NATIONAL PATHOLOGY ACCREDITATION ADVISORY COUNCIL

REQUIREMENTS FOR THE ESTIMATION OF MEASUREMENT UNCERTAINTY

(2007 Edition)

ISBN: 1 74186 164 0
Online ISBN: 1 74186 165 9
Publications Approval Number: 3960

First published 2007

Publications Production
Australian Government Department of Health and Ageing

Contents

National Pathology Accreditation Advisory Council

The National Pathology Accreditation Advisory Council (NPAAC) was established in 1979 to consider and make recommendations to the Australian, state and territory governments, on matters related to the accreditation of pathology laboratories and the introduction and maintenance of uniform standards of practice in pathology laboratories throughout Australia. An ongoing function of NPAAC is to formulate standards, and initiate and promote guidelines and education programs about pathology tests.

Publications produced by NPAAC are issued as accreditation material to provide guidance to laboratories and accrediting agencies about minimum standards considered acceptable for safe laboratory practice.

Failure to meet these minimum standards may pose a risk to public health and patient safety.

Preface

Medical laboratories are responsible for ensuring that test results are fit for clinical application by defining the required analytical performance goals and selecting appropriate measurement procedures.

Measurement uncertainty (MU) provides quantitative estimates of the level of confidence that a laboratory has in the analytical precision of test results, and is therefore an essential component of a quality system for medical laboratories.

The authoritative reference for MU cited in International Organization for Standardization (ISO) ISO 15189 is the *Guide to the Expression of Uncertainty in Measurement* (GUM), published in 1995 by a collaboration of national and international standards bodies. The theory and implementation of MU described in GUM was developed specifically for calibration and testing laboratories undertaking measurements in fields such as analytical chemistry and physical testing (e.g. mechanical, electrical, temperature), and does not address the special nature of much of quantitative medical testing. This NPAAC standard and accompanying guidelines provide general guidance for the practical implementation of MU in medical laboratories, taking account of the limitations of biological measurement and the basic principles of MU. It should be recognised that the approach to applying MU principles to measurements in medical laboratories is still evolving, and that discipline-specific aspects may not be addressed fully in this edition.

The use of the word 'must' in each standard within this document indicates a mandatory requirement for pathology practice; 'should' is used to indicate guidelines or recommendations where compliance would be expected for good laboratory practice. Notes and commentaries provide guidance on the document, and examples are intended to illustrate the text and provide guidance on interpretation.

- A standard is the minimum standard for a procedure, method, staffing resource or laboratory facility that is required before a laboratory can attain accreditation; standards are printed in **bold** type and prefaced with an 'S' (e.g. **S2.2**).
- A guideline is a consensus recommendation for best practice and should be used if a higher standard of practice is appropriate, particularly when setting up or modifying a laboratory test, or when contamination problems have occurred; guidelines are prefaced with a 'G' (e.g. G2.2) and are numbered to correspond with their associated standard.
- A commentary is provided to give clarification to the standards and guidelines, and may include examples, references and guidance for interpretation.

This document is based on a document prepared by the Uncertainty of Measurement Working Group, which was established under the auspices of the Scientific and Regulatory Affairs Committee of the Australasian Association of Clinical Biochemists (White and Farrance 2004). NPAAC acknowledges the work undertaken by this working group.

This document is for use in the accreditation process. Comment from users can be directed to:

The Secretary
NPAAC
Australian Government Department of Health and Ageing
MDP 107
GPO Box 9848
CANBERRA ACT 2601

Phone: +61 2 6289 4017
Fax: +61 2 6289 8509
Email: npacc@health.gov.au

Abbreviations

CLSI	Clinical and Laboratory Standards Institute (formerly NCCLS), Wayne, Philadelphia, US.
CRM	certified reference material
CV	coefficient of variation
FISH	fluorescent in situ hybridisation
GUM	*Guide to the Expression of Uncertainty in Measurement* (GUM) (1993). BIPM, IEC, ICC, ISO, IUPAC, IUPAP, OIML, 1st edition (corrected and reprinted in 1995).
ISO	International Organization for Standardization, Geneva, Switzerland.
IVDs	in vitro diagnostic devices
MU	measurement uncertainty
P	Plasma
RCPA	Royal College of Pathologists of Australasia
SD	standard deviation
VIM	International vocabulary of basic and general terms in metrology (see reference ISO 1993).

Definitions

Accuracy of measurement	Closeness of the agreement between the result of a measurement and a true value of the measurand. [VIM: 1993, definition 3.5]
Analyte	Component represented in the name of a measurable quantity. [ISO 17511, ISO, Geneva, Switzerland]
Analytical interference	System effect on a measurement caused by an influence quantity which does not by itself produce a signal in the measuring system, but which causes an enhancement or depression of the value indicated. [ISO/WD 15193:2006; 3.10, ISO, Geneva, Switzerland]
Analytical specificity	Ability of a measurement procedure to determine solely the quantity it purports to measure. [ISO/WD 15193:2006; 3.9, ISO, Geneva, Switzerland]
Bias	Difference between the expectation of the test results and an accepted reference value. [ISO 3534-1, ISO, Geneva, Switzerland]
Certified reference material (CRM)	Reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes metrological traceability to an accurate realisation of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence. [Harmonised Terminology Database, CLSI, http://www.clsi.org]
Combined standard uncertainty (u_c)	Standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities. [GUM 1995]
Commutability of a reference material	Property of a given reference material demonstrated by the closeness of agreement between the relation among the measurement results, for a stated quantity in this material, obtained according to two given measurement procedures, and the relation obtained among the measurement results for other specified materials. [VIM] Note: The material in question is usually a calibrator. At least one of the two given measurement procedures is usually a high-level measurement procedure.
Coverage factor (k)	Numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty. [GUM 1995]

Expanded uncertainty (U)	Quantity defining an interval about a result of a measurement expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. Note 1: The fraction may be regarded as the coverage probability or level of confidence of the interval. Note 2: To associate a specific level of confidence with the interval defined by the expanded uncertainty requires explicit or implicit assumptions regarding the probability distribution characterised by the measurement result and its combined standard uncertainty. The level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions can be justified. [GUM 1995]
Imprecision	Dispersion of independent results of measurements obtained under specified conditions. [Harmonised Terminology Database, CLSI, http://www.clsi.org]
Influence quantity	Quantity that is not the measurand but that affects the result of a measurement. [VIM 1993]
Kind-of-quantity	See Table 1 for examples.
Matrix (of a material system)	All components of a material system, except the analyte. [ISO 15193, 15194, ISO, Geneva, Switzerland]
Matrix effect	Influence of a property of the sample, independent of the presence of the analyte, on the measurement and thereby on the value of the quantity being measured. Note 1: A specified cause of a matrix effect is an influence quantity. Note 2: A matrix effect depends on the detailed steps of the measurement as described in the lmeasurement procedure. For example, the measurement of the amount-of-substance concentration of sodium ion in plasma by flame emission spectrometry may be influenced by the viscosity of the sample.
Measurand	Quantity intended to be measured. [VIM 1993]
Measurement	A set of operations having the object of determining a value of a quantity. [VIM: 1993, definition 2.1]
Measurement method	Generic description of a logical sequence of operations used in a measurement (e.g. two-site sandwich immunoassay).
Measurement procedure	Set of operations, described specifically, used in the performance of particular measurements according to a given method [VIM 1993]. For example, specific procedures as marketed by specific manufacturers. A measurement procedure is usually documented in sufficient detail to enable an operator to perform a measurement.

Measurement uncertainty (u)	Parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand. [VIM 1993] Note: The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence (ISO 15195). [1]
Nominal property	Property that can be compared for equality or identity with another property of the same kind-of-property, but has no magnitude.
Ordinal quantity scale	Quantity scale defined by formal agreement. An ordinal quantity scale may be established by measurements according to a measurement procedure.
Precision	Closeness of agreement between independent test results obtained under stipulated conditions. Note 1: Precision depends only on the distribution of random errors and does not relate to the true value or the specified value. Note 2: The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a higher standard deviation. [Note 3 omitted] Note 4: Precision of measurement is a qualitative concept. [ISO 3534-1]
Primary sample specimen	Set of one or more parts initially taken from a system. [ISO 15189:2003(E); 3.14, ISO, Geneva, Switzerland]
Quantity	Attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively. [VIM 1993, definition 1.1]
Quantity scale	Ordered set of values of quantities of a given kind used in ranking quantities of the same kind (e.g. celsius temperature scale).
Reference measurement procedure	Thoroughly investigated measurement procedure, described in detail in a written document, shown to yield values having a measurement uncertainty commensurate with its intended use, especially in assessing the trueness of other measurement procedures for the same quantity and in characterising reference materials. [ISO/WD 15193:2006; 3.7]
Relative measurement uncertainty	Standard uncertainty (units) expressed as a coefficient of variation (CV; or %CV; dimensionless).

1 ISO (International Organization for Standardization) (2003), *Laboratory Medicine – Requirements for Reference Measurement Laboratories*. ISO 15195, Geneva

Repeatability	Precision under repeatability conditions. That is, conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time. [ISO 3534-1, ISO, Geneva, Switzerland]
Reproducibility	Precision under reproducibility conditions. That is, conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment. [ISO 3534-1, ISO, Geneva, Switzerland]
Sample	One or more parts taken from a system and intended to provide information on the system, often to serve as a basis for decision on the system or its production. Example: A volume of serum taken from a larger volume of serum. [ISO 15189:2003(E); 3.14]
Standard uncertainty ($u(x_i)$)	Uncertainty of the result of a measurement expressed as a standard deviation. [GUM 1995]
Traceability	Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties. [VIM: 1993, definition 6.10]
Trueness	Closeness of agreement between the average value obtained from a large set of test results and an accepted reference value. Note: The measure of trueness is normally expressed in terms of bias. The reference to trueness as 'accuracy of the mean' is not generally recommended. [ISO 3534-1]
Trueness of measurement	Closeness of agreement between the average value obtained from a large series of results of measurements and a true value. Note: Adapted from ISO 3534-1:1993, definition 3.12
Uncertainty budget	List of sources of uncertainty and their associated standard uncertainties, compiled with a view to evaluating a combined standard uncertainty associated with a measurement result. Note: The list often includes additional information, such as sensitivity coefficients (rate of change of result with change in quantity affecting the result), degrees of freedom for each standard uncertainty, and an identification of the means of evaluating each standard uncertainty in terms of a Type A or Type B evaluation. [ISO/TS 21748:2004(E); 3.13]

Principles and relevance of measurement uncertainty

All types of measurement have some inaccuracy due to bias and imprecision, and therefore measurement results can be only estimates of the values of the quantities being measured. To properly use such results, medical laboratories and their clinical users need some knowledge of the accuracy of such estimates. Traditionally, this has been by using the concept of error, but the difficulty with this approach is that the term 'error' implies that the difference between the true value and a test result can be determined and the result corrected, which is rarely the case. In contrast, the more recent concept of measurement uncertainty (MU) assumes that significant measurement bias is either eliminated, corrected or ignored, evaluates the random effects on a measurement result, and estimates an interval within which the value of the quantity being measured is believed to lie with a stated level of confidence.

Estimates of MU provide a quantitative indication of the level of confidence that a laboratory has in each measurement and are therefore a key element of an analytical quality system for medical laboratories. The principles of measurement uncertainty contribute to ensuring test results are fit for clinical application by:

- defining the quantity intended to be measured (measurand)
- indicating the level of confidence a laboratory has in a given measurement
- providing information essential for the meaningful interpretation of measurement results and their comparison over space and time
- identifying clinically appropriate goals for imprecision
- identifying significant sources of MU and opportunities for their reduction.

Measurement uncertainty and traceability

The long-term goal for any field of measurement is to be able to meaningfully compare quantitative test results for a given quantity (analyte) produced by any laboratory at any time. To achieve this, all routine measurement procedures for a given measurand must have a quantifiable relationship (see Figure 1) to an internationally recognised (certified) reference material. This relationship is established through a hierarchy of method procedures and calibrators, typically from a high-order international reference material via secondary reference materials and procedures to local method calibrators and procedures. At each stage, the value and the estimated MU of the method calibrator and the local routine procedure must be known. Thus, MU is an essential component of traceability.

For quantitative medical testing, such method traceability facilitates patient mobility between laboratories, use of common clinical decision values and local application of clinical research data. At present, few measurement procedures in medical testing are traceable, but with increasing clinical application of international expert group decision limits and the desirability of common reference values, there is an increasing need for method traceability.

Many methods presently lack certified reference materials and therefore do not have traceability; however, it is often practical to use conventional reference materials, conventional reference methods and external proficiency testing to facilitate the comparability of measurements between users of the same measurement procedure within and between laboratories. It should be noted that traceable calibrators may facilitate, but do not guarantee, that measurement results are transferable or comparable, unless also shown to be commutable across all methods and procedures for the given analyte.

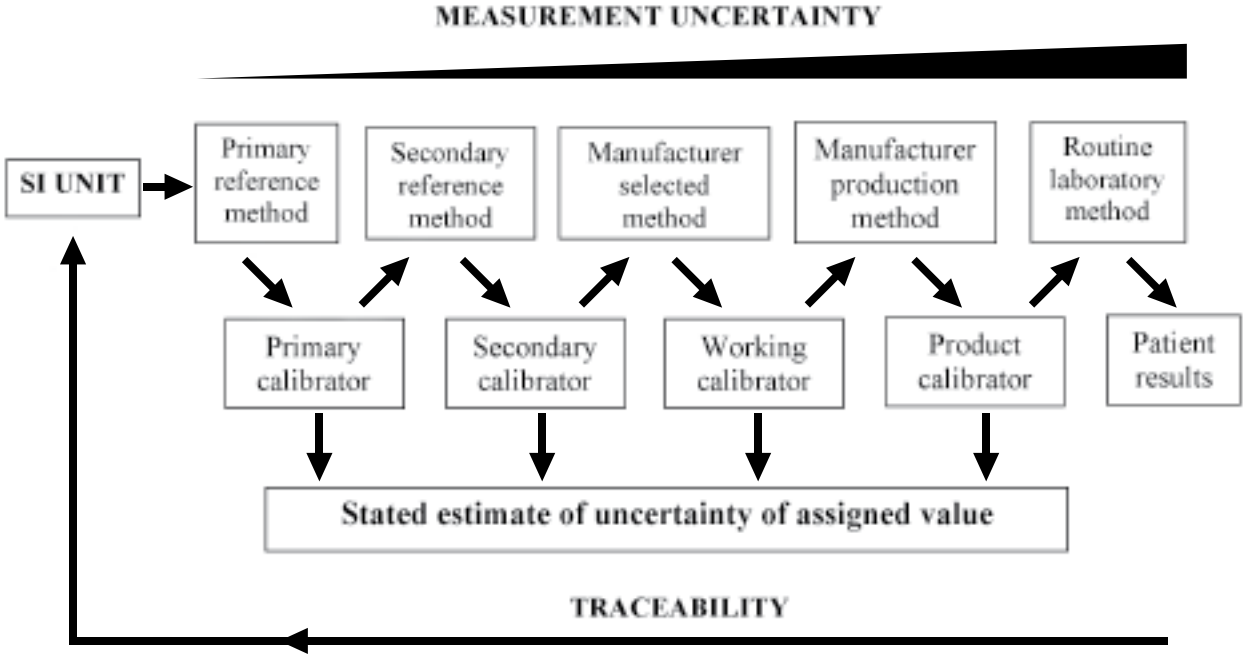

Figure 1 Relationship of traceability and measurement uncertainty

Measurement uncertainty and bias

MU (random effects; imprecision) and bias (systematic effects; inaccuracy) are two critical determinants of the quality of measurements, and although they are separate concepts, it is good laboratory practice to document both parameters together for methods where bias can be assessed (see Figure 2 below). This approach is followed in these guidelines.

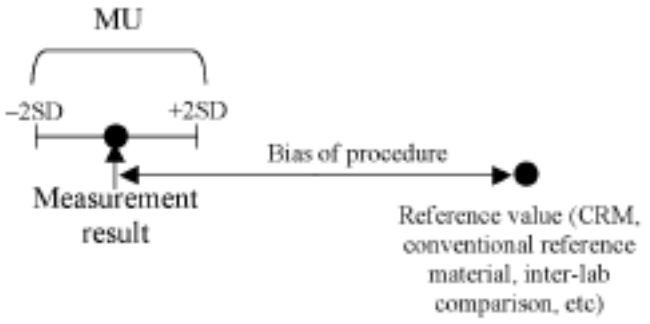

Figure 2 Measurement uncertainty and bias

Guide to uncertainty in measurement and medical testing

The *Guide to the Expression of Uncertainty in Measurement* (GUM) was developed primarily for estimating MU in fields such as physical and chemical testing (e.g. electrical, materials, optics, etc). These types of measurements generally lend themselves to a bottom-up approach to estimating MU, because the potential sources of uncertainty are usually readily identifiable, and their magnitudes can be estimated and combined.[2] However, medical laboratory methods can generally use quality control materials to monitor whole-of-procedure performance, and therefore the quality data generated can be used to directly estimate their combined measurement uncertainties (top-down approach). Therefore, where technically possible, this document recommends the use of quality control data to estimate measurement uncertainties. The need for further action depends on whether the MU estimates for a given method suggest that the results produced will be fit for clinical application. It is therefore essential to set appropriate MU goals for each method procedure. If an MU goal is not met, the method procedure may need to be analysed to identify significant and modifiable uncertainty sources based on the bottom-up approach. The effort and cost of such analysis should be commensurate with the technical and clinical requirements.

Sources of uncertainty and the interpretation of patient results

Medical laboratories have a good understanding of the many non-disease influences that can affect a patient result and its clinical interpretation (Figure 3). Whether such factors will have a significant effect will depend largely on the value of the result and its clinical application. For practical convenience, these factors are usually grouped according to where they may act in the request–test–report cycle. In the following summary, it is assumed that all technical steps are conducted according to standard operating procedures and without nonconformances.

Pre-analytical sources

Differences in patient preparation, specimen collection technique, transportation and storage time, and preparation of primary sample, etc, may alter the measurable amount of an analyte in a sample. Laboratories should have standard operating procedures in place to eliminate or minimise these influences to acceptable levels for given measurement procedures. Other factors that may influence a measurement are generally patient specific (e.g. heterophilic antibodies, jaundice, drugs and other factors, as shown in Figure 3).

2 The Gum bottom-up approach uses a variety of data sources, such as experiments, manufacturer information, validation data and professional judgment, to assemble a model of the component uncertainties for a given procedure, from which a combined (i.e. total) uncertainty is calculated. GUM recommends that, where possible, uncertainty models and estimates should be compared with actual whole procedure data.

Measurement uncertainty

In this document MU is considered to encompass the inputs and influences on a measurement result that occur within the technical bounds of the measurement procedure itself.

The measurement process typically commences when an acceptable specimen interacts with the first technical step of the measurement procedure (e.g. placement in an automated analyser, commencement of an extraction step before measurement). Typical MU sources include uncertainty of the calibrator value and dispensed volumes, reagent and calibrator batch variations, equipment ageing and maintenance, changing operators, environmental fluctuations, etc. There may also be uncertainty associated with the component(s) in the measurand itself (e.g. different molecules can carry the same epitope detected by a given antibody).

Post-analytical sources

Patient results should comprise an appropriate number of significant figures, as reporting an inappropriate number may adversely affect clinical interpretation (see Appendix 2). However, for some purposes (e.g. quality control data, comparing results, clinical trials), limiting the number of significant figures reported may adversely affect their statistical use.

Sources of uncertainty and result interpretation

Disease and physiological factors such as biological variation, stress and diet may have significant effects on the amount of an analyte present in the specimen at the moment of collection. Depending on the definition of the measurand, its clinical application and the amount reported, some of the modifying influences can bring uncertainty to result interpretation. If a test value is distant from a clinical decision value, the non-disease factors are generally of little or no importance, but as results approach clinical decision values, or a previous result, their optimal interpretation may need to account for the effects of the relevant influences.

In summary, MU is one of the major potential contributors to the uncertainty of result interpretation, and laboratories should have such data available for clinical users.

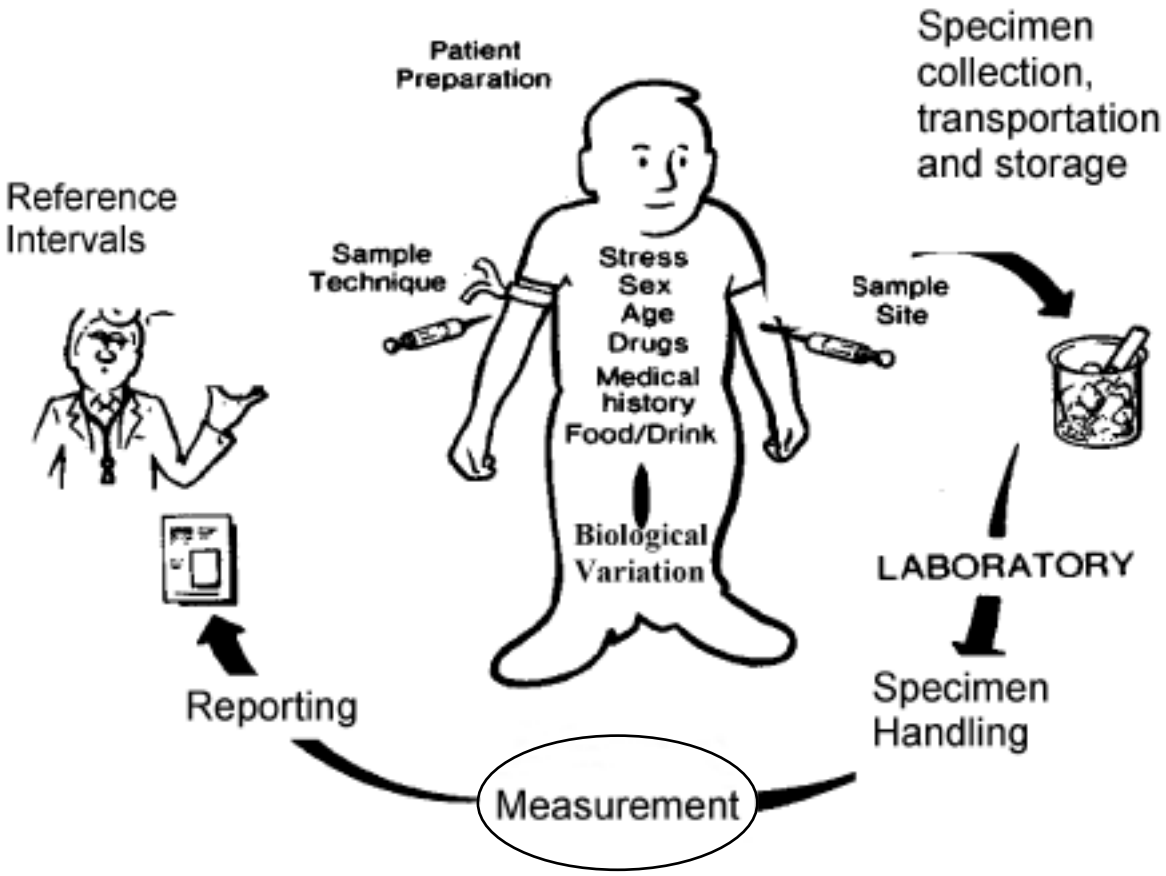

Figure 3 **Main sources of uncertainty in the request–test–report cycle (modified from Walmsley and White (1985) *Pocket Diagnostic Clinical Chemistry,* Blackwell Scientific Publications)**

Standard

S1 Laboratories must estimate measurement uncertainty where relevant and possible.

Commentary

C1.1 Estimates of measurement uncertainty should be made for all measurements. The complexity and cost of obtaining an estimated MU should be commensurate with the quality requirements of the clinical application of the results. If a laboratory decides that MU is not relevant or possible to estimate, then the laboratory should document the reasoning.

C1.2 Estimates of MU allow measurements to be compared meaningfully with each other and with clinical decision values. Within laboratories, such estimates are a critical parameter for quantifying and monitoring the quality of measurements, and for understanding their technical limitations.

C1.3 This standard does not apply to qualitative tests which are not derived from a numerical value.

C1.4 Some qualitative methods generate numerical values during the procedure that are reported finally in relation to preset values (cut-offs) (i.e. ordinal quantities). For such methods, the laboratory should estimate an MU for that part of the procedure that generates the numerical values.

Guidelines

Measurands

G1 Laboratories should define the measurand of each of their measurement procedures and record clinically important limitations and interferences.

Commentary

C1.1 A measurand is defined as the particular quantity subject to measurement, where the quantity is the attribute of a substance that can be distinguished and determined quantitatively. It is essential to define as fully as possible the quantity that is measured (i.e. the measurand) by a given procedure. There are four aspects of a measurand that should be described (see Table 1):

(a) quantity intended to be measured

(b) system

(c) kind-of-quantity and measurement unit

(d) method.

Some measurands require further definition, which may include parameters such as time, temperature, specimen site and specific measurement procedure (e.g. total serum calcium by atomic absorption spectrophotometry, serum total calcium by *O*-cresolphthalein complexone).

C.1.2 For some types of methods, there can be significant limitations to defining a measurand adequately. For example, a monoclonal antibody to a specified epitope may result in measurement of the concentration of a variety of molecules, the relative proportions of which may vary from individual to individual (e.g. hCG species, prolactin/macroprolactin). Where lack of measurand definition may have relevance to the clinical interpretation of patient results, such limitations should be recorded. Similarly, it is also important to identify limitations or interferences that can cause clinically important effects on measurement of the specified measurand (e.g. heterophilic antibody interference with an immunoassay, detection of non-specific or cross-reacting antibodies in serological methods).

Table 1 Examples of measurand definition

Quantity intended to be measured	System	Kind-of-quantity	Measurement unit	Method
Sodium	Venous plasma	Amount of substance concentration	mmol/L	Flame photometry
Calcium ion	Arterial whole blood	Amount of substance concentration	mmol/L	Ion-selective electrode
Creatine kinase MB	Serum	Mass concentration	µg/L	two-site immunoassay
Creatine kinase MB	Serum	Activity concentration	mIU/L at 37°C	Immuno-inhibition
FMR1 gene	Genomic DNA	Number of CCG repeats in the FMR1 gene		Capillary electrophoresis
Chromosome 21	Cell	Number of FISH signals for chromosome 21 probe per cell		FISH
Haemoglobin	Venous whole blood	Mass concentration	g/L	Spectrophotometry
White cell count	Urine	Number concentration of white cells in urine	White cells per volume	Microscopy
Prolactin/ macroprolactin	Serum	Mass concentration	µg/L	two-site immunoassay
Rubella IgG	Serum	Rubella IgG + cross-reacting IgGs	Arbitrary units, IU/L	Immunoassay
Gentamicin	Serum	Trough mg/L–trough (8 hours post-dose)	mg/L	Immunoassay

Imprecision

G2 Imprecision should be included in estimates of the uncertainty of measurement procedures.

Commentary

C2.1 MU provides a quantitative estimate of the variability in results a laboratory would normally expect if the measurement were to be repeated at another time. For most measurement procedures, random effects are the major contributors to MU, and therefore quantifying imprecision provides the most reasonable estimate of the combined standard measurement uncertainty (u_c). Estimates should, where possible, include levels of the measurand at or near clinical decision values.

C2.2 Estimated combined standard measurement uncertainties (u_c) are expressed as either 1SD (units) or as relative u_c (CV, CV%), and should include an indication of the range of measurement values to which they are applicable.

C2.3 To define intervals that enclose larger fractions of expected dispersions of results, coverage factors (k) may be applied to u_c to provide expanded measurement uncertainties (U).

Step 1 The recommended first step is to make a reasonable estimate of the imprecision for the whole measurement procedure (u_C). For procedures already in routine laboratory service, the most efficient approach to estimating the expected dispersion of results is to calculate the standard deviation (SD) of results achieved for the appropriate quality control material(s). The laboratory should be satisfied that the material used behaves in the measurement procedure in a similar way to that of patient samples. A statistically valid number of results should be collected across all routinely encountered events that are reasonably expected to have a detectable influence on the results produced (e.g. calibrator and reagent batch changes, different operators, equipment maintenance, environmental fluctuations). The laboratory should ensure that the estimate of u_c is applicable across the reporting range; an estimate at more than one measurand value may be necessary.

For new methods undergoing evaluation or verification, an interim estimate of imprecision should be made from a statistically valid number of results produced by several different analytical runs.

Where use of routine quality control materials is not possible, an estimate may be achievable using the laboratory's results from an external assessment program. However, it should be noted that such estimates may not comprise sufficient data, not adequately cover all routine measuring conditions, or not be applicable to all clinical decision values for a given procedure.

Step 2 Whole-of-procedure imprecision can be used as the reasonable estimate of u_c. Estimates of combined uncertainty (u_c) should be expressed as a standard uncertainty (i.e. SD in the reported units of measurement) or as a percentage relative standard uncertainty (i.e. CV%) for a stated value of the given measurand).

Step 3 The clinical use of MU is in either comparing two results from the same patient or comparing a result with a clinical decision value that, by definition, is without uncertainty (see examples).

Bias

G3.1 The MU concept assumes that significant measurement bias is eliminated, corrected for or ignored.

G3.2 If a bias value or a correction factor is applied, then an estimate of the uncertainty of the value used should be assessed for inclusion in the estimate of combined uncertainty for the procedure.

G3.3 Although bias and MU are separate components of the quality of a measurement result, it is good laboratory practice, where relevant and possible, to record bias data together with MU data. Laboratories would be expected to possess the necessary bias data from their method evaluation studies (e.g. in-house IVD manufacturer).

Commentary

C3.1 Bias and MU are separate components of a measurement result, where MU is the variability expected if a measurement were to be repeated, and 'bias is the difference between the expectation of the test results and an accepted reference value' (ISO 3534-1). Ideally, such an accepted reference value would be provided by a commutable certified reference material, but this option is presently available for only a minority of methods. For those lacking certified reference materials (CRMs), it is often clinically and technically useful to align results produced by different laboratories using the same measurement procedures by estimating bias relative to conventional reference materials, reference methods, interlaboratory comparisons, etc.

For those procedures producing results that are interpreted by comparison with clinical decision values or previous test results produced by the same procedure conducted by the same laboratory, any significant bias should be comparable for all similar values and therefore should cancel out.

Methods with traceable calibrators

The bias of a measurement procedure calibrated with a traceable calibrator can be estimated in various ways (e.g. measuring a commutable and traceable certified reference material, spiking studies, reference method procedure, etc). Whatever approach is used, the final step is for the mean value generated by the routine method to be compared with the reference value to assess if they are significantly different (t test). If the bias is small relative to measured values then it can be ignored; otherwise, it should be investigated and if possible, eliminated or corrected by recalibration of the measurement procedure. The uncertainty of the bias value or correction factor used should be assessed for inclusion in the estimate of combined uncertainty.

The best approach to assuring the traceability and uncertainty of results of commercial methods is to obtain traceability certificates or statements from the manufacturer for the values and uncertainties assigned to their calibrators and use these data to adjust for bias if required. However, *bias corrections* should only be made when the provider of the CRM can provide data that demonstrates the commutability of the reference materials for the measurement procedures being used. It should be noted that if the matrix of a CRM and that of typical routine samples is very different, the estimated uncertainties may not be relevant to routine practice.

Methods without traceable calibrators

Many measurement procedures lack reference materials traceable to a higher metrological order (e.g. certified reference material or a recognised international standard), and they may also suffer inadequate measurand definition.

Where results generated by non-traceable methods are interpreted relative to clinical decision values determined by a different measurement procedure (e.g. defined by an expert group), an estimate of bias may be needed to achieve clinically acceptable interlaboratory agreement. In such cases, it may be useful to use an appropriate material that has been assigned a conventional reference value or an appropriate group mean from an external proficiency-testing program, or conduct an interlaboratory comparison. Where none of these approaches is practical, bias is unknown, and therefore ignored.

Measurement uncertainty goals

G4 Laboratories should set routine performance goals for measurement uncertainty based on the clinical use of the test results.

Commentary

C4.1 Currently, few methods have internationally agreed performance goals (e.g. cholesterol, haemoglobin A1c). In the absence of such goals, various approaches have been used to set clinically relevant targets. A widely used approach to setting an MU goal is to define the upper acceptable limit as a proportion of the intra-individual biological variation of the measurand. The principle of this approach is that with the correct choice of the proportionality factor, imprecision should not contribute significant additional variation to the test result when compared with the natural variation of the component.

This approach should be used with caution because there is limited evidence concerning biological variation of the measurand in healthy individuals and its applicability to the sick patient. It is also advisable to consult recent literature to ensure the most relevant data are used.

C4.2 For some measurement procedures, depending on physiological considerations and clinical applications, more than one imprecision goal may be appropriate (e.g. use of serum hCGs for pregnancy testing, monitoring threatened miscarriage or for the management of testicular tumours).

C4.3 For many methods and measurands, imprecision goals based on biological variation are inappropriate because they may not be:

(a) achievable by current routine laboratory technology (e.g. plasma sodium concentration)

(b) relevant to the clinical application (e.g. urine sodium concentration)

(c) relevant physiologically (e.g. serum hCG concentration in normal early pregnancy).

In such cases, goals can be set using other criteria (e.g. expert group recommendation, clinical or laboratory opinion). Some external proficiency testing programs use clinically based criteria for assessment, and performance comparisons are a useful guide to current 'state-of-the-art' MU for routine measurement procedures.

Measurement uncertainty outcomes

G5.1 Where the goal for MU is met, the major sources of uncertainty need not be individually identified or their magnitude estimated.

G5.2 If a measurement procedure does not meet its uncertainty goal, the laboratory should identify and attempt to reduce significant sources of uncertainty, or consider changing the method, to ensure the goal is met.

Commentary

C5.1 Sources that contribute to uncertainty may include sampling[3], sample preparation, sample portion selection, calibrators, reference materials, input quantities, equipment used, environmental conditions, condition of the sample and changes of operator (ISO 15189, 5.6.2).

C5.2 If the analytical goal is not met, then likely major contributors to the combined uncertainty should be identified and their magnitudes estimated. The required uncertainty data may be estimated from sources such as direct experimentation (e.g. pipette imprecision), manufacturer data and the literature. Another helpful source of data is the identification of trends and shifts in the QC data that can be related to specific events, such as reagent stability, lot-to-lot variation or preparation differences, stability of calibration and maintenance programs, etc. Reviewing these data can lead the user, and sometimes the manufacturer, to improved practices that can reduce the uncertainty of a procedure.

C5.3 Opportunities for reducing MU should be sought (e.g. replace manual pipettes with automated system). Fully automated instrumentation generally limits such opportunities, and therefore failure to meet an MU goal may result in a range of outcomes (i.e. changing individual technical steps within a measurement procedure to replacement of the method). If one or more technical steps are modified to reduce their uncertainty, then the combined uncertainty of the modified measurement procedure must be estimated and assessed for fit-for-purpose.

3 Sample: example - a volume of serum taken from a larger volume of serum. ISO 15189:2003(E); 3.10

Numerical significance

G6 Laboratories should report test results to the number of significant figures consistent with the MU of the method.

Commentary

C6.1 Patient results should be reported to the appropriate number of significant figures, as use of an inappropriate number may adversely affect clinical interpretation (see Appendix 2).

Clinicians may not be aware of the true imprecision of the results they use, and can be misled by the inappropriate use of significant figures in patient reports. In addition, they will only appreciate the implied significance of reporting significant figures if all laboratories use the same approach.

C6.2 Laboratories should report results in rounding intervals that are commensurate with the MU of their measurement procedures.

C6.3 For a given measurement procedure, the uncertainty – and hence the rounding interval – may vary significantly across the reportable range. Care is therefore required to ensure the chosen rounding interval is appropriate across a reporting range.

C6.4 Significant digits and rounding: For a given uc, the number of significant figures should generally be one (e.g. $u_c = 0.039$ becomes 0.04; $u_c = 7.5$ becomes 8).

C6.5 A measurement value should be rounded to the same decimal place as its measurement uncertainty (e.g. measurement value of 151.4, $u_c = 4$ should be reported as 151) (ISO GUIDE 31 1992).

C6.6 Rounding may affect the statistical use of results (e.g. quality control data, comparison of results, clinical trials) and should be deferred until the final result is calculated.

Clinical applications

G7 Laboratories should ensure relevant MU information is available.

Commentary

C7.1 Depending on their relative values, measurements of a given measurand often cannot be meaningfully compared with each other or with a clinical decision value without knowledge of their uncertainty.

In medical testing, measurement results are generally interpreted by comparison with other values. Such comparisons are for the purpose of either assessing whether the two values are measurably different by the procedure used, or whether they are not only measurably different but also biologically different. For both types of assessment, knowledge of the measurement uncertainty of the patient result is necessary.

The value with which a patient result is compared is usually either a previous result for the same patient, or is a clinical decision value. In the first situation, the expected dispersion of both results must be taken into account in assessing whether the value difference between them is probably due to just measurement uncertainty, or because they are measurably different (see Appendix 2a). In the second situation, a clinical decision value is generally a fixed value with no dispersion, and therefore the only measurement uncertainty to consider is that of the patient result (see Appendix 2b).

From Appendix 2a, it can be seen that if two patient results are separated by greater than $2½ \times 1.96 \times u_c$, $(2.77 \times u_c)$, then there is about 95% probability that they are measurably different by the measurement procedure used.

If the result of a measurement is compared with a reference value (fixed value with no MU) then the calculation is $2 \times u_c$ (Appendix 2b).

C7.2 It is important for laboratories to understand the clinical implications of the results of the measurements they report and to be aware of those where MU could affect clinical interpretations and patient management.

C7.3 Laboratories should consider providing relevant MU information with patient reports where it may be of clinical utility (e.g. tumour marker monitoring).

Appendices

The appendices contain a number of examples of determinations of measurement uncertainty across various medical laboratory disciplines. The examples are provided by different laboratories, and it should be noted that a variety of formats may be used.

A1 Combining uncertainty estimates

When the measurement value is derived from more than one input, the uncertainty of the result is calculated by combining the uncertainties of the significant contributing inputs. There are mathematical rules that must be followed when adding individual uncertainty estimates. Two formulae are relevant, and the choice depends on how the final result is calculated from the contributing inputs.

1. For the estimate of combined measurement uncertainty calculated from a sum and/or a difference of independent inputs (i.e. inputs without covariance)

   If a result (R) is derived from two (or more) independent inputs (X and Y) by their addition and/or subtraction, then the imprecision of the contributing inputs must be summed as their variances (SD^2):

   Let: $R = X + Y$ or $R = X - Y$,

   Then: $u_R = ((u_X)^2 + (u_Y)^2)^{1/2}$

   where: u_R, u_X and u_Y are the respective input standard uncertainties (e.g. technical steps within a measurement procedure, other measurements), expressed as standard uncertainties (e.g. imprecision).

   Example: Measurement uncertainty of anion gap (u_{cAG})

   Anion gap (AG) is derived by combining the measurements of serum (plasma) sodium, potassium, chloride and bicarbonate.

   $AG = ([Na^+] + [K^+]) - ([Cl^-] + [HCO_3^-])$

   The uncertainty of a result is related to the sum of the individual standard uncertainties (ux1, ux2, etc), which occur at each stage of the measuring process. For results derived from a sum and/or a difference, the combined uncertainty can be expressed mathematically by adding together the variances of the contributing measurements (CV cannot be used for summing):

   $(u_{AG})^2 = (u_{Na}+)^2 + (u_K+)^2 + (u_{Cl}^-)^2 + (u_{HCO_3}^-)^2$

   Let:

   $u_{Na+} = 1.2$ mmol/L; $u_K+ = 0.1$ mmol/L; $u_{Cl}- = 1.3$ mmol/L; $u_{HCO_3}- = 1.2$ mmol/L

   Then: $(u_{AG})^2 = (1.2)^2 + (0.1)^2 + (1.3)^2 + (1.2)^2$

   $(u_{AG})^2 = 4.58$; $\mathbf{u_{AG}} = \mathbf{2.14} = \mathbf{\sim 2}$ **mmol/L** (see C6.4 – Rounding).

2. For the estimate of measurement uncertainty of a measurement calculated from a product and/or a quotient of independent inputs (i.e. without covariance)

If a result (R) is derived from two (or more) independent measurands (X and Y) by their multiplication and/or division, then the imprecision of the contributing measurements must be summed using their coefficients of variation $(CV)^2$:

Let: $R = X \times Y$ or $R = X/Y$ then,

$$(u_R/R)^2 = (SD_X/X)^2 + (SD_Y/Y)^2 = (CV_R)^2 = (CV_X)^2 + (CV_Y)^2$$

where: CV_R, CV_X and CV_Y are the respective relative uncertainties (e.g. CV%).

Example: Calculation of fasting spot urine calcium-to-creatinine ratio

Let: $u_{ca} = 7.71$ mmol/L; $u_{creat} = 3.1$ mmol/L

$CV_{Uca} = 5.2\%$; $CV_{Ucreat} = 3.9\%$

$uUca/creat = ((5.2)^2 + (3.1)^2)^{1/2} = 6.5\%$

MU = 7% (see C1.6.4 – Rounding).

A2 Application of measurement uncertainty to result interpretation

A patient has a serum prostatic specific antigen (PSA) result of 4.2 µg/L; 12 months ago, the result by the same laboratory and measurement procedure was 3.8 µg/L.

1. Has the PSA increased?

Considering only the measurement uncertainty

The combined measurement uncertainty (u_cPSA) for the laboratory PSA method at a concentration of 2.9 µg/L is 0.15 µg/L (rel u_c = 5.0%). Assume the PSA assay has had no significant change in bias.

The PSA has increased by 4.2 – 3.8 = 0.4 µg/L = 0.4/3.8 × 100 = 10.5%

Is the PSA increase greater than the laboratory would expect from the measurement imprecision?

It can be shown that two serial results are measurably different at a confidence level of 95% if they differ by > 2½ × 1.96 x u_A (2.77 × CV_A), where u_A = combined uncertainty of a method (SD), and CV_A = relative SD (rel u_c) of a method.

For the above example: 2.77 × 5.0 = 14%, i.e. the two results should differ by > 14% (> 0.53 µg/L) of the first result (3.8 + 0.53 = > 4.33 µg/L) for there to be 95% confidence that they are measurably different.

For a 99% confidence level:

2½ × 2.58 × u_A = 3.65 × u_A = 18.3% (i.e. > 0.7 µg/L); second result should be > 4.5 µg/L.

The laboratory could use its MU data in several ways to assist the referring practitioner with an interpretative comment (e.g. *'Taking account of measurement variability, this result is not significantly different at a confidence level of 95%'*).

Considering both measurement uncertainty and biological variation

In practice, the effect of individual biological variation (CV_I) should be included in the statistical comparison. As the two results are similar to the reference interval, it would be reasonable to assume that published data can be applied (CV_I = 14.0%). The measurement and biological dispersions are summed in the usual way.

For a 95% confidence level:

2½ × 1.96 × $[(CV_A)^2 + (CV_I)^2]^{½}$ = 2.77 × $[(5.0)^2 + (14.0)^2]^{½}$

= 2.77 × 14.87 = 41.2% (i.e. the first result must increase by at least 41.2%, or by 1.57 µg/L to ~5.4 µg/L, for there to be 95% confidence that the second result is both measurably and biologically different from the first result).

The relative effects of MU and biological variation on the PSA results can be seen.

The laboratory could assist interpretation with a comment such as *'Taking account of both measurement and biological variation, this result would need to be > 5.3 µg/L for 95% confidence that it has significantly increased from the previous result'.*

2. Is the latest PSA value significantly above the clinical decision value of 4.0 µg/L?

The clinical decision value of 4.0 µg/L does not have a known uncertainty associated with it. The rel u_c = 5.0% = 0.21 µg/L at a level of 4.2 µg/L. The 95% confidence interval for the patient result is ± 1.96 × 0.21 = 0.41 = 0.4 µg/L. Thus the 95% confidence interval for the latest result = 3.8–4.4 µg/L.

A3 International normalized ratio

International normalized ratio (INR) is derived by adjusting the results of a prothrombin time test with a factor, the International Sensitivity Index (ISI). The INR result corrects for the bias that a specified test thromboplastin reagent has when compared with a WHO standard thromboplastin.

In the following example, the calculation of combined uncertainty includes traceability data from the manufacturer regarding the precision of the ISI value of the current reagent batch. However, there is no correction for bias in this calculation as a reagent-specific reference interval has been determined by the laboratory. Thus bias has already been negated. The calculation of MU is done as follows.

1. **Identify measurand** – P, prothrombin; relative time.

2. **Calibrator** – The ISI factor is set by the manufacturer. The uncertainty shall be available.

3. **Set an analytical goal** – This is based on a proportion of biological variation (CV_I).

For many tests, the laboratory may decide to use the values for CV_I in the database on the Westgard website. Alternatively, the laboratory may use another standard published analytical goal or determine a goal from a specific multilaboratory study.

Thus, for prothrombin time, the goal might be: 0.5 CV_I = 2.0% (from database http://www.westgard.com/biodatabase1.htm).

4. **Identify all measurement uncertainties**

a. **Imprecision** can be calculated from the laboratory's own CV value of internal QC. This is known as CV_A and is derived from the internal QC data of a control plasma close to the clinical decision point. Data should be from a large number of consecutive determinations (e.g. control for abnormal P prothrombin time).

Prothrombin time = 39.5 seconds ± 1.1 seconds (mean ± SD for n = 200 samples)

As CV% = 100 x SD/mean %

CV_A = 1.1/39.5 = 2.8%

Desirable analytical goal has not been met as $CV_A > 0.5\ CV_I$ from database (i.e. 2.8% > 2.0%). However, the minimum goal of 0.75 CV_I (3%) has been met.

b. **Uncertainty of ISI value.** It is not always possible to obtain these data from the manufacturer. If the manufacturer does provide full traceability of calibration details, then the uncertainty of the ISI can be included in the calculation of uncertainty (e.g. manufacturer states that the thromboplastin reagent ISI = 1.26 ± 0.03 (CV = 2.3%)).

c. **Bias.** As the laboratory has determined a reagent-specific reference interval, the effect of bias on measurement uncertainty is already corrected.

5. Combined relative uncertainty

$u_C = [(CV_1)^2 + (CV_2)^2]^{0.5}$

$u_C = [(2.8)^2 + (2.3)^2]^{0.5} = 3.6\%$

6. Expanded relative uncertainty

$U = u_C \times k$

Where k = 2 for a 95.5% coverage factor

U = 7%

For a confidence level of approximately 95%, use a coverage factor of $k = 2$.

Notes:

1. In this example, the measurand is P-Prothrombin time (e.g. 39.5 seconds).
2. The prothrombin ratio (PR) is calculated as

   PR = time for test sample/time for normal 39.5/12.0 = 3.3.
3. INR is derived from the PR and ISI

   $INR = PR^{ISI}$ $3.2^{1.26} = 4.3$
4. ISI is specific for each batch of thromboplastin reagent and may be quoted on the reagent data sheet or obtained by request from manufacturer (e.g. ISI = 1.26 ± 0.03).
5. If ISI traceability data are not available, only imprecision of the laboratory can be calculated then report $CV_A\%$ = u = 2.8% at a mean prothrombin time of 39.5 seconds.
6. Values for many haematology analytes published in the Westgard database for biological variation are not appropriate for current use. Targets can be based on approaches based on Westgard (e.g. intralaboratory comparisons).
7. CV% of INR is not the same as CV% of prothrombin time.

A4 Haemoglobin

In the following example, the calculation of combined uncertainty does not include traceability data from the manufacturer, because these were not available. However these data could be added when they become available. There is correction for bias in this calculation as laboratory-specific bias has been determined from end-of-cycle RCPA quality assurance program (QAP) reports.

The calculation of MU is done as follows:

1. Identify the measurand

Venous blood haemoglobin concentration, vB – haemoglobin mass concentration.

2. Set an analytical goal that the laboratory should achieve

This is usually a relative uncertainty (e.g. CV%). The laboratory may decide to use the values in the database on the Westgard website. Alternatively, the laboratory may use another accepted analytical goal or determine a goal from a specific multilaboratory study. Thus, for haemoglobin the goal might be: 0.5 CV_I = 1.4% (from database http://www.westgard.com/biodatabase1.htm).

3. Identify all measurement uncertainties

a. **Imprecision** can be calculated from the laboratory's own internal QC. This is known as CV_A% and is derived from the internal QC data of a control sample close to the clinical decision point. Data should be from a large number of consecutive determinations.

For example, CV_A% = 1.1% for 'control X' (n = 200 samples).

Desired analytical goal has been met as 1.1% is less than the analytical goal of 1.4%.

This could be reported as u_c = 1.1% at a mean haemoglobin of xxx g/L.

b. **Uncertainty** of haemoglobin calibrator. If the manufacturer provides traceability to a reference standard, these data can be used in the determination of combined uncertainty (as in the previous example for ISI). However this is not included in the example below.

c. **Bias.** Data are available on the bias of an individual laboratory in End-of-Cycle reports of RCPA QAP proficiency testing.

For example, uncertainty CV for haemoglobin (mean of the last five cycles) = 1.2%

4. Relative combined uncertainty

$u_C = [(CV_1)^2 + (CV_2)^2 \ldots.]^{0.5}$

$u_C = [(1.1)^2 + (1.2)^2]^{0.5} = 1.6\%$

5. Relative expanded uncertainty

$U = u_C \times k$

Where $k = 2$ (95.5% coverage factor)

$u_c = 3.2\%$ ($k = 2$)

For a confidence level of approximately 95%, use a coverage factor of $k = 2$.

A5 Leucocytes

In this example, the calculation of combined uncertainty will not have traceability because there is no leucocyte primary reference. It should be noted that the matrix of control blood is not the same as patient samples and that imprecision of control material is often greater, especially for differential leucocyte counts. There is correction for bias in this calculation as laboratory-specific bias data is obtainable from End-of-Cycle QAP reports.

The calculation of MU is done as follows:

1. Measurand

B leucocytes; number concentration.

2. Set an analytical goal that the laboratory should achieve

This is usually a relative uncertainty (e.g. CV%) and is known as CV_I%.

The laboratory may decide to use the values in the database on the Westgard website. Alternatively the laboratory may use another standard published analytical goal or determine a goal from a specific multi-laboratory study.

Thus, for leucocytes the goal might be: 0.5 CV_I% = 5.5% (from database http://www.westgard.com/biodatabase1.htm).

3. Identify all measurement uncertainties

a. Imprecision can be calculated from the laboratory's own CV value from internal QC. This is known as CV_A% and is derived from the internal QC data of a control sample close to the clinical decision point. Data should be from a large number of consecutive determinations.

For example, CV_A = 2.6% for 'Control X' (n = 200 samples).

Desired analytical goal has been met as 2.2% is less than the analytical goal of 5.5%.

This could be reported as uc = 2.6% at a mean leucocyte count of y.y $10^9/_L$.

b. Bias. Data are available on the bias of an individual laboratory in End-of-Cycle reports of RCPA QAP proficiency testing.

For example, average BIAS for leucocyte count (mean of the last five cycles) = 3.2%.

4. Relative combined uncertainty

$u_C = [(CV_1)^2 + (CV_2)^2 \ldots.]^{0.5}$

$u_C = [(2.6)^2 + (3.2)^2]^{0.5} = 4.12\%$

5. Relative expanded uncertainty

$U = u_C \times k$

Where $k = 2$ (95.5% coverage factor)

$U = 8\%$ ($k = 2$)

For a confidence level of approximately 95%, use a coverage factor of $k = 2$

An appropriate way to report, if required, would be leucocytes = $6.6 \pm 0.5 \times 10^9$/L. However, this should only be provided on request.

A6 Measurement procedure for fragile X (A) syndrome (CGG repeats)

Measurement of uncertainty report

Measurement procedure for fragile X (A) syndrome (CGG repeats)

Measurand	CGG repeats in the *FMR1* gene.
Mnemonic	FRAXA PCR screening.
Test principle	Molecular diagnosis is made by PCR of the relevant part of the *FMR1* gene and measurement of the number of CGG repeats using capillary electrophoresis.
Units	CGG repeats.
Reference intervals	Normal alleles: 5–44 repeats Intermediate alleles: 45–58 repeats Pre-mutation alleles: 59–200 repeats Mutant alleles: >200 repeats
Test limitations	PCR technique does not amplify full mutations. Cannot distinguish between homozygous females from those heterozygous with a normal sized allele and a full mutation allele. If an individual is a mosaic for a full mutation and a normal allele or pre-mutation allele, then the smaller allele will be amplified preferentially and the larger allele missed.
Clinically significant interferences	None
Calibrator traceability uncertainty	Lower range control specimens have been sequenced to accurately determine repeats size. Verified upper control (56 repeats) obtained from the Centers for Disease Control and Prevention (Atlanta, US).
Analytical bias	Analytical bias is corrected using linear regression.
Analytical imprecision (CV_A) Applied Biosystem 310 Genetic Analyser	Internal QC data for 2/09/04 to 20/09/05 QC SD CV 23 repeats 0.30 1.3% 29 repeats 0.34 1.24% 42 repeats 0.47 1.32% 55 repeats 0.53 0.93%
Analytical goal	To distinguish alleles as: • less than, equal to, or greater than 45 CGG repeats • less than, equal to, or greater than 59 CGG repeats.
μ	23 repeats 1.3% 29 repeats 1.24% 42 repeats 1.32% 55 repeats 0.93%
Fit-for-purpose action	Assay is fit-for-purpose for alleles < 43 CGG repeat; and for alleles > 47 CGG repeats and < 57 repeats; and alleles > 61 CGG repeats. Alleles outside these regions should be sequenced where possible.
MU for clinical users	± 2 repeats at 45 repeats ± 2 repeats at 59 repeats

Notes

- Where possible, individuals with an allele between 43–47 CGG and 57–61 CGG repeats should be sequenced to size the allele precisely. However, this may not be technically possible in all cases. Clinicians should be informed of the measurement of uncertainty in these critical regions and should recommend genetic counselling.
- The clinical significance of the reference intervals is not precise due to variable penetrance/stability, and thus careful clinical counselling is essential in the intermediate and pre-mutation ranges.
- Individuals with alleles > 53 repeats should be sent for Southern blotting to remove the possibility of mosaicism for a full mutation. However, this will not provide an accurate sizing in the critical range between 53 and 61 CGG repeats.

A7 Serology Rubella IgG

Measurement Uncertainty

Name of assay	AxSYM® Rubella IgG assay
Manufacturer	Abbott Diagnostics
Sample type	Human serum or plasma (EDTA, heparin or sodium citrate)
Measurand	Plasma/serum rubella antibodies; arbitrary concentration
Interfering factors	Specimens with particulate matter should be clarified by centrifugation. Samples that have been heat treated, are lipaemic or grossly haemolysed, or have obvious microbial contamination should not be used.
Sources of variation	• Mixing of samples before testing • Ambient and incubation temperatures • Volume of reagent and sample pipetted • Time of incubation, delay in pipetting and reading of results • Reagent batch changes • Calibration of instrument • Operator • Optical assembly reading

Test units:	**IU/mL**
Reference intervals:	**< 5.0** IU/mL negative **5.0 to 9.9** IU/mL equivocal (grey-zone) **> 10.0** IU/mL positive
MU as estimated from testing of quality control sample:	**RUB1**
Period of time of QC testing:	From **28/05/2002 to 12/03/2003**
Number of QC sample results in the period:	**122**
Weighted mean value of peer group:	**25.3 IU/mL**
Mean value of laboratory:	**23.4 IU/mL**
ucu (SD):	**4.83 IU/mL**
Expanded uncertainty (U), $k = 2$:	**10. IU/mL**

Document version
QC Officer authorisation
Date

A8 Microbiology

In clinical microbiology, the main contributors to uncertainty for a given measurement procedures are:

- incomplete definition of the particular quantity under measurement (see G1 Measurands)
- uncertainty related to calibration processes (see G4 Measurement uncertainty goals)
- inappropriate calibration function used by an analyser (see G4 Measurement uncertainty goals), interferences (see G1 Measurands), and imprecision (see G2 Imprecision)
- rounding of results especially for cell counts for urines (see G6 Numerical significance).

These sources of uncertainty do not apply in all cases; for each measurement procedure, it is necessary to identify which of these sources should be taken into account.

In the following example, measurement of a common and typical microbiological quantity is presented step by step.

Urine microscopy — white cell count

1. Measurand — urine white blood cells (U-WBC); number concentration

White blood cells (WBCs) appear in urine in response to urinary tract infection. The number of WBCs measured in a urine specimen will determine which culture media are inoculated, and will also influence the interpretation and reporting of bacterial cultures and susceptibility results.

Urine WBC counts are reported in ranges ($< 10 \times 10^6$/L; $10–50 \times 10^6$/L; $50–100 \times 10^6$/L; $> 100 \times 10^6$/L). Algorithms based partly on WBC counts determine which comments are added to final reports. WBC counts that are less than or greater than 50×10^6/L influence choice of media inoculation while WBC counts that are less than or greater than 10×10^6/L influence reporting and comments. For this reason, determining the degree of uncertainty of this measurand is important.

2. Measurement procedure

A sample of the urine specimen is loaded into a counting chamber and the number of WBCs in a given volume of urine is counted manually under light or phase-contrast microscopy.

3. Possible sources of uncertainty

Variable	Within the control of the laboratory?	Need to estimate uncertainty?
Collection technique	No *	No
Transport of specimen	No *	No
Storage of specimen	No *	No
Sampling of specimen	Yes	Yes
Volume of counting chamber	No	No
Phase or light microscopy	Yes	No
Manual counting of cells	Yes	Yes
Operator	Yes	Yes

* These variables are not involved in determining the MU of this measurand

4. Sources of uncertainty to be estimated

Mixing and sampling of specimen, operator, manual counting of cells and calculations will be examined. All of these sources of uncertainty are estimated by repeatability measurements across operators, specimens and days of procedure.

5. Method

Estimates of uncertainty are made for two urine specimens encountered in routine laboratory work — one specimen with a high WBC count (approximately 50×10^6/L) and another with a low WBC count, near the important interpretative cut-off of 10×10^6/L. Boric acid (final concentration of 1.8%) is added to the specimen to preserve the cellular components.

Measurements of uncertainty are performed using three scenarios:

- Urine with high WBC count (different chambers) – over several days, the urine specimen is sampled, counting chambers are loaded and eight operators perform up to 50 WBC counts on the specimen using the routine method of microscopy.
- Urine with low WBC count (different chambers) – over several days, the urine specimen is sampled, counting chambers are loaded and eight operators perform up to 50 WBC counts on the specimen using the routine method of microscopy.
- Urine with low WBC count single sampling (same chamber) – to measure mixing, sampling and equipment variation, six operators perform counts from the same counting chamber after a single sampling procedure.

6. Results

The data for each of the above scenarios are entered into a spreadsheet.

The data can be shown to be normally distributed. The mean, median, SD, CV% and measurement uncertainty (MU), typically 95% confidence level or CV% × 2 are calculated.

7. Summary

Results will indicate that considerable variations exist in these measurements, and that sampling and equipment variations as well as operator factors contribute to MU.

Guidelines as to an acceptable level of uncertainty are not available but awareness of MU and techniques to minimise these variations will improve the quality of results. The laboratory should continue to examine these issues. Estimates of uncertainty based on category intervals (e.g. < 10, 10–50, etc) may provide more realistic results and should be examined. Training of staff and variation between operators should also be explored. Laboratory protocols influenced by cell count results may require adjustment.

A9 Plasma/serum creatinine

Quantity	Creatinine
Measurand	Plasma/serum creatinine; substance concentration
Units	µmol/L
Method	Jaffe method — kinetic colorimetry of alkaline picrate reactivity, rate-blanked with compensation for non-creatinine chromogens at a concentration of 26.5 µmol/L
Measurement procedure	Roche® Modular P unit; Roche® creatinine reagents used as per manufacturer's instructions
Test limitations	Not used for neonatal specimens due to bilirubin interference and foetal haemoglobin if sample haemolysed
Clinically significant interferences	Cephalosporin antibiotics may cause significant false positives Gross haemolysis
Calibrator traceability	Isotope dilution mass spectrometry
Calibrator uncertainty	3.71 µmol/L at 388 µmol/L CV: 0.97% (CI define: 95.5%) Data supplied by manufacturer
Bias	Assumed negligible based on manufacturer method of calibration To be verified using commutable serum-based reference material
Imprecision (CVa) P unit: Instrument 1	Internal QC data for 1/09/05–20/03/06 QC Mean — SD — CV% 62 µmol/L — 2.35 — 3.81 509 µmol/L — 10.01 — 1.97
Analytical goal	CV_I = 4.3% from Westgard database. Desirable goal: 2.2% Minimum goal: 3.2%
u_c	Calibrator MU: 0.97% – not significant Imprecision: 3.8% at 62 µmol/L; 2.0% at 509 µmol/L
Fit-for-purpose action	Suboptimal at clinical decision values Acceptable at high creatinine concentration Maintain performance within top 20% of RCPA QAP participants
MU data made available to clinical users	u = ± 2.5 µmol/L at ~100 µmol/L; ± 10 µmol/L at ~500 µmol/L

A10 Creatinine clearance

Creatinine clearance is derived from measurements of serum (plasma) creatinine, a timed (usually 24 hour) urine collection with measurement of urine creatinine (which, for the purpose of this example, are all assumed to be independent). The total uncertainty of a result is related to the sum of all individual uncertainties, which are produced at each stage of the measuring process. For results derived by multiplication and/or division, the overall uncertainty must be expressed mathematically using fractional standard deviation or CV:

$(u_R/R)2 = (u_X/X)^2 + (u_Y/Y)^2 + (u_Z/Z)^2 + \ldots$

Summation of uncertainties for creatinine clearance calculation, where:

	C = creatinine clearance	mL/second
	P = plasma creatinine	µmol/L
	U = urine creatinine	µmol/L
	V = urine volume	mL
	T = collection period	second
	C = (U × V)/(P × T)	mL/second
Let:	P = 100	uP = 2.5; CV% = 0.025
	U = 10,000	uU = 250; CV% = 0.025
	V = 1500	uV = 15; CV% = 0.01
	T = 24 hours (86,400 seconds)	uT = assume no error

C = (10,000 × 1500)/(100 × 86,400) = 1.74 mL/seconds

u clearance = $C \times \{(u_U/U)/^2 + (u_V/V)^2 + (u_P/P)^2 + (u_T/T)^2\}^{½}$

Then C = 1.74 ± 0.13 mL/second (95.5% CI)

Let:	P = 100	u_P = 5.0; CV% = 0.05
	U = 10,000	u_U = 250; CV% = 0.025
	V = 1,500	u_V = 15; CV% = 0.01
	T = 24 hours (86,400 seconds)	u_T = assume no error

Then C = 1.74 ± 0.18 mL/second (95.5% CI)

Bibliography

ISO (International Organization for Standardization) (1992), *Quantities and Units – Part 0: General Principles* (ISO 31-0:1992(E)), ISO, Geneva.

ISO (1993), *International Vocabulary of Basic and General Terms in Metrology* (VIM), 2nd edition, ISO, Geneva.

ISO (1995), *Guide to the Expression of Uncertainty in Measurement,* ISO, Geneva.

ISO (1999), *General Requirements for the Competence of Calibration and Testing Laboratories* (ISO/IEC 17025), ISO, Geneva.

ISO (2003), *Medical Laboratories – Particular Requirements for Quality and Competence,* (ISO/IEC 15189), ISO, Geneva.

Westgard JO, *Desirable Specifications for Total Error, Imprecision and Bias, Derived from Biologic Variation.*

http://www.westgard.com/biodatabase1.htm (Accessed 7 September 2006).

White GH and Farrance I (2004), Uncertainty of measurement in quantitative medical testing: a laboratory implementation guide, *The Clinical Biochemist Reviews* 25(4):S1–24.

Further reading

Badrick T, Wilson SR, Dimeski G and Hickman PE (2004), Objective determination of appropriate reporting intervals, *Annals of Clinical Biochemistry* 41(5):385–390.

EA (European Co-operation for Accreditation of Laboratories) (2003), EA-4/16. EA *Guidelines on the Expression of Uncertainty in Quantitative Testing,* December 2003 rev00. http://www.european-accreditation.org (Accessed June 2005).

Eurachem (2000), *Quantifying Uncertainty in Analytical Measurement,* Eurachem/CITAC Guide, 2nd edition
http://www.measurementuncertainty.org/mu/guide/index.html (Accessed September 2005).

Fuentes-Arderiu X (2004), Uncertainty of Measurement in Clinical Microbiology, eJIFCC (electronic *Journal of the International Federation of Clinical Chemistry and Laboratory Medicine*) 13(4)
http://www.ifcc.org/ejifcc/vol13no4/130401006.htm (Accessed December 2004).

Hawkins RC and Johnson RN (1990), The significance of significant figures, *Clinical Chemistry* 36(5):824.

Kristiansen J (2001), Description of a generally applicable model for the evaluation of uncertainty of measurement in clinical chemistry, *Clinical Chemistry and Laboratory Medicine* 39:920–931.

Kristiansen J (2003), The guide to expression of uncertainty in measurement approach for estimating uncertainty: an appraisal, *Clinical Chemistry* 49:1822–1829.

Kristiansen J and Christensen JM (1998), Traceability and uncertainty in analytical measurements, *Annals of Clinical Biochemistry* 35(3):371–379.

Krouwer JS (2003), Critique of the guide to the expression of uncertainty in measurement method of estimating and reporting uncertainty in diagnostic assays, *Clinical Chemistry* 49:1818–1821.

NPAAC (2004), *Standards for Pathology Laboratory Participation in External Proficiency Testing Programs*, Appendix 1, NPAAC (National Pathology Accreditation Advisory Council), Australian Government Department of Health and Ageing, Canberra.

Ricos C, Alvarez V, Cava F, García-Lario JV, Hernández A, Jiménez CV, Minchinela J, Perich C and Simón M (1999), Current databases on biological variation: pros, cons, and progress, *Scandinavian Journal of Clinical and Laboratory Investigation* 59:491–500.