

Supplement (ii)

Uncertainty of Measurement in Quantitative Medical Testing

- A Laboratory Implementation Guide -

G.H. White, I. Farrance

on behalf of the

AACB Uncertainty of Measurement Working Group

November 2004

Australasian Association of Clinical Biochemists

PO Box 278,

Mt Lawley, Western Australia, 6929, Australia.

Telephone: +61-8-9370 5224

Facsimile: +61-8-9370 4409

Email: office@aacb.asn.au

Preface

These guidelines have been produced in response to the Australian requirement for medical testing laboratories to be accredited under ISO 15189 as from 1st July 2005. The aim is to provide a general overview of the uncertainty of measurement concept, with minimal metrological terminology, and also practical guidelines to assist pathology laboratories comply with this accreditation requirement.

The guide is not a definitive statement on uncertainty of measurement and may not conform in every aspect to a formal metrological approach. It is not intended to be used as an alternative to more rigorous procedures if these are required (for example; calibration laboratories, manufacturers of reagents and calibrators, reference laboratories, or for the characterisation of definitive or reference methods).

The guide was 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.

In addition to the Australasian Association of Clinical Biochemists (AACB), membership of the working group included representatives from the Australian Institute of Medical Scientists (AIMS), the National Association of Testing Authorities (NATA), the National Pathology Accreditation Advisory Council (NPAAC) and the Royal College of Pathologists of Australasia (RCPA).

Graham White (Convenor)	Flinders Medical Centre, Adelaide.	AACB
Tom Hartley	Royal Hobart Hospital, Hobart.	AACB
Ken Sikaris	Melbourne Pathology, Melbourne.	AACB
John Whitfield	Royal Prince Alfred Hospital, Sydney.	AACB
Ian Farrance	PathCare, Geelong.	NPAAC
John Glasson	IMVS, Adelaide.	AIMS
Tony Barker	LabPLUS, Auckland, New Zealand.	RCPA
Jenny Kox (from March, 2004)	Medical Testing, Melbourne.	NATA
Georgina Kanizaj-Clark (to March, 2004)	Medical Testing, Melbourne.	NATA

The guide provides practical suggestions and worked examples to assist pathology laboratories meet the uncertainty of measurement requirement of ISO 15189 and of ISO/IEC 17025. Currently, the requirement is only applicable to quantitative tests, but many of the concepts can also be applied to test procedures which produce qualitative results. In providing these guidelines, the Working Group recognises that the theoretical and practical aspects of uncertainty of measurement in medical testing are still evolving and that this Guide will require modification in due course.

The Working Group welcomes comment on this Guide

(email: graham.white@fmc.sa.gov.au).

Contents

PART 1	UNCERTAINTY OF MEASUREMENT - OVERVIEW	
	Introduction	S4
	Terminology	S4
	Uncertainty of measurement, traceability and numerical significance	S4
	What is traceability?	S5
	Why uncertainty of measurement?	S5
	What is uncertainty of measurement?	S6
	What is uncertainty of measurement in Medical Testing?	S6
	What does uncertainty of measurement include?	S7
	Working Group Recommendations for the Medical Testing Laboratory	S7
	Analyte and Measurand – an important difference	S8
	Analytical bias (systematic variation)	S8
	Uncertainty of measurement and the number of significant figures reported	S8
	Uncertainty of measurement and fitness for purpose	S8
	Uncertainty of measurement and clinical use	S9
	Summary of Working Group Recommendations for Estimating Uncertainty of Measurement	S9
PART 2	IMPLEMENTING UNCERTAINTY OF MEASUREMENT - PROCEDURES IN THE CLINICAL LABORATORY	
	The uncertainty of measurement data base	S10
	Step 1: Define the measurand	S10
	Step 2: Estimating uncertainty of measurement	S10
	Step 3: Assessing uncertainty of measurement for fitness for clinical purpose	S10
PART 3	THE CLINICAL USE OF UNCERTAINTY INFORMATION	
	Providing uncertainty of measurement data for clinical use	S12
PART 4	APPLYING UNCERTAINTY OF MEASUREMENT TO A LABORATORY METHOD	
	Appendix A: Summary Flow Chart for Applying Uncertainty of Measurement to a Laboratory Method	S14
	Appendix B: Example Records of Uncertainty of Measurement Information for Chemical Pathology Methods	S15
	Appendix C: Combining Uncertainty Components	S18
	Appendix D: Uncertainty in Haemoglobin A _{1c} Measurement	S19
	Appendix E: Inclusion of Uncertainty of Measurement Information with Test Reports – Examples	S21
PART 5	DEFINITIONS	S21
PART 6	REFERENCES AND BIBLIOGRAPHY	S23

PART 1

UNCERTAINTY OF MEASUREMENT - OVERVIEW

Introduction

“A measurement result is complete only when accompanied by a quantitative statement of its uncertainty. The uncertainty is required in order to decide if the result is adequate for its intended purpose and to ascertain if it is consistent with other similar results.”¹

Uncertainty of measurement provides a quantitative estimate of the quality of a test result, and therefore is a core element of a quality system for calibration and testing laboratories. To reflect this, various international metrological and standards bodies jointly developed a Guide to the Expression of Uncertainty in Measurement (GUM) to provide such laboratories with a framework of formal metrological terminology and methodology for expressing uncertainty of measurement.² Subsequently, international standards ISO/IEC 17025 and ISO 15189 (17025 re-written for medical testing), have required complying laboratories to provide estimates of uncertainty for their test measurements, referring to GUM for the appropriate methodology.

The GUM approach was developed primarily for physical measurements, such as length, temperature, weight, electrical conductivity etc., and uses mathematical theory and experimental observation to estimate standard uncertainties for all relevant components of a test procedure. As it was unclear how the GUM approach could be easily applied to pathology testing, the National Pathology Accreditation Advisory Council (NPAAC) deferred compliance with the uncertainty of measurement requirement when it introduced ISO/IEC 17025 as the standard for the accreditation of Australian pathology laboratories in January 2000.

ISO/IEC 17025 and ISO 15189 together briefly outline the two inter-dependent metrological concepts of uncertainty of measurement and traceability. Neither concept should be new to those who work in medical testing; for example, clinical biochemists have for many years sought to achieve traceability by reference to primary standards which have international recognition, and to define uncertainty of measurement by determining the various components of total analytical error. An overview of the merits of a “comprehensive measurement system in clinical chemistry” was provided by Tietz in 1979,³ who described a measurement system comprising a hierarchical structure of Definitive, Reference and Field methods, in association with Primary Reference Materials (Standards), Secondary Reference Materials and Control Materials. National and international proficiency testing

programmes have assisted significantly with conformity to such a measurement system.

In the long term, the practical realisation of traceability of routine methods to internationally recognised standards and estimation of uncertainty of measurement for reported results will bring the benefits of common reference intervals and the comparability of patient results across laboratories and methods. However, at the present time full traceability is limited to a minority of analytical methods routinely used in the clinical laboratory.

In this document the AACB Uncertainty of Measurement Working Group has attempted to provide a practical framework for estimating and reporting the uncertainty of measurement of routine quantitative medical testing procedures which recognises both the special nature of biological measurement and the uncertainty of measurement principles of ISO/IEC 17025 and GUM. With ISO 15189 (*Medical Laboratories – Particular requirements for quality and competence*) replacing ISO/IEC 17025 in July 2005, it is timely for Australian pathology laboratories to commence providing working estimates of uncertainty of measurement for their quantitative test procedures.

Terminology

Medical scientists are not generally familiar with the metrological terminology used in GUM, nor are all such terms necessarily applicable to clinical laboratories. This guide uses pathology laboratory terminology where possible, but retains some key formal terms to assist conformity (see Definitions). The term *uncertainty of measurement* is preferred to *measurement uncertainty*, the former being the term of choice in ISO/IEC 17025 and ISO 15189, and of organisations such as the International Federation of Clinical Chemistry (IFCC) and the National Committee for Clinical Laboratory Standards (NCCLS).

Uncertainty of measurement, traceability and numerical significance

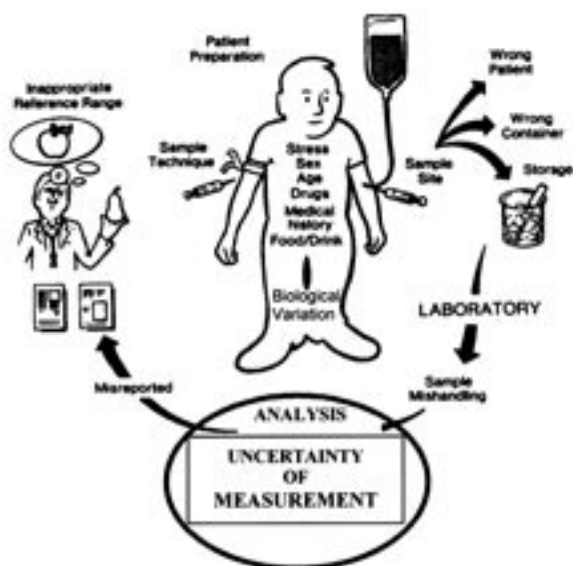
Uncertainty of measurement, traceability and numerical significance are separate but closely related concepts that affect both the format and the information conveyed by a quantitative test result. In addition, the use of SI units provides a consistent basis for the reporting of clinical laboratory data.

- *Uncertainty: “a parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand” (VIM).*

- **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” (ISO 15189).
- **Numerical Significance:** The significant figures of a number are those that have some practical meaning. The significant figures of a number expresses its magnitude to a specified degree of accuracy.⁴
- **Systeme Internationale (SI) units:** The system of metric units which has been adopted by agreement in all major countries for use in science, medicine, industry and commerce. SI is a coherent system based on the seven basic quantities of length (metre, m), mass (kilogram, kg), time interval (second, s), electric current (ampere, A), thermodynamic temperature (degree Kelvin, K), luminous intensity (candela, cd) and amount of substance (mole, mol).

In medical testing there are many potential “uncertainties” that can significantly affect test results (for example; poor specimen collection or transport, patient related factors such as biological variation and the presence of drugs, clerical and reporting errors, etc). Although it is important to identify and minimise such factors (for example, ISO 15189, 5.8.5; “The report shall indicate if the quality of the primary sample received was unsuitable for examination or could have compromised the result”), pre- and post-analytical influences do not affect the inherent uncertainty of the testing procedure itself, and therefore such factors are excluded from the estimation of uncertainty of measurement (Figure 1).

Figure 1.



There is an on-going debate as to how uncertainty should be determined and expressed for measurements of biological substances, with many theoretical and practical issues still needing to be resolved. An outline of the most relevant and controversial factors which affect uncertainty and traceability have recently been highlighted by the opposing views of J.S. Krower⁵ and J. Kristiansen.⁶ In particular, the article by Kristiansen⁶ gives an excellent overview of the basic concepts of uncertainty and traceability and their inter-dependence.

What is traceability?

Traceability and uncertainty are fundamental properties of all quantitative measurements. Because such measurements are made relative to some scale or defined standard, they are by definition *traceable* to this scale or standard. *Traceability* relates a measurement result to a stated metrological reference through an unbroken chain of calibrations or comparisons, each of which may contribute a stated level of uncertainty to the final test result. This unbroken chain of comparisons, which leads back to a known reference value, allows different laboratories (or the same laboratory at different times) to compare results and also relate them to a common measuring scale. The common measuring scales recommended are those of the SI units of measurement.

For example, the defined primary standard for a medical testing method may be an internationally agreed hormone preparation, against which the calibrator value for a commercial kit has been assigned via a chain of intermediate reference preparations. Only when the uncertainty of the value assignment (uncertainty of measurement) for each intermediate reference preparation and for the kit calibrator is known, can traceability of the kit test results be assured (Figure 2).

Why uncertainty of measurement?

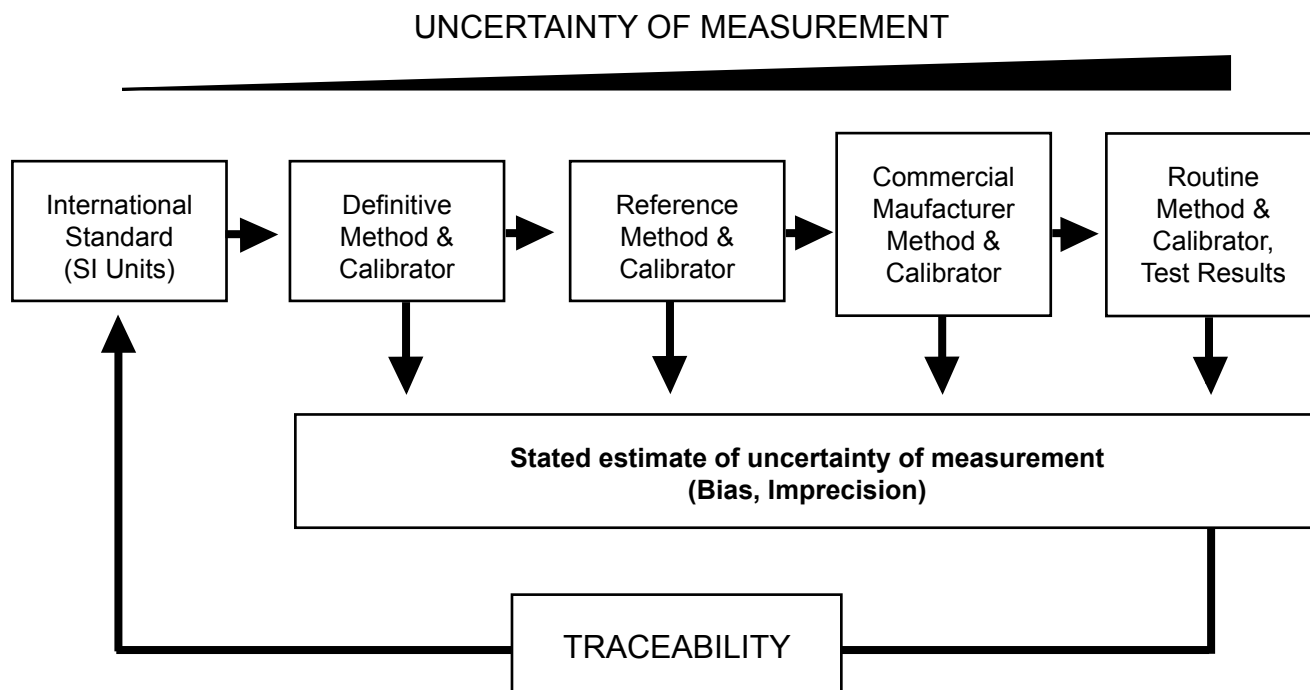
ISO 15189, 5.6.2 requires that “The laboratory shall determine the uncertainty of results, where relevant and possible”.

*The expression of the uncertainty of a result allows comparison of results from different laboratories, or within a laboratory or with reference values given in specifications or standards.*⁷

Laboratories are responsible for ensuring that test results are fit for their clinical purpose by setting and maintaining the quality of their analytical methods, and that the methods used are appropriate for the given clinical application. The principles of estimating uncertainty of measurement contribute to ensuring test outputs are fit for their clinical purpose by:

- Defining what an analytical method measures

Figure 2.



- Meeting a defined analytical goal
- Indicating the confidence that can be placed in a test result
- Contributing to defining, monitoring and indicating where a test procedure may be improved.

What is uncertainty of measurement?

ISO 15189 (3.17): The uncertainty of measurement is a parameter associated with the result of a measurement, that characterises the dispersion of the values that could be reasonably attributed to the measurand.

What is uncertainty of measurement in Medical Testing?

There are two major sources of uncertainty which contribute to the total uncertainty of measurement of a routine quantitative diagnostic method. Firstly, there is uncertainty associated with the numerical value assigned to the measurand present in the calibrator material used in the routine method. This uncertainty should be estimated by the commercial supplier of the calibrator, or by the laboratory if the calibrator has been prepared in-house. The method for estimating the uncertainty of calibrator value(s) will depend on how the value is

determined (for example, gravimetry, definitive method, etc), but for most methods the Type A and B bottom-up approaches described in GUM will be required. At the present time only some commercial manufacturers of calibration materials provide the necessary uncertainty estimates of assigned values.

Secondly, there is uncertainty associated with the value of a test result due to the random errors that normally occur when conducting the testing procedure. This uncertainty component is demonstrated by the *dispersion* of values observed when a measurand in the same specimen is repeatedly measured by a properly conducted test method. In the medical testing laboratory this dispersion is termed *imprecision*, and has long been used as the basic quantitative estimate of the confidence that can be placed in a result.

For practical purposes, imprecision data obtained from the routine application of internal quality control is recommended as the quantitative estimate of the uncertainty of measurement. For laboratory clients (clinicians), the dispersion of test results around a clinical decision value is the major uncertainty that has the potential to affect interpretation and clinical management.

Where the estimate of uncertainty is known for both the calibrator and the routine analytical imprecision of a test

procedure, the total estimate of uncertainty of measurement of the test results can be calculated by summing the two estimates (*as squares of the variances; see Appendix C*).

What does uncertainty of measurement include?

ISO 15189, 5.6.2: “Sources that contribute to uncertainty may include sampling, sample preparation, sample portion selection, calibrators, reference materials, input quantities, equipment used, environmental conditions, condition of the sample and changes of operator”.

The uncertainty of measurement of a test procedure is the sum of the uncertainties associated with the technical steps required to conduct a test according to the standard operating procedure of the method. Where an estimate of uncertainty for a calibrator value is available, then this forms part of the uncertainty of measurement for the testing procedure. Uncertainty components which do not form part of the actual test procedure are excluded from this definition of uncertainty of measurement.

Working Group Recommendations for the Medical Testing Laboratory

The Working Group recognises that the estimation of uncertainty of measurement is a fundamental characteristic of the quality of a quantitative medical testing method, and therefore is an essential requirement for laboratory accreditation. The Working Group also recognises that the implementation of the uncertainty of measurement requirement offers opportunities for pathology laboratories to value-add to their diagnostic services, particularly in educating users to better understand the limitations of tests, and in recognising when clinically significant changes in patient results have or have not occurred.

The GUM is generally accepted worldwide as the master document describing the theory and implementation of uncertainty of measurement. It is based on sound mathematical theory and utilises probability density functions and the law of propagation of uncertainty as the basis for modeling. It outlines procedures for estimating and summing the standard uncertainties of all inputs to the final result of a measurement of a well characterised measurand. GUM does, however, recognise that the formal metrological approach may be difficult to apply to some types of testing. ISO/IEC 17025 and ISO 15189 also recognise that the rigour of estimating uncertainty of measurement may be based on the needs of the client. The Working Group is therefore of the view that in applying the concept of uncertainty of measurement to medical testing, it must be of practical relevance to both the laboratory and the clinical users of the test results. In

this context, the following sections describe the Working Party recommendations, recognising that the theoretical and practical aspects of estimating uncertainty of measurement in medical testing are still evolving and that this Guide will require modification in due course.

As a laboratory generally employs a measurement procedure for long periods of time, the uncertainty of measurement information most relevant to interpreting its test results against fixed reference values is the imprecision of the test results across as many routine operating conditions as possible (for example; multiple calibrator and reagent batches, multiple operators, equipment maintenance, summer/winter etc). With the *caveat* that quality control materials may not totally reflect the analytical behaviour of patient specimens, this imprecision is most easily derived from long-term internal quality control (QC) data, calculated as standard deviation (SD) or coefficient of variation (CV%). For the purpose of recording estimates of uncertainty of measurement the imprecision should be documented as the 95% confidence interval (± 1.96 SD; or ± 1.96 CV%). It should be noted that imprecision derived from the performance of a laboratory in an external quality assurance programme is not recommended for estimating uncertainty of measurement, because generally far fewer data points are available on which to base the uncertainty estimate relative to the number available from internal QC.

Depending on the range of reportable values and clinical use of the test, it may be appropriate to record the estimate of uncertainty of measurement (imprecision) at more than one level of quality control. SD imprecision should be quoted to an appropriate number of significant figures in the same units as the test result (or CV% imprecision to the nearest convenient whole integer), preferably at a level close to a critical clinical decision limit.

For well established methods, it is recommended a minimum of six months internal QC data should be used to calculate routine imprecision, updated at least annually where possible. For new methods, evaluation data comprising at least 30 data points for each level of QC across at least two different batches of calibrator and reagents should be used to provide an interim estimate of uncertainty of measurement.

As part of the initial and ongoing review process, a laboratory should determine whether the uncertainty of measurement estimate for each method is fit for the clinical purpose for which the test results will be used (*see below; Uncertainty of measurement and fitness for purpose*). Reasons for not proceeding for a given method should be documented.

Analyte and Measurand – an important difference

Analyte is a term used to identify the substance or constituent of interest that is the subject of measurement. However, a substance can have a number of properties, some or all of which can be utilised to quantify the substance in an appropriate measuring system. The particular quantifiable property of the analyte used in the measuring system is called the **measurand**.

Analyte	Measurand
Sodium	Urine sodium concentration
Sodium	Plasma sodium activity concentration
Creatine kinase MB	Plasma creatine kinase MB concentration
Creatine kinase MB	Plasma creatine kinase B activity concentration
Alkaline phosphatase	Plasma alkaline phosphatase activity concentration

It is therefore important to accurately identify what is being measured by an analytical method, and this is usually straightforward, as illustrated above. However, there may be significant uncertainty as to the exact nature of the measurand when analytical principles such as immunoassay are considered (for example; prolactin/macroprolactin; PTH/PTH species, hCG/hCG species). Clinically significant cross-reactivities which contribute to negative or positive interference should also be identified and documented.

Analytical bias (systematic variation)

When test results are clinically interpreted by comparison with reference or previous values produced by the same analytical method, analytical bias should not introduce uncertainty additional to the imprecision of the method. However, if results are interpreted using clinical decision limits determined by an analytical method other than the one generating the result, an estimate of *analytical bias* may need to be included in the total estimate of uncertainty of measurement (*Total Analytical Error*) of the method. Where complete traceability is available for a method (for example, stated calibrator bias and imprecision relative to a recognised international standard), it may be appropriate to apply a correction factor so that reported results do not reflect systematic bias. The estimate of uncertainty of measurement of such a method would comprise the stated imprecision of the calibrator and the long-term imprecision of the method, summed as their variances.

Uncertainty of measurement and the number of significant figures reported

The number of significant figures used to report a quantitative result conveys not only the value, but also implies a certainty with which the result has been determined. A perusal of reports from medical testing laboratories often reveals that many results are reported with an apparently high, but misleading, level of certainty. Laboratories need to be aware that clinicians are often unaware of the real imprecision of the results they use, and can be misled by inappropriate use of the number of significant figures with which a test result is reported.

For example, consider serum creatinine results. At a concentration of 150 $\mu\text{mol/L}$, most laboratories will have an SD of approximately 4 $\mu\text{mol/L}$. That is, the result will be within the range 142 – 158 $\mu\text{mol/L}$, 95% of the time. Thus, taking analytical imprecision into account and reporting to three significant figures, it is necessary to round to the nearest 5 $\mu\text{mol/L}$ to adequately convey the range within which the result actually lies.⁸ Analytical precision, and hence the number of significant figures used to report a result, may vary for the one analyte depending upon the concentration of that analyte. The number of significant figures will change at higher concentrations as the imprecision changes, for example, creatinine concentrations above 400 to 500 $\mu\text{mol/L}$ should probably be reported to only two significant figures.

An alternative approach for reporting quantitative results in a format which takes account of the analytical imprecision and the appropriate number of significant figures, is to report the result as a numerical interval.⁹ With this approach, the incremental value chosen for the reporting interval of a given measurand is based on the statistical confidence that two results are analytically different. This type of statistical analysis is dependent on the analytical method and its imprecision.

Uncertainty of measurement and fitness for purpose

ISO 15189, 5.5.1: “*The laboratory shall use examination procedures, ...which meet the needs of the users of laboratory services and are appropriate for the examinations*”.

The fundamental role of medical testing laboratories is to routinely produce test results that are fit for their purpose; that is, they have analytical accuracy and precision that is appropriate for the clinical purpose(s) to which they are applied.

In order to determine whether a method is routinely producing ‘fit for purpose’ results, there needs to be an appropriate analytical goal against which the estimated uncertainty

of measurement (for example, long-term imprecision from internal QC, bias) can be compared. Some methods have internationally agreed analytical goals (for example; cholesterol and haemoglobin A_{1c}), but in their absence various approaches have been used to set relevant goals for bias and imprecision. A widely used and internationally recommended concept is to define the upper acceptable limit for imprecision as a proportion of the intra-individual biological variation of the analyte. With correct choice of the proportionality factor, analytical imprecision should not contribute significant additional variation to the test result when compared with the natural variation of the analyte being measured. Where relevant, a similar approach to goal-setting can be used for Total Analytical Error (bias + imprecision).

ISO/IEC 17025, 5.4.6.2:

NOTE 1 The degree of rigor needed in an estimation of uncertainty of measurement depends on factors such as:

- *the requirements of the test method;*
- *the requirements of the client;*
- *the existence of narrow limits on which decisions on conformance to a specification are based.*

Where the analytical goal for total imprecision is met, the contributing uncertainty components need not be individually identified and estimated unless there is a specific clinical purpose. If the goal is not met, then major contributors ($\geq 30\%$) to the total imprecision should be identified and opportunities for reduction sought. This process may result in a range of outcomes, from change of a step within a method to a change of method. For some methods, analytical goals set by biological variation are unachievable by current technology, or are not relevant to the clinical application. Comparison of a laboratory's internal quality control results with method-related external proficiency testing data will assist with this review process.

Uncertainty of measurement and clinical use

ISO/IEC 17025, 5.10.3.1

- c) *where applicable, a statement on the estimated uncertainty of measurement; information on uncertainty is needed in test reports when it is relevant to the validity or application of the test results, when a client's instruction so requires, or when the uncertainty affects compliance to a specification limit;*

Much uncertainty of measurement data may appear not to have direct clinical value to requesting doctors, but in some specific clinical settings it does have the potential to contribute to patient care. It is therefore important for laboratories to understand the clinical uses of the tests they report, and identify those where uncertainty of measurement information if reported could significantly affect clinical interpretations and patient management. In any case, such information should be readily available through the laboratory on request. Laboratories should also consider clinical uses where it may be appropriate to provide some uncertainty of measurement information as part of individual patient reports. Procedures for informing requesting doctors of relevant uncertainty of measurement information in a clinically meaningful way is a challenge that should be addressed if medical testing laboratories are to fully discharge their accreditation and clinical governance responsibilities.

Summary of Working Group Recommendations for Estimating Uncertainty of Measurement

The estimate of uncertainty of measurement most relevant to the clinical clients of medical testing laboratories is the total imprecision of a quantitative method as reflected by routine QC.

- Define measurand(s) of method, and clinically important limitations and interferences.
- Record mean long-term imprecision of QC as estimate of uncertainty of measurement (± 1.96 SD or ± 1.96 CV%). The uncertainty of the value assigned to a calibrator(s) should be included if available.
- If clinically relevant, set analytical goal for the estimate of the uncertainty of measurement based on clinical use of test results, using biological variation where appropriate.
- If clinically relevant, methods that fail their analytical goal may be reviewed to identify significant contributors to total analytical imprecision.
- Report test results to the number of significant figures that reflect the estimate of the uncertainty of measurement of the method.
- Clinically relevant uncertainty of measurement information should be made available to users, and may be usefully provided with patient test reports in some clinical situations.

PART 2

IMPLEMENTING UNCERTAINTY OF MEASUREMENT PROCEDURES IN THE CLINICAL LABORATORY

The uncertainty of measurement data base

Most of the information required to satisfy accreditation requirements will already be available within the working records of the laboratory. It is recommended however, that a separate uncertainty of measurement database, either electronic or paper, be created. This will facilitate updating uncertainty of measurement information, demonstrating accreditation compliance, and meeting client requests for uncertainty of measurement information. The minimum required fields for such a database are identifiable from the following steps.

Step 1: Define the measurand

Some routine methods have very high analytical specificity for the substance they are designed to measure (for example, methods based on the analytical principle of mass spectrometry), whilst others may reflect the presence of related metabolites or unrelated substances with similar molecular structure or chemical cross-reactivity. Such interfering substances may be naturally or pathologically present in patient specimens, or result from the administration of therapeutic or diagnostic substances. However, test results are usually identified in patient reports by the name of the *analyte* (see Definitions) of interest, and clinical users may not always be aware of the property of the analyte actually measured. Cross-reacting species or interfering substances that can significantly alter the test result may contribute in a variable and often unknown manner. It is therefore important for the laboratory to record:

- the analytical principle
- the substance(s) the method is designed to measure, what is actually measured (measurand), and the unit of measurement
- the diagnostic limitations of the method
- other substances which may cross react sufficiently to affect clinical interpretation.

Step 2: Estimating uncertainty of measurement

Quantitative test results are usually interpreted by comparing the reported value against a reference or clinical decision value, or against a previous test value. For most methods the

reference values used for interpretation have been determined or verified using the same method, and therefore uncertainty of measurement is most usefully estimated by the long-term imprecision obtained from in-house routine quality control data, expressed with 95% confidence limits as ± 1.96 SD or ± 1.96 CV%. The term 'long-term' is arbitrarily defined as the mean of QC values accumulated over a six month period, but should ensure accumulation of sufficient data points across most working conditions to satisfactorily reflect the routine uncertainty of measurement of the method. For newly introduced methods, the imprecision determined during the initial evaluation provides an interim estimate of uncertainty of measurement (a minimum of 30 data points across two or more different batches of reagents and calibrator). Uncertainty of measurement information should be updated at least annually.

For methods that require several levels of quality control material, the laboratory should determine whether the imprecision at the different levels is sufficiently different as to require separate quotation for clinical purposes. If not, a mean ± 1.96 SD (± 1.96 CV%) can be recorded as the uncertainty of measurement estimate.

For some methods, test results are interpreted against reference or clinical decision values that have been determined by a different method. In this situation, the uncertainty of the result includes not only the analytical imprecision of the method, but also any systematic error (method *bias*). For such methods the long-term bias should be recorded, ideally as full calibrator traceability and uncertainty data from the commercial supplier, or in its absence, from proficiency testing (external quality assurance) reports.

Step 3: Assessing uncertainty of measurement for fitness for clinical purpose

Having estimated the uncertainty of measurement of a method in routine use (as long-term imprecision), its *fitness for purpose* with respect to method imprecision should be assessed by comparing it to an appropriate clinical goal. For some measurands, an analytical goal may not be clinically or physiologically relevant. The goal for comparison should be relevant to the clinical application of the test result. An internationally recognised approach for such goal-setting is based on the intra-individual biological variation of the measurand.

There are three levels of analytical goal for imprecision based on intra-individual biological variation:

$$\textit{Optimum: } CV_A = < 0.25 \times CV_I$$

$$\textit{Desirable: } CV_A = < 0.50 \times CV_I$$

$$\textit{Minimum: } CV_A = < 0.75 \times CV_I$$

where: CV_A = Coefficient of variation (analytical), derived from long-term imprecision. The level(s) selected should be close to clinical decision points wherever possible. If CV_A differs markedly at different levels, it may be important that separate CV_A estimates are used at each level.

CV_I = Coefficient of variation (intra-individual), derived from the intra-individual biological variation of the specified measurand (analyte). (*Refer to references for additional information.*)

The most clinically and technically appropriate goal should be set as the minimum for imprecision. If the goal selected compares unfavourably with the imprecision recorded by other methods and laboratories as indicated in external proficiency testing programmes, a more realistic goal or an alternative method should be considered. For analytes where CV_I data is unavailable or the goal is beyond current technology, other criteria may be used (for example, relative performance in external proficiency testing programmes, proportion of reference interval, clinical opinion etc.). For some applications an analytical imprecision goal based on intra-individual biological variation may not be appropriate (for example, serum hCG).

If: $CV_A = > (\text{factor selected}) \times CV_I$

- The method steps/processes contributing $\geq 30\%$ to CV_A should be identified and assessed for opportunities to reduce imprecision. This may not be feasible for fully automated commercial analytical systems.
- If CV_A reduction is unsuccessful or not feasible, it may be appropriate to consider a change of method.

For therapeutic drug assays the intra-individual biological variation component can, if clinically useful, be replaced by the pharmacokinetic variables of drug plasma half-life (t) and dosing interval (T):

$$\text{Desirable } CV_A = < 0.25 [(2^{T/t} - 1) / (2^{T/t} + 1)] \times 100$$

If test results are interpreted using reference or clinical decision values determined by a different method, **bias** should be considered as part of the estimate of uncertainty of measurement and an appropriate analytical goal set.

There are three levels of analytical goal for bias based on biological variation:

$$\text{Optimum: } B_A = < 0.125 (CV_I^2 + CV_G^2)^{1/2}$$

$$\text{Desirable: } B_A = < 0.250 (CV_I^2 + CV_G^2)^{1/2}$$

$$\text{Minimum: } B_A = < 0.375 (CV_I^2 + CV_G^2)^{1/2}$$

where: B_A = Bias (accuracy, systematic variation)

CV_I = Coefficient of variation (intra-individual), derived from the intra-individual biological variation of the specified measurand (analyte).

CV_G = CV of between - subject (inter-individual) biological variation. (*Refer to references for additional information.*)

The most clinically and technically appropriate goal should be set as the minimum for bias. If the goal selected compares unfavourably with the bias recorded by other methods and laboratories in external proficiency testing programmes, a more realistic goal or an alternative method should be considered. For analytes where CV_I/CV_G data is unavailable or the goal is beyond current technology, other criteria may be considered.

For methods where an analytical goal has been recommended by a recognised international authority, this goal should be adopted as the minimum requirement.

For methods where bias and imprecision must both meet performance criteria for clinical applications, the two parameters are conveniently combined as Total Error Allowable (Te_a), for which various levels of analytical goal may be set:

$$\text{Optimum: } Te_a = < 1.65 (0.25 CV_I) + 0.125 (CV_I^2 + CV_G^2)^{1/2}$$

$$\text{Desirable: } Te_a = < 1.65 (0.50 CV_I) + 0.250 (CV_I^2 + CV_G^2)^{1/2}$$

$$\text{Minimum: } Te_a = < 1.65 (0.75 CV_I) + 0.375 (CV_I^2 + CV_G^2)^{1/2}$$

PART 3

THE CLINICAL USE OF UNCERTAINTY INFORMATION

Providing uncertainty of measurement data for clinical use

A summary of the key uncertainty of measurement information for all quantitative routine methods, in a “user friendly” and understandable format, should be available within the laboratory and available to clients of the laboratory service as required. Examples of this availability and the manner in which this information could be distributed may include; display on selected hard copy reports, included in electronic reports, form part of the information available in a departmental or electronic handbook.

When presenting test result and uncertainty data, the laboratory should review the application of SI units and the relevant number of significant figures used for reporting both the numerical result and any uncertainty estimate. The number of significant figures used in reporting a result has the capacity to impart an incorrect impression of the uncertainty of the test measurement if appropriate rounding does not occur. The articles by Badrick⁹ and Hawkings and Johnson¹⁰ provide guidance for reporting results to the appropriate number of significant figures.

Data available to clients should, where appropriate, include:

- Measurand and units of measurement (analyte and the property of the analyte which is actually measured; for example, concentration or activity)
- Method principle
- Traceability, where available, to an international standard (for example, WHO or IRP hormone preparations)
- Limitations to diagnostic use
- Clinically significant negative and positive interferences
- Imprecision at stated concentration and/or key clinical decision limits (uncertainty of measurement)
- Total Error Allowable at stated concentration and/or key clinical decision limits.

For some specific methods or clinical applications, the provision of uncertainty data together with the test result may reduce the potential for significant clinical misinterpretation (for example, immunological-based methods, where antibody specificity, cross-reactivity with closely related species or clinically significant interfering substances are probably unknown to the requester).

It is understandable that for a quantitative pathology test both the clinician and the laboratorian focus on the actual numerical value of the result, neglecting the potential implication of the uncertainty surrounding the value.

In addition to the clinical application of a test result, there are two important aspects which also need to be considered. The most important of these is the *in vivo biological variability* of the measurand, as this is the signal that may differentiate health from disease. The second is the imperfection in the analytical method that may lead to different results on different occasions. It is vitally important that variation due to imperfect analysis (the analytical uncertainty) is less than the measurement signal we are trying to discriminate.

As a general principle, it has been widely suggested that the analytical goal for imprecision of a test method remain below half the intra-individual biological variation ($CV_A < 0.5 CV_I$). If this condition is satisfied and the analytical variability is appropriately less than the biological variability, the test can be confidently used for clinical diagnosis and monitoring. The impact of uncertainty does not end here however, as diagnostic decisions may be made by comparison to a reference population (reference interval or limit) or compared to a diagnostic cut-off. The methods used to establish these diagnostic decision points have their own imperfections, but once established they become set values without variation. Analytical uncertainty will change the “distance” between the test result and the particular cut-off used for comparison. If the “distance” between the test result and the diagnostic cut-off point is less than 1.96 SD, then it cannot be stated (at the usual 95% confidence level) that a repeat analysis would not produce an analytically valid result on the other side of that diagnostic cut-off. This analytical uncertainty should be conveyed to the clinician who might otherwise see the result in more absolute terms.

Clinical monitoring of a patient using quantitative results is different to diagnosis. Firstly, constant analytical bias (systematic error) is cancelled out in monitoring. It does not matter if the initial result is artificially high, when the follow up result will also be higher by the same amount. Secondly, both the initial and final result has an uncertainty, thereby increasing the overall uncertainty when comparing these two values. Statistically, two results need to be more than 2.77

analytical CV_A 's apart (that is, $\sqrt{2} \times 1.96$) before there can be 95% confidence that they are significantly different from an analytical perspective.

If we wish to know if two results on a patient are significantly different also from a biological point of view, we need to additionally allow for the biological variation of the two results. To do this, the analytical variation and the biological variation of one of the results for the measurand are first summed (*see Appendix C, 3.*). The two results being compared need to be more than 2.77 analytical and biological CV 's apart (that is, $2.77 \times \sqrt{(CV_A^2 + CV_I^2)}$) before there can be 95% confidence that the patient's condition may have changed. (It should be noted that such calculations are based on the assumption that measurands show the same biological variation in healthy and ill individuals, for which currently there is little evidence).

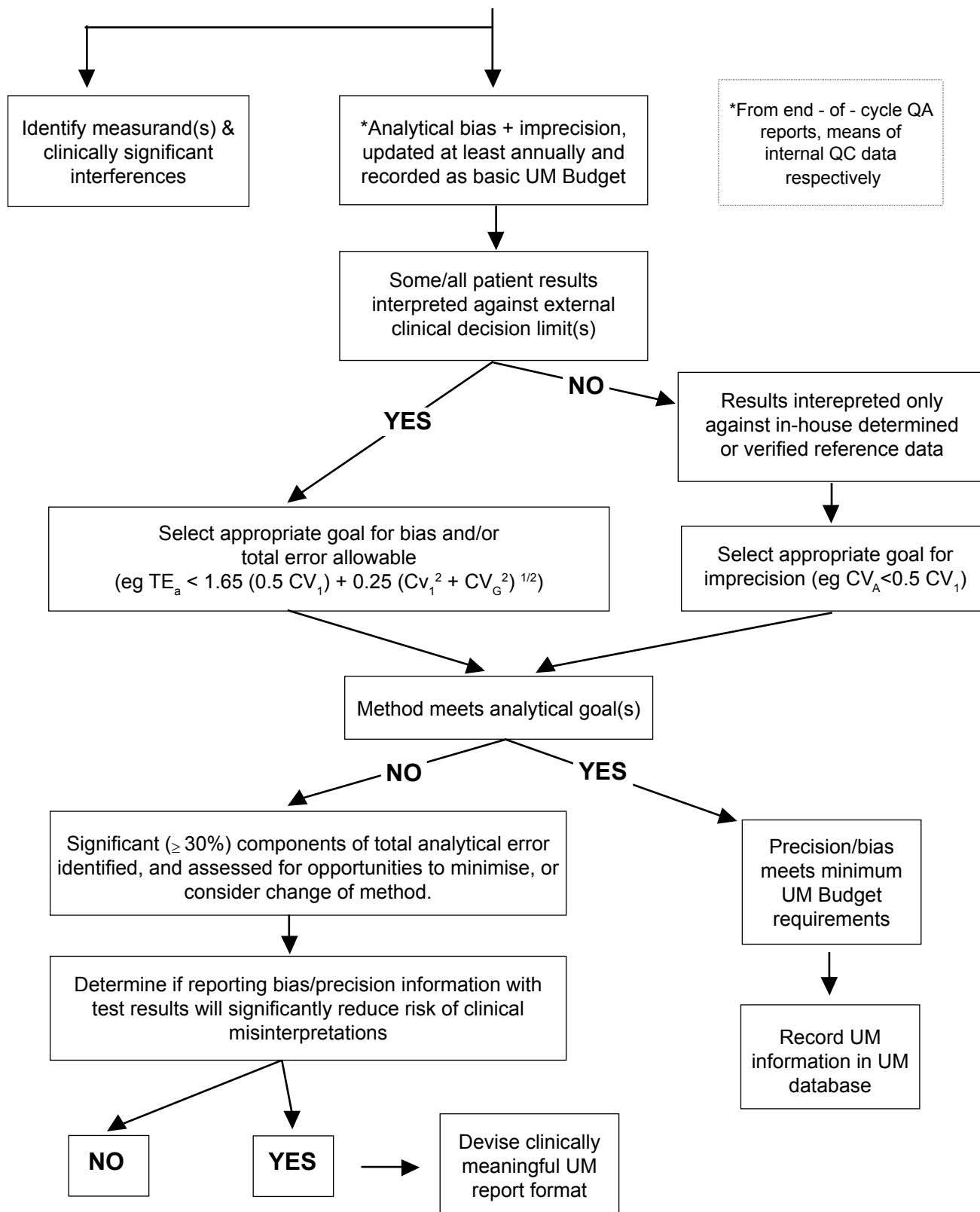
In addition to uncertainty of measurement, it must be remembered that the manner in which numbers are reported from the laboratory also implies a degree of (numerical) uncertainty for the test result. A single ALT result of 125 U/L cannot be analytically differentiated from another result of 126 U/L. Yet, a difference is actually implied when the results are reported to three significant figures in the manner described. Even if these ALT results were rounded to the nearest 10 (that is, to 120 U/L and 130 U/L), the analytical uncertainty limit may still not have been reached. Nevertheless, the use of appropriate rounding and significant figures may be the simplest way to clearly convey laboratory analytical uncertainty.

Comments in patient reports on significant changes in test results can cause confusion as to what is meant. Is an analytical change of any clinical significance? Once again it is important to be mindful of biological variability before claiming that there has been a clinically significant change in a patient's result. When such commenting is used it should be clear and helpful.

Finally, for laboratories to acquire and retain the highest confidence of clinicians and patients, it is vital that they detect quality control violations, monitor and appropriately report the uncertainty of their assays with equal confidence.

PART 4 APPLYING UNCERTAINTY OF MEASUREMENT TO A LABORATORY METHOD

Appendix A: Summary Flow Chart for Applying Uncertainty of Measurement (UM) to a Laboratory Method



Appendix B: Example Records of Uncertainty of Measurement Information for Chemical Pathology Methods

1. Examples of Measurand Information

ANALYTE	hCG	Alkaline Phosphatase	Potassium	Albumin
MEASURAND	Serum/plasma intact β hCG, β hCG subunit concentration	Serum/plasma total alkaline phosphatase activity concentration	Serum/plasma /urine concentration (See separate entry for whole blood potassium)	Serum/plasma/urine concentration
MNEMONIC	β hCG	ALP	K	ALB
TEST PRINCIPLE	2 site sandwich microparticle immunoassay	PNPP/AMP	Valinomycin ion-selective electrode	Bromocresol purple dye-binding spectrophotometry
UNITS	IU/L	U/L	mmol/L	g/L
REFERENCE VALUES (Adults) see handbook for paediatric values	<5 IU/L - non-pregnant adult >25 IU/L - consistent with pregnancy	30-110 U/L - adult	Plasma: 3.2-4.3 mmol/L Serum: 3.5-5.0 mmol/L	In-patient: 31-44 g/L Out-patient: 33-46 g/L
TEST LIMITATIONS	For use in normal pregnancy only. Not to be used as a sole criterion for the diagnosis/management of trophoblastic or non-trophoblastic tumours.	None reported	Haemolysed specimens	Grossly haemolysed specimens
CLINICALLY SIGNIFICANT INTERFERENCES	Heterophilic antibodies; Murine antibodies used for imaging/therapy.	None reported	EDTA	None reported

2. Examples of Uncertainty of Measurement Information

ANALYTE	hCG	Alkaline Phosphatase	Potassium	Albumin
CALIBRATOR TRACEABILITY	WHO 3 rd IS 75/537 – calibrator values (IU/L): 0, 75, 250, 500, 1,000 Uncertainty data not provided by manufacturer	IFCC manual method, 1983 Calibrator value: 224 ± 5.03 U/L (95% confidence) Data from manufacturer	Gravimetry of analytical grade K salt. Low calibrator: 3.00 ± 0.02 mmol/L High calibrator: 7.00 ± 0.05 mmol/L (95% confidence) Data from manufacturer	470 Certified Reference Material (CRM), European Calibrator value: 38.2 ± 0.72 g/L (95% confidence) Data from manufacturer
ANALYTICAL BIAS	Mean 4.5% Cycle 21, QAP Endocrine.	N/A	N/A	N/A
UNCERTAINTY OF MEASUREMENT	Internal QC data for: 1/01/04 - 20/07/04	Internal QC data for: 1/01/04 - 20/07/04	Internal QC data for : 1/01/04 - 20/07/04	Internal QC data for: 1/01/04 - 20/07/04
	QC Mean 6.0 IU/L: SD: 0.47 IU/L CV: 8.2% 22.0 IU/L: SD: 1.2 IU/L CV: 5.4% 1754 IU/L: SD: 81 IU/L CV: 4.7%	QC Mean 87 U/L: SD: 1.28 U/L CV: 1.45% 352 U/L: SD: 3.4 U/L CV: 0.97%	QC Mean 4.2 mmol/L: SD: 0.04 mmol/L CV : 1.06% 6.2 mmol/L: SD: 0.05 mmol/L CV: 0.86%	QC Mean 26.7 g/L: SD: 0.52 g/L CV: 1.95% 38.4 g/L: SD: 0.67 g/L CV: 1.75%

3. Examples of Uncertainty of Measurement and Fitness for Clinical Purpose

ANALYTE	hCG	Alkaline Phosphatase	Potassium	Albumin
ANALYTICAL GOAL	CV _A <5% at 25 IU/L Clinical requirement	CV _i = 6.4% (Westgard website) Desirable: < 0.5 CV _i = <3.2%	CV _i = 4.8% (Westgard website) Desirable: < 0.5 CV _i = <2.4%	CV _i = 3.1% (Westgard website) Desirable: < 0.5 CV _i = <1.6%
CV_A	5.4% at QC mean 22 U/L	1.45% at QC mean 87 U/L	1.06% at QC mean 4.2 mmol/L	1.75% at QC mean 38.4 g/L
FIT FOR PURPOSE	Borderline acceptable	Acceptable	Acceptable	Borderline unacceptable; assess relative QAP performance
REPORTABLE INTERVALS (Approximately 50% confidence – see Reference 8)	<20 IU/L: 1 U/L 20-100 IU/L: 2 U/L >1000 U/L: 100 U/L	< 300 U/L: 1 U/L > 300 U/L: 5 U/L	<10 mmol/L: 0.1 mmol/L	20-50 g/L: 1 g/L
ESTIMATION OF UNCERTAINTY OF MEASUREMENT FOR CLIENT INFORMATION	± 2 IU/L at 25 U/L ± 10% at >1,000 U/L	± 3 U/L for normal values ± 10 U/L at 500 U/L	± 0.1 mmol/L	± 1.5 g/L

Appendix C: Combining Uncertainty Components

When the reported result is derived from more than one actual measurement, the uncertainty of the final result can be calculated by combining the uncertainty components of the contributing measurements. There are mathematical rules that must be followed when adding the individual uncertainty components. There are two formulae which are relevant in this context, and the choice depends on how the final test result is calculated from the contributing measurements.

1. For the estimate of uncertainty of measurement of a final result calculated from a sum and/or a difference of independent measurements (that is, measurements without covariance).

If a result (R) is derived from two (or more) independent measurands (X and Y) by their addition and/or subtraction, then the imprecision of the contributing measurements must be summed as their variances (SD²), that is,

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

Then:

$$(SD_R)^2 = (SD_X)^2 + (SD_Y)^2 \therefore SD_R = \sqrt{(SD_R)^2}$$

where: SD_R, SD_X and SD_Y are the respective analytical SD's.

Example: Uncertainty of measurement for anion gap (AG).¹¹

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 all individual uncertainties which are produced 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):

$$(SD_{AG})^2 = (SD_{Na^+})^2 + (SD_{K^+})^2 + (SD_{Cl^-})^2 + (SD_{HCO_3^-})^2$$

Let: $SD_{Na^+} = 1.2$ mmol/L; $SD_{K^+} = 0.1$ mmol/L
 $SD_{Cl^-} = 1.3$ mmol/L; $SD_{HCO_3^-} = 1.2$ mmol/L

Then:

$$(SD_{AG})^2 = (1.2)^2 + (0.1)^2 + (1.3)^2 + (1.2)^2$$

$$(SD_{AG})^2 = 4.58$$

$$\therefore SD_{AG} = 2.14 \approx 2$$

Anion gap uncertainty (SD_{AG} x 2) = ± 4 mmol/L

2. For the estimate of uncertainty of measurement of a final result calculated from a product and/or a quotient of independent measurements (that is, measurements 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 fractional standard deviations or coefficients of variation (CV):

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

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

where: CV_R, CV_X and CV_Y are the respective analytical coefficients of variation.

Example: Uncertainty of measurement for 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:-

$$(SD_R/R)^2 = (SD_X/X)^2 + (SD_Y/Y)^2 + (SD_Z/Z)^2 + \dots$$

Summation of uncertainties for creatinine clearance calculation, where:-

- C = creatinine clearance ml/sec
- P = plasma creatinine mmol/L
- U = urine creatinine mmol/L
- V = urine volume ml
- T = collection period second
- $C = (U \times V) / (P \times T)$ ml/sec

Let: P = 0.1 SD_p = 0.01
 U = 10.0 SD_u = 0.25
 V = 1500 SD_v = 15

$$T = 24 \text{ hours (86400 secs)} \quad SD_T = \text{assume no error}$$

$$C = (10.0 \times 1500)/(0.1 \times 86400) = 1.74 \text{ ml/sec}$$

SD clearance

$$= C \times \{(SD_U/U)^2 + (SD_V/V)^2 + (SD_P/P)^2 + (SD_T/T)^2\}^{1/2}$$

Then C = 1.74 ± 0.36 ml/sec (clearance ± 2 SD)

Let: P = 0.1	SD _P = 0.02
U = 10.0	SD _U = 0.25
V = 1500	SD _V = 15
T = 24 hours (86400 secs)	SD _T = assume no error

Then C = 1.74 ± 0.70 ml/sec (clearance ± 2 SD)

3. Summation of analytical uncertainty and biological variation.

If the combined estimate (SD_T) of analytical imprecision (SD_A) and within-individual biological variation (SD_I) is required (for example, to compare two test results from a patient), this can be calculated by summing the respective variances as described in 1. above. That is,

$$(SD_T)^2 = (SD_A)^2 + (SD_I)^2.$$

However, if the analytical imprecision is determined at around the same level of measurand as the biological variation (or approximately within the range covered by the biological variation), then CV terms can be used instead of SD. This only applies when the components used within this type of calculation have the same mean.

Thus, for terms with the same mean (or close to the same mean),

$$(CV_T)^2 = (CV_A)^2 + (CV_I)^2$$

This can be shown indirectly as follows,

Assume that: $(CV_T)^2 = (CV_A)^2 + (CV_I)^2$ then,

$$(SD_T/T)^2 = (SD_A/A)^2 + (SD_I/I)^2.$$

If $T = A = I$ (same mean), then $(SD_T)^2 = (SD_A)^2 + (SD_I)^2$.

Example of a clinical application:

Plasma alkaline phosphatase (ALP) activity for Patient:
95 U/L, and two days later 108 U/L.

Uncertainty of measurement (mean imprecision of long-term internal QC) for ALP:

$$CV_A = 1.45\% \text{ at QC mean } 87 \text{ U/L}$$

ALP intra-individual biological variation (CV_w from Westgard website): $CV_I = 6.4\%$

Sum analytical and biological variations as CV's:

$$CV_T = \sqrt{((1.45)^2 + (6.4)^2)} = 6.6\%$$

If the two results are analytically and biologically different, they need to differ by

$$2.77 \times \sqrt{(CV_A^2 + CV_I^2)}, \text{ (for 95\% confidence)}$$

$$2.77 \times 6.6\% = 18.3\%$$

Thus,

$$95 \text{ U/L} + (95 \text{ U/L} \times 18.3\%) = 95 + 17.4 = 112.4 = 112 \text{ U/L}$$

That is, the second result would have to be at least 112 U/L for there to be 95% confidence that it was both analytically and biologically different.

Therefore, 95 U/L and 108 U/L are analytically different, but probably not biologically different.

Appendix D: Uncertainty in Haemoglobin A_{1C} Measurement

Effective control of blood glucose has been shown to significantly reduce complications in patients with diabetes mellitus. The widely regarded Diabetes Control and Complications Trial (DCCT) clearly demonstrated significantly improved outcomes for patients with tight blood glucose control.

The measurement of haemoglobin A_{1C} (HbA_{1C}) is an important marker for long term glycaemic control provided assay procedures and conditions are appropriate and strictly monitored. Only assay procedures which are calibrated with standards traceable to the DCCT and which demonstrate consistently high precision (low analytical CV) are recommended. Based principally on the results of the DCCT, the American Diabetes Association (and supported by a consensus statement from the Australian Diabetes Society, the Royal College of Pathologists of Australasia and the Australasian Association of Clinical Biochemists)¹² has recommended that a primary goal of diabetic therapy is a HbA_{1C} level of less than 7.0%, with re-evaluation of treatment in patients with a HbA_{1C} level consistently greater than 8.0%. These HbA_{1C} values only apply to assay methods that are traceable to the DCCT reference procedure to ensure that the study conclusions and decision points are applicable for interpreting patient HbA_{1C} values. The non-diabetic reference range is generally accepted as 4.0% to 6.0%.

To provide a clear analytical distinction between the recommended HbA_{1c} treatment levels of 7.0% and 8.0%, an analytical (method) uncertainty expressed statistically as a coefficient of variation (CV) of less than 3% is recommended. Methods which produce analytical CV's of 4.0% or greater are not considered appropriate, as this degree of analytical imprecision cannot distinguish changes in the HbA_{1c} level of over 1% (that is, can not distinguish a HbA_{1c} level of 7.0% from 8.0%).

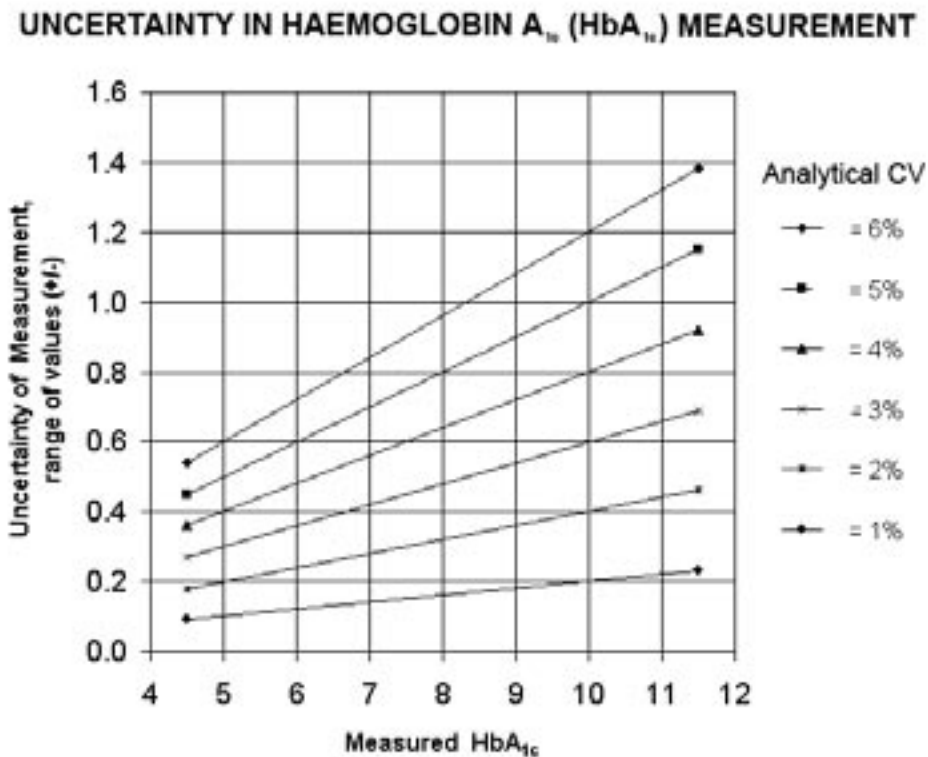
When the analytical CV (uncertainty of measurement) is applied to the actual measured HbA_{1c} value, the possible range of HbA_{1c} values applicable to this measurement can be determined. It is general practice to calculate the 95% confidence interval for pathology measurements and this is achieved by using the plus or minus (\pm) uncertainty range calculated as 1.96 SDs or 1.96 coefficients of variation about the measured value.

For example, if the analytical CV is 3%, a measured HbA_{1c} can only be regarded as falling within a range of values

defined by \pm twice the CV (or \pm 6% of the measured value in this example). Additional examples for a nominal analytical CV of 3% (uncertainty of measurement 6%) and various measured HbA_{1c} levels are shown in the table below.

Measured HbA _{1c} %	Uncertainty of measurement range for an analytical CV of 3%
5.0	4.7 – 5.3
6.0	5.6 – 6.4
6.5	6.1 – 6.9
7.0	6.6 – 7.4
7.5	7.1 – 8.0
8.0	7.5 – 8.5
8.5	8.0 – 9.0
9.0	8.5 – 9.5
10.0	9.4 – 10.6

The chart also provides similar information and a shorthand method of calculating uncertainty or the range for a given HbA_{1c} measurement for a known analytical CV.



Appendix E: Inclusion of Uncertainty of Measurement Information with Test Reports - Examples

1. Prolactin

Macroprolactin is not uncommon in the population and is detected to a variable extent by most serum prolactin immunoassays. This interference can be of sufficient magnitude to cause misdiagnosis and mismanagement of hyperprolactinaemia and prolactinomas. Laboratories using affected assays should consider advising users of this measurand uncertainty when reporting prolactin results above the upper reference interval.

For example:

This assay also detects macroprolactin, which if present in abnormal quantities can falsely raise prolactin results. If this result does not fit clinical expectations, please contact the laboratory.

2. Human Chorionic Gonadotrophin (hCG)

hCG immunoassays vary widely in their ability to detect and quantify hCG and the various hCG-related fragments that arise during pregnancy, and as tumour products. Clinical users are often not aware of the various hCG species detected/not detected by an assay, and of manufacturers warnings concerning test limitations. Laboratories should consider advising clinical users of measurand uncertainties concerning hCG by including an appropriate comment with test results.

For example:

This assay is intended for normal pregnancy applications only, and should not be used as the sole criterion for the diagnosis or management of trophoblastic or non-trophoblastic malignancies.

3. Monoclonal Murine Sandwich Immunoassays

Despite technical precautions some assays remain susceptible to significant positive or negative interference by high levels of animal and heterophilic antibodies in patient specimens. Laboratories should consider advising clinical users of this measurand uncertainty by including an appropriate comment with each report.

For example:

This assay can very occasionally produce falsely high or low results if interfering antibodies are present in the specimen. These occur naturally in individual patients, but can also be due to administration of murine antibodies for imaging/treatment purposes.

4. Digoxin

Some digoxin immunoassays are subject to significant negative interference by high dose steroids. Laboratories should consider advising selected clinical users (e.g. intensivists) of this measurand uncertainty.

For example:

Patients receiving high dose spironolactone or prednisolone may cause significant (up to 50%) negative interference with this digoxin assay.

PART 5 DEFINITIONS

There are many variations of the definitions presented below, but all describe the same essential features, even if presented in a slightly different manner or with different descriptive words. The following definitions are provided for consistency and to assist interpretation.

Accuracy

Closeness of the agreement between the result of a measurement and a true value of the measurand (IUPAC).

Closeness of agreement between a quantity value obtained by measurement and the true value of the measurand (VIM).

Agreement between the best estimate of a quantity and its true value.

See also Inaccuracy.

Analyte

The component of a system to be analysed (IUPAC).

Bias

Systematic error of indication of a measuring system (VIM).

Numerical difference between the mean of a set of replicate measurements and the true value. This difference (positive or negative) may be expressed in the units in which the quantity is measured or as a percentage of the true value.

Error (of measurement)

Difference of quantity value obtained by measurement and true value of the measurand (VIM).

Difference between the estimated value of a quantity and its true value. This difference (positive or negative) may be expressed either in the units in which the quantity is measured or as a percentage of the true value.

Imprecision

Variation of the result in a set of replicate measurements (IUPAC).

Standard deviation or coefficient of variation of the results in a set of replicate measurements. The mean value and

number of replicates must be stated, and the design used must be described in such a way that it can be repeated by other workers. This is particularly important whenever a specific term is used to denote a particular type of imprecision, such as between-laboratory, within-day or between-day.

Inaccuracy

A quantitative term to describe the (lack of) accuracy of a measurement process (IUPAC).

This difference (positive or negative) may be expressed in the units in which the quantity is measured or as a percentage of the true value.

See also Accuracy.

Measurand

Quantity intended to be measured (VIM).

The quantity (property of a body, substance or phenomenon, to which a magnitude can be assigned) subject to measurement. For example the *analyte* may be serum sodium; the *measurand* may be serum sodium concentration or serum sodium activity concentration (as determined by the measurement process).

Measurement method

Generic description of a logical sequence of operations used in a measurement (VIM).

Measurement procedure

Detailed description of a measurement according to one or more measurement principles and to a given measurement method (VIM).

Measurement system

Set of measuring instruments and other devices or substances assembled and adapted to the measurement of quantities of specified kinds within specified intervals of values (VIM).

Metrology

Field of knowledge concerned with measurement (VIM).

Precision

Closeness of agreement between quantity values obtained by replicate measurements of a quantity, under specified conditions (VIM).

The closeness of agreement between independent test results obtained by applying the experimental procedure under stipulated conditions. The smaller the random part of the experimental errors which affect the results, the more precise the procedure (IUPAC).

See Imprecision.

Proficiency Testing

The use of inter-laboratory comparisons to determine the performance of a laboratory with respect to individual

test(s), measurement(s) or observation(s), and to monitor a laboratory's continuing performance. (Standards for Pathology Laboratory Participation in External Proficiency Testing Programs; NPAAC).

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 (ISO 15189, VIM).

A process whereby the indication of a measuring instrument (or a material measure) can be compared with a national or international standard for the measurand in question (ILAC-G2:1994, Traceability of Measurements).

In principle, traceability and uncertainty of measurement are closely interrelated. If the standard or calibrator used in an assay is not traceable to an amount of pure substance then the value of the *measurand* cannot be accurately known. In practice, traceability to an international standard is often outside the control of the laboratory, with reliance on the commercial supplier of the method or reagent kit to establish the chain of traceability. For more specialised and non-automated tests it may be possible for a laboratory to purchase and use a pure substance as a standard; in which case the traceability resides much more within the method and within the control of the laboratory (selection, weighing, preparation and dilution of the standard material).

There are also difficulties associated with the analysis of mixtures, as often occur in biological systems. Method systems which rely on reagent enzymes, antibodies or antigens, are particularly troublesome. Components with similar reactivity in the test matrix, or proteins which have undergone post-translational modification or partial degradation may cross-react to a varying degree, whilst reagent antibodies from different sources may differ in their ability to discriminate between the various components which may be present. Furthermore, a chain of traceability will not be perfect, as errors may be introduced at each of the multiple stages of the analytical process. This is why the definition provides for an unbroken chain of comparisons, *all having stated uncertainties.*

Uncertainty of Measurement

Parameter, associated with the result of a measurement, which characterises the dispersion of the values that could reasonably be attributed to the measurand (ISO 15189 and VIM).

Medical laboratory staff are familiar with the concept that an analytical result is subject to error (uncertainty), and repeated

measurements of the same constituent in the same material will vary. Often, the frequency distribution of these replicate results will approximate the normal (Gaussian) distribution and this can be characterised by a mean, and a variance or SD (and perhaps by skewness and kurtosis, although these are less relevant to this situation). Either the variance or the SD of this distribution of repeated measures is a statistic which “characterises the dispersion of the values that could reasonably be attributed to the measurand”.

The term *measurand* (see definition) may be less familiar than similar terms used previously. It is simply the quantity property of the substance (property, for example; optical density, fluorescence, voltage etc.) which is to be measured. This however, should be taken in context, as it represents (for medical testing) the concentration (or other measurement property) of the target substance in the particular sample. It is not the same as the *analyte*, which is the substance, compound or element being measured.

Medical laboratories have traditionally calculated and quoted the SD of multiple measurements as the assay *precision*, or more correctly the *imprecision*, of the method. Uncertainty of measurement may be expressed as the SD estimated from multiple measurements on internal quality control material or on patient material. The uncertainty of measurement is “associated with the result of a measurement”, and is likely to vary with the magnitude of the result. It may relate linearly to the value of the result, in which case the coefficient of variation (CV) may be constant across the range of values encountered, but this should be checked rather than assumed. At very low levels, the confidence interval for the uncertainty of measurement will overlap with zero and this determines the detection limit of the method. The uncertainty of measurement may be affected by the sample matrix; for example the uncertainty of measurement may differ between urine and serum even if the concentrations in these two matrices are similar.

PART 6 REFERENCES AND BIBLIOGRAPHY

Text References

1. National Institute of Standards and Technology, Background to Uncertainty of Measurement; www.physics.nist.gov.
2. Guide to the Expression of Uncertainty in Measurement (1995) ISO, Geneva. ISBN 92-67-10188-9.
3. Tietz NW. A model for a comprehensive measurement system in Clinical Chemistry. Clin Chem 1979;25:833-9.
4. Porkess R. Dictionary of Statistics. Collins Reference. 1988. ISBN 000-434354-9.
5. Krouwer JS. *Point*. Critique of the Guide to the Expression of Uncertainty in Measurement Method of Estimating and Reporting Uncertainty in Diagnostic Assays. Clin Chem 2003;49:1818-21.
6. Kristiansen J. *Counterpoint*. The Guide to Expression of Uncertainty in Measurement Approach for Estimating Uncertainty: An Appraisal. Clin Chem 2003;49:1822-9.
7. European Co-operation for Accreditation of Laboratories, EAL-G23, Edition 1, August, 1996.
8. Badrick T, Hickman PE. Reporting of creatinine concentration. Ann Clin Biochem 2002;39:162.
9. Badrick T, Wilson SR, Dimeski G, Hickman PE. Objective determination of appropriate reporting intervals. Ann Clin Biochem 2004;41:385-90.
10. Hawkings RCW, Johnson RN. The significance of significant figures. Clin Chem 1990;36:824.
11. Badrick T, Hickman PE. The anion gap – a reappraisal. Am J Clin Pathol 1992;98:249-52.
12. Coleman G, Goodall GI, Garcia-Webb P, Williams PF, Dunlop M. Glycohaemoglobin: a crucial measurement in modern diabetes management. Progress towards standardisation and improved precision of measurement. MJA 1997;167:96-8.

General References and Bibliography

Australian Standard AS ISO/IEC 17025-1999, General Requirements for the competence of testing and calibration laboratories. ISBN 0-337-3304-2. Standards Australia International, Strathfield, NSW 2135, Australia.

European Co-operation for Accreditation, EA-4/16. EA guidelines on the expression of uncertainty in quantitative testing. December 2003 rev00.

Fraser CG. Biological Variation: From Principles to Practice. AACC Press, 2001; ISBN 1-890883-49-2.

Fraser CG, Petersen PH, Libeer J-C, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem* 1997;34:8-12.

Fuentes-Arderiu X. Uncertainty of measurement in Clinical Microbiology. *eJIFCC* vol 13, no 4: <http://www.ifcc.org/ejifcc/vol13no4/130401006.htm> (accessed November 2004).

ILAC G-2: 1994. International Laboratory Accreditation Cooperation (ILAC). www.ilac.org (accessed November 2004).

International Standard ISO 15189; First edition 2003-02-15. Medical laboratories – Particular requirements for quality and competence. Reference number ISO 15189:2003(E).

IUPAC Compendium of Chemical Terminology; International Union of Pure and Applied Chemistry (IUPAC). www.iupac.org/publications/compendium/index.html (accessed November 2004).

International Vocabulary of Basic and General Terms in Metrology (VIM). Second Edition, ISO; 1993; ISBN 92-67-10175-1.

Kristiansen J, Christensen JM. Traceability and uncertainty in analytical measurements. *Ann Clin Biochem* 1998;35:371-9.

Standards for Pathology Laboratory Participation in External Proficiency Testing Programs, 2004; Appendix 1. National Pathology Accreditation Advisory Council (NPAAC), Dept. of Health and Ageing, Canberra, ACT 2601, Australia.

Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, Minchinela J, Perich C, Simon M. Current databases on biologic variation: pros, cons, and progress. *Scand J Clin Lab Invest* 1999;59:491-500.

Westgard JO. Desirable specifications for Total Error, Imprecision, and Bias, derived from biologic variation. www.westgard.com/biodatabase1.htm (accessed November 2004).