- degree of mixing at each dilution step
- number of steps in a serial dilution
- precision, accuracy and appropriate use of diluting equipment
- pipette volume used
- micro-organisms adhering to pipettes

6.5 Media and reagents

- quality of raw materials
- accurate weighing of materials
- water quality including pH and conductivity
- personal error in preparation and use of culture media (including appropriate temperature when adding supplements)
- heat processing and control
- adequate mixing
- degree of dryness of solid media
- performance of media and reagents such as selectivity and sensitivity
- shelf life

6.6 Inoculation of media

- volume of inoculum
- equipment used in dispensing, spreading and filtering
- temperature of molten agar in pour plate techniques

6.7 Incubation conditions

- duration
- temperature
- humidity

- atmospheric conditions

6.8 Reading and interpretation of results

- recognition of target colonies
- number of colonies counted
- dilutions chosen for counting (one dilution or more than one dilution)
- proportion of colonies confirmed
- properties of media especially when using automated counters

7 SOURCES OF INFORMATION ON UNCERTAINTY

7.1 There are many sources of information on uncertainty associated with test methods including:

- results from inter-laboratory validations published in the scientific press and international standards
- published data on the effects of sample storage
- published and in-house data on test performance
- in-house validation data including shelf life trials
- manufacturers' performance data
- internal quality control data which may include duplicate testing using the same member of staff and using different members of staff, as well as processing spiked samples
- data from external quality assessment schemes (proficiency schemes)
- calibration certificates and in-house calibration data for relevant equipment

8 SAMPLE MATRICES

8.1 It is important to consider the sample matrix as this may affect the uncertainty of measurement. For example a liquid sample such as water or milk will show greater homogeneity than a solid sample such as a composite meal.

8.2 Most foods received in HPA laboratories are ready to eat and have often been cooked. They are examined by similar methods and it is therefore feasible to determine UM for a test method using all the different food matrices together. Matrices might include high fat products, high protein products, processed food such as ready meals, vegetables, raw foods and so on. The selection of matrices should be representative of the laboratory's typical workload and at least five types of matrix included.

8.3 If it becomes clear that a particular matrix gives widely varying results (several logs) between duplicate sets of counts the UM for this matrix should be evaluated separately. This may happen, for example, with a multicomponent food such as a pasta or rice salad containing raw and cooked ingredients.

8.4 Water samples are usually examined by different enumeration methods to food methods and should be considered separately. This group of samples also consists of several types, for example potable water, surface water, bathing water, pool water, cooling tower water. All these types have different chemical properties and microbiological profiles. Guidance is available[11] on preparation of appropriate spiked samples of potable water that may be used in evaluation of uncertainty.

9 EVALUATION OF UNCERTAINTY OF MEASUREMENT FOR MICROBIOLOGICAL TESTING

9.1 For most laboratories the sampling procedure to provide laboratory samples is not part of the test procedure and the sample is "as received in the laboratory". It should then be made clear whether a subsampling step to provide the test portion is included in the test procedure. For most tests only part of the sample submitted to the laboratory is examined, but sometimes, eg for some water tests, the entire laboratory sample may be examined, in which case the distribution of the organisms within the matrix is less likely to affect the outcome of the result.

9.2 For each test, specify the measurand. This requires a statement of what is being measured and the conditions of testing. Eg: for an aerobic colony count, this statement is "the calculation of the number of bacteria, yeasts and moulds within the 25 g test portion (if this is the start point) or within the 10^{-1} dilution of the test portion (if the sample homogenate is the start point) which are capable of growing aerobically under the test conditions of an incubation period of 48 ± 2 hours at 30°C ± 1°C".

9.3 For each test, list the components that contribute to the uncertainty of the result and establish that they are under control.

9.4 General enumeration methods

a) identify the individual components of uncertainty

b) demonstrate that they are under control

c) state that the individual components are inherent within the evaluation process

d) identify the start point for uncertainty evaluations

e) use repeatability and reproducibility data as the basis for expression of uncertainty

The individual components of uncertainty should be identified and demonstrated to be under control, for example by regular checking of equipment performance and remedial action if outside specification. It is useful to record the tolerances allowed by the laboratory for the various measurements forming part of the test, for example ± 2% for diluent volumes, ± 5% for inoculum volumes. Generally for microbiological evaluation of uncertainty the individual components do not contribute a significant amount. The most likely component to contribute significantly is an extended dilution series (>5 dilution steps), but under controlled conditions even this may not make a significant contribution.

Establish the start point of the uncertainty evaluation. If natural samples are used (eg: water) the start point may be the test portion. If spiking is required, the start point may be the test portion homogenate. If the test portion used to prepare the homogenate or dilution is the start point in the procedure, the start point is as received by the laboratory after the appropriate storage prior to laboratory sampling. For laboratories using a gravimetric diluter in the preparation of their sample homogenates the contribution of the homogenate preparation should be negligible and artificial inocula may be added to the sample homogenate. This then becomes the start point for spiked samples. If however the laboratory weighs their sample to fall within a specified weight range rather than an exact weight and then adds a volume of diluent calculated for the target weight rather than the exact weight, there may be a significant contribution to uncertainty and this should be evaluated.

9.5 Evaluation of uncertainty using test results

International consensus supports the use of reproducibility and repeatability data in the evaluation of uncertainty in food microbiology, recognising the difficulties in calculating the contributions of individual components. However this will not determine laboratory bias. This may be assessed by using proficiency test samples from external quality assessment schemes.

It is preferable to use real samples for provision of data although this may not always be realistic. However every attempt should be made to cover the full range of sample matrices examined by the laboratory. All the data from all the sources (actual samples, spiked samples, EQA samples) may be used in the uncertainty assessment. In general when using spiked food samples the level of the spike should be such that around one hundred colonies are counted.

Sample examination would normally be done during routine laboratory operation and this would reflect all the operating factors contributing towards uncertainty. The duplication should normally be performed by different operators (reproducibility). Ideally the replication should be performed using as many variables as possible, for example different operators, different pipettors, different incubators operating at the same temperature, different batches of media. This will help to reflect the variations that may occur naturally in the laboratory.

Replicate (2 or more) results from each sample can then be entered into an EXCEL spreadsheet and the data analysed to obtain the reproducibility standard deviation RSD_R (see appendix 1 for details). Results to at least three significant figures should be used in order to minimise rounding errors. Levels of similar magnitude should be used, for example $10^2 – 10^4$ for indicator parameters or $10^4 – 10^6$ for aerobic colony counts. Further results should be added to the spreadsheet on a regular and rolling basis, so that the laboratory's ongoing performance can be assessed in a contemporaneous way. At least 30 sets of samples are required in order to apply a coverage factor of 2 to obtain the expanded uncertainty; below this number the coverage factor should be determined by using the Student *t* value for the number (n) of sample sets and the 0.975 quantile (two-tailed 95% test) with n-1 degrees of freedom. These values are shown in Table 1. Note that "less than" values (below the limit of detection) cannot be included.

Assess the uncertainty at least annually to ensure that it still reflects the conditions in the laboratory such as staffing and equipment in use. If the uncertainty is required by a client or for an interpretation of a result against a legal standard and the calculation has not been performed for over six months it is advisable to update the calculation. If there is a significant change in the operating conditions of the laboratory, such as significant staff changes or changes of equipment, or if significant changes are made to the method used (such as a change of plating medium), the effect of these changes on UM should be evaluated.

9.6 Contribution of multiple replicate plating

Sets of replicate results from a number of samples can be pooled together to calculate a pooled variance and pooled standard deviation. This is often done as a procedure separate to the normal operation of the laboratory, but can produce useful information on an operator's repeatability (within operators) and reproducibility between operators. This may detect certain biases such as differences between workers or equipment but will not detect laboratory bias. Detectable variation is likely to be due to both UM and the random variation due to the distribution of the organisms in the test portions of the sample examined[10].

9.7 Presence/absence tests

Presence/absence tests do not result in an enumeration; therefore uncertainty of measurement cannot be calculated using the above approach. However the factors affecting the variability of the results must be considered and shown to be under control. For each pathogen or organism of concern the lower limit of detection that can be achieved consistently should be determined and the sensitivity and specificity of the test calculated.

It is not normally realistic to use naturally contaminated samples for these determinations as they are rarely encountered in the routine operation of the laboratory. However if naturally contaminated samples are identified they may be repeated in replicate. Different sample matrices should be used for spiking to encompass the types received by the laboratory. The uncertainty of the inoculum size may need to be considered when assessing the limit of detection.

9.8 Most Probable Number (MPN) tests

If contributory components can be shown to be under control, the theoretical 95% confidence intervals given in tables can be used to provide the expanded uncertainty range. Results obtained using EQA samples should be reviewed to establish that they are within 95% confidence limits at least 95% of the time.

9.9 Membrane filtration tests

These tests may either be enumeration tests or presence/absence tests, depending on the specification for the test and the target organism. Additional contributory factors specific to this type of test should be identified. Spiked samples and naturally contaminated samples of water containing 20-50 colonies per test volume[11] should be used when possible for calculating the uncertainty of a membrane filtration enumeration method. Naturally contaminated water may be used to spike dechlorinated tap water to avoid the use of organism cultures. Representatives of the different types of water examined (drinking, river, surface, chlorinated pool etc) should be included. For low numbers of target organisms the method may be treated as a presence/absence method and the lower limit of detection determined.

9.10 Phosphatase testing

Phosphatase testing measures levels of the enzyme mammalian alkaline phosphatase, which is likely to be evenly dispersed through a well-mixed sample (unlike micro-organisms). Because it is a biological product the greatest sources of uncertainty are animal to animal variation, seasonal variation and the age of the sample since production. A similar approach can therefore be used to assess UM as for microbiological tests. Uncertainty can be determined using replicate test results in which as many operators as possible have examined aliquots of the same sample or by using duplicate sets of results for a series of samples as described in section 9.5. Appropriate samples are EQA samples or routine samples containing significant levels (>100 mU/L) of phosphatase activity.

9.11 Confirmation/ Identification tests

These are qualitative tests and so the concept of uncertainty measurement cannot be applied directly. However the individual sources of variability should be identified and shown to be under control. As the proportion of suspect colonies that are submitted to identification tests increases the associated uncertainty diminishes, but a pragmatic view needs to be taken on the number of such colonies identified and this is specified in each method.

10 LABORATORY BIAS

Laboratory bias can only be assessed by review of EQA results, results from inter-laboratory collaborative trials or use of certified reference materials. If bias is detected the reasons for it require investigation. Any bias that is present can be included in the uncertainty evaluations by including the results from proficiency testing schemes and other appropriate collaborative trials

11 CONCORDANCE

Presence/absence testing does not result in numerical values that can be used to calculate uncertainty. However reproducibility can be determined for presence/absence tests by calculating the concordance[12]. Taking each organism sought by presence/absence testing separately, list the laboratory's results in EQA and IQC samples and compare them with the

expected result. Include the negative results as well as the positive results. The concordance is the percentage of all pairings giving the same results. Record the concordance achieved. The time frame for the results used should be sufficiently contemporaneous to reflect the conditions in the laboratory such as the staffing and equipment in use.

12 EXPRESSION OF RESULTS

12.1 When reporting the result of a test the extent of information given depends mainly on the client's requirements, the specification and the intended use of the result. The methods used to assess the results should be available either in the report or in the records of the test.

Information should include sufficient documentation to demonstrate how the uncertainty was obtained and enable a repeat of the calculation if necessary.

12.2 When reporting enumeration tests the result of the test should be reported together with the expanded uncertainty appropriate to the 95% level of confidence, and expressed to no more than two significant figures. For most enumeration tests $\log_{10}$ values should be used. These results should be accompanied by the following statement if at least 30 sets of data have been used and a normal distribution can be assumed:

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor of $k = 2$ (or other appropriate t *value), providing a level of confidence of approximately 95%.*

12.3 For fewer sets of data and if the conditions of normality apply a *t* distribution may be assumed. In this case the following statement should be used:

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor k = XX, which for a t *distribution with YY effective degrees of freedom provides a level of confidence of approximately 95%*

12.4 The conditions for assuming a normal distribution are explained in EA 4/16[4].

12.5 For further guidance on special cases, see LAB 12[7]

a) when reporting results of MPN tests, report the test result and the 95% confidence limits (obtained from tables)

b) for presence/absence tests, report the sensitivity and specificity or concordance together with the lowest level of organisms that can be recovered consistently.

13 ASSESSMENT OF COMPLIANCE WITH SPECIFICATION

13.1 If a statement of compliance is required, for example against a numerical limit specified in a standard, the count obtained in the test should be expanded by the uncertainty interval at a level of confidence of 95% before comparison with the numerical limit.

13.2 For microbiological tests and phosphatase tests, maximum limits are usually specified.

Compliance is achieved if the expanded count is lower than the maximum limit. A statement of compliance with the specification can be made.

If the maximum limit is exceeded even when the measured count is decreased by half the uncertainty interval, a statement of non-compliance with the specification can be made.

If the measured count exceeds the maximum limit but when decreased by half the uncertainty interval it is less than the maximum limit, it is not possible to confirm compliance or non-compliance with the specification. The test result and expanded uncertainty should be reported together with a statement that compliance was not demonstrated. If the measured result is below the maximum limit but when expanded exceeds the maximum limit a similar statement should be made. For example:

The measured result is above (below) the specification limit by a margin less than the measurement uncertainty; it is therefore not possible to state compliance based on the 95%

level of confidence. However the result indicates that non-compliance (compliance) is more probable than compliance (non-compliance) with the specification limit[7].

14 CONTACT

Any queries relating to this Guidance Note should be made to:

Melody Greenwood, Laboratory Director, Wessex Environmental Microbiology Services.
Andre Charlett, Head of Statistics Unit, Statistics, Modelling and Bioinformatics Department, CFI
Ruhi Siddiqui, Acting Head of Standards Unit, Evaluations & Standards Laboratory, CFI

15 TABLE 1 STUDENT *t* DISTRIBUTION FOR THE 0.975 QUANTILE

Degrees of freedom*	0.975 quantile
1	12.71
2	4.303
3	3.182
4	2.776
5	2.571
6	2.447
7	2.365
8	2.306
9	2.262
10	2.228
11	2.201
12	2.179
13	2.160
14	2.145
15	2.131
16	2.120
17	2.110
18	2.101
19	2.093
20	2.086
21	2.080
22	2.074
23	2.069
24	2.064
25	2.060
26	2.056
27	2.052
28	2.048
29	2.045
30	2.042

* degrees of freedom = n-1 data sets

16 APPENDIX 1: THE CALCULATION OF RELATIVE STANDARD DEVIATIONS OF REPRODUCIBILITY (RSD_R) TO ASSESS THE UNCERTAINTY OF MEASUREMENT FOR COUNTS

This appendix describes calculations of the relative standard deviation of reproducibility (RSD_R) using a type A evaluation to estimate the uncertainty of measurement for counts.

Duplicate counts

If duplicate counts are made on the same test material by two different operators, then they can be used to provide an estimate of the RSD_R. While the calculations can be performed on the raw counts it is usual to take common logarithms of the count. Note that common logarithms are logarithms to the base 10 e.g. $\log_{10}100 = 2.00$

If we consider for example that for operators, A and B, counts of 1089 and 1211 were obtained respectively, when testing the same material. The RSD_R is calculated as follows:

Take common logarithms of the counts

$x_A = \log_{10}(1089) = 3.03702787976$

and $x_B = \log_{10}(1211) = 3.08314414314$

The RSD_R is calculated using the following formula

$$RSD_R = \sqrt{2\left(\frac{(x_A - x_B)}{(x_A + x_B)}\right)^2} \qquad \text{… (1)}$$

and substituting x_A and x_B into equation (1) we obtain

$$\sqrt{2 \cdot \left(\frac{(3.037028 - 3.083144)}{(3.037028 + 3.083144)}\right)^2} = \sqrt{2 \cdot \left(\frac{-0.046116}{6.120172}\right)^2} = 0.01065627$$

It is not immediately clear that equation (1) is the usual estimate of the RSD_R, i.e. standard deviation divided by the mean. It is shown below that this is so.

The general formula for the arithmetic mean ($\bar{x}$) is:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \text{… (2)}$$

and for the pair of observations, x_A and x_B, this is: $\bar{x} = \frac{(x_A + x_B)}{2}$ … (3)

The general formula for the standard deviation (s) is:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \qquad \text{… (4)}$$

For the pair (n=2) of observations x_A, and x_B, substitution into equation (3) gives:

$$\sqrt{\frac{\left(x_A - \frac{x_A + x_B}{2}\right)^2 + \left(x_B - \frac{x_A + x_B}{2}\right)^2}{2-1}}$$, which can be simplified as follows

$$\sqrt{\left(\left(\frac{x_A}{2} - \frac{x_B}{2}\right)^2 + \left(\frac{x_B}{2} - \frac{x_A}{2}\right)\right)^2} = \sqrt{\frac{(x_A - x_B)^2}{4} + \frac{(x_B - x_A)^2}{4}}$$

as $(x_A - x_B)^2 = (x_B - x_A)^2$, the equation simplifies further to

$$\sqrt{\frac{2(x_A - x_B)^2}{4}} = \sqrt{\frac{(x_A - x_B)^2}{2}} = \frac{(x_A - x_B)}{\sqrt{2}} \qquad \dots (5)$$

The RSD_R is defined as

$$RSD_R = \frac{s}{\bar{x}} \qquad \dots (6)$$

Substitution of (5) and (3) into equation (6) and then simplifying gives:

$$RSD_R = \frac{\frac{(x_A - x_B)}{\sqrt{2}}}{\frac{(x_A + x_B)}{2}} = \sqrt{2}\left(\frac{(x_A - x_B)}{(x_A + x_B)}\right) = \sqrt{2 \cdot \left(\frac{(x_A - x_B)}{(x_A + x_B)}\right)^2} \qquad \dots (7)$$

The formula for the RSD_R as obtained in equation (7) is identical to equation (1).

It is often the situation that the RSD_R is multiplied by 100 to obtain a notional percentage; in this example the RSD_R is then expressed as 1.0656%.

It is not recommended that just a single pair of results from two different operators be used to estimate the laboratory specific RSD_R. To obtain what could be considered as a “typical” laboratory specific RSD_R a reasonably large number of pairs of results would be required. It is difficult to recommend exactly how many pairs of counts are required to have a reasonably precise estimate of the average RSD_R without additional information. Ideally at least 30 sets of paired counts should be used, but a reasonable estimate may be obtained using 10 sets of paired counts.

Obtaining a “typical” RSD_R

Each of the individual RSD_R will provide an estimate of the actual laboratory specific reproducibility. However, there is likely to be some variation between these estimates. One often-used approach used to provide a point estimate is to obtain the quadratic mean or root mean square of the individual RSD_R

estimates. To do this, each RSD_R is squared and the arithmetic average is then taken, and the result is square rooted to give an estimate of the combined reproducibility RSD_{RC}.

$$RSD_{RC} = \sqrt{\frac{\sum_{i=1}^{n} RSD_{R_i}^2}{n}} \quad \ldots (8)$$

Example of the calculation of RSD_{RC} using four RSD_R obtained from pairs of counts

First count (A)	Second count (B)	$\log_{10}A$	$\log_{10}B$	$\log_{10}A - \log_{10}B$	$\log_{10}A + \log_{10}B$	RSD_R
1089	1211	3.037028	3.083144	-0.046116	6.120172	0.010656
122000	142000	5.08636	5.152288	-0.06593	10.23865	0.009106
32500	29000	4.511883	4.462398	0.049485	8.974281	0.007798
28000	35020	4.447158	4.544316	-0.097158	8.991474	0.015281

Substituting these four values into equation (8) we obtain:

$$RSD_{RC} = \sqrt{\frac{\sum_{i=1}^{n} RSD_{R_i}^2}{n}} = \sqrt{\frac{0.010656^2 + 0.009106^2 + 0.007798^2 + 0.015281^2}{4}}$$

$$= \sqrt{\frac{0.0001135503 + 0.0000829192 + 0.0000608088 + 0.0002335090}{4}} = \sqrt{\frac{0.0004907873}{4}}$$

and thus, RSD_{RC} = 0.0110768604, i.e. 0.0111

More than two operators

For some laboratories, the data upon which the laboratory specific reproducibility is estimated will not necessarily be pairs of counts but may be replicated counts made by more than two operators. An extension to the method described above is then required. If we assume that up to *n* operators may evaluate the test material, then we can represent the resultant counts as below:

Test material	Operator					
	A	B	C	D	…	N
1	C_{A1}	C_{B1}	C_{C1}	C_{D1}	…	C_{n1}
2	C_{A2}	C_{B2}	C_{C2}	C_{D2}	…	C_{n2}
…						
K	C_{Ak}	C_{Bk}	C_{Ck}	C_{Dk}	…	C_{nk}

Clearly, it is not always the situation that all results will be obtained and some of the counts will be missing for a variety of reasons. However, for each separate test material the standard deviation and mean can be calculated using equations (4) and (2) respectively, after taking common logarithms of the counts. For example, four operators estimated the colony count on the same test material to be 5532, 6012, 6060, and 6150. After taking common logarithms, these are 3.742882, 3,779019, 3.782473, and 3.788875, respectively. Substitution of these into equation (2) gives:

$$\bar{x} = \frac{(3.742882 + 3.779019 + 3.782473 + 3.788875)}{4} = \frac{15.093249}{4} = 3.773312$$

and substitution in (4) gives:

$$s = \sqrt{\frac{(3.742882 - 3.773312)^2 + (3.779109 - 3.773312)^2 + (3.782473 - 3.773312)^2 + (3.788875 - 3.773312)^2}{3}}$$

thus s = 0.0206935, and:

$$RSD_R = \frac{s}{\bar{x}} = \frac{0.0206935}{3.773312} = 0.005484 \approx 0.0055$$

Again it would not be sensible to use an estimate of the RSD_R from a single test material, and a "typical" laboratory specific RSD_R can be obtained by calculating the quadratic mean of a large set of RSD_R obtained from different test material.

Example

Test material	Operator				(after taking common logarithms)		
	A	B	C	D	sd	mean	RSD_R
1	5532	6012	6060	6150	0.0206935	3.773312	0.005484
2	564	313	.	412	0.1279618	2.620574	0.048830
3	3021	5123	.	.	0.1621917	3.594838	0.045118
4	.	10432	9121	9560	0.0295967	3.986289	0.007425

Substituting these four values into equation (8) we obtain;

$$RSD_{RC} = \sqrt{\frac{\sum_{i=1}^{n} RSD_{R_i}^2}{n}} = \sqrt{\frac{0.005484^2 + 0.048830^2 + 0.045118^2 + 0.007425^2}{4}}$$

$$= \sqrt{\frac{0.000030076 + 0.002384339 + 0.002035629 + 0.000055115}{4}} = \sqrt{\frac{0.004505169}{4}}$$

and thus, RSD_{RC} = 0.033560279, i.e. 0.0336

Application of the RSD_{RC} in uncertainty of measurement (UM)

The RSD_{RC} as calculated above is the average relative standard deviation of reproducibility, after common logarithms ($\log_{10}$) of the observed counts have been taken. To obtain the UM for a particular count, *c*, then use equation (9),

$$UM = \log_{10}(c) \pm k \cdot RSD_{RC} \quad \ldots (9)$$

where *k* is the appropriate coverage factor, usually 2.

This will produce a range of the common logarithms of the count after UM has been applied. This range is that which could reasonably be expected if this material were to be tested by different operators within this laboratory.

To obtain the UM limits on the natural count scale, the antilog of these two $\log_{10}$ values can be taken to obtain UM limits on the count scale. This can either be done using computer software or a hand-held calculator, the mathematical formula for the antilog is:

$\text{antilog}(v) = 10^v$ e.g. if the resultant quantity was 5.034, then the antilog would be $10^{5.034} \approx 108143$

Example

If the RSD_{RC} is 0.011 and the count is 6.76 x 10^4, and we assume a coverage factor of k = 2. The common logarithm of 6.76 x 10^4 is 4.8299. The upper and lower limits are:

4.8299±2.0.011 = 4.8299±0.022 = 4.8079 to 4.8519

If a range for counts were required, it would be 6.43 x 10^4 – 7.11 x 10^4.

Note that this results in an asymmetric range around the count value.

Important points to note

- Only use actual counts, so exclude any results that are either ‘less than’ or ‘greater than’ values.
- During any intermediate stages in the calculations, e.g. when transforming counts to $\log_{10}$ values, calculating the mean, etc try and keep figures as accurate as possible and only round the final results to the desired precision.
- It is useful to include columns in the Excel spreadsheet that allow the date of testing, sample matrix and identity of the operators to be recorded.

Excel worksheets are available on the Standards Unit Website at http://www.hpa-standardmethods.org.uk/documents/qsop/pdf/rsd.xls

REFERENCES

1. Department of Health NHS Executive: The Caldicott Committee. Report on the review of patient-identifiable information. London. December 1997.

2. BS EN ISO/IEC 17025: 2000. General requirements for the competence of testing and calibration laboratories. London: British Standards Institution; 2000. p. 1-26

3. EA-04/10. Accreditation for Microbiological Laboratories, revision 2. European co-operation for Accreditation; 2002. p. 1-26

4. EA-04/16. EA guidelines on the expression of uncertainty in quantitative testing. European co-operation for Accreditation; 2003. p. 1-27

5. M3003. The expression of uncertainty and confidence in measurement. United Kingdom Accreditation Service; 1997.

6. Eurachem / CITAC Guide. Quantifying Uncertainty in Analytical Measurement. 2nd ed. London: Laboratory of the Government Chemist; 2000. p. 1-126

7. LAB 12. The Expression of Uncertainty in Testing. Feltham, Middlesex: United Kingdom Accreditation Service; 2000. p. 1-13

8. Niemela SI, editor. Uncertainty of quantitative determinations derived by cultivation of micro-organisms. Helsinki: Centre for Metrology and Accreditation; 2003.

9. Guideline No. 47. Jewell K, editor. Microbiological measurement uncertainty: a practical guide. Campden & Chorleywood Food Research Association; 2004.

10. DD 260:2003. Water quality - enumeration of micro-organisms in water samples - Guidance on the estimation of variation of results with particular reference to the contribution of uncertainty of measurement. London: British Standards Institution; 2003.

11. Standing Committee of Analysts. The Microbiology of Drinking Water (2002) - Part 3 - Practices and procedures for laboratories. Environment Agency. http://www.dwi.gov.uk/regs/pdf/micro.htm.

12. Langton SD, Chevennement R, Nagelkerke N, Lombard B. Analysing collaborative trials for qualitative microbiological methods: accordance and concordance. Int J Food Microbiol 2002;79:175-81.