Basics of Estimating Measurement Uncertainty

Graham H White
SA Pathology, Flinders Medical Centre, Bedford Park, Adelaide, SA 5042, Australia.
For correspondence: Assoc Prof Graham White e-mail: Graham.White@fmc.sa.gov.au

Summary

- All measurements are imperfect and have many potential sources of variation.
- An estimate of measurement uncertainty (MU) provides an interval of values within which the true value is believed to lie with a stated probability, and is therefore a quantitative indication of the reliability of a measurement.
- MU estimates are essential for assessing whether methods are suitable for clinical use and for comparison of results of a similar type.
- MU estimates can help identify method limitations and opportunities for improvement.

Introduction

Regardless of method, repeated measurements on the same sample will generally produce different results if the system is sufficiently sensitive. The dispersion of results obtained from such repeated measurements (imprecision) can be described approximately by a normal probability (Gaussian) distribution, with some 95% of the results falling within ± 2 standard deviations (SD) of the mean value. Measurement results are thus unreliable and should be regarded as best estimates of the true value of the quantities being measured. Some knowledge of the result variability expected from a given measurement procedure is required if results are to be meaningfully compared with other results of the same kind or with decision and legal limits. For many years the error concept has been widely used, including by clinical biochemistry, to quantify measurement unreliability and to set quality goals for methods.

Error Concept

Briefly, total error comprises systematic (bias) and random error (imprecision), e.g. if a reference material is repeatedly measured by a routine method under replicate conditions, the difference between the mean replicate value and the assigned reference value is the bias of the procedure, and the dispersion of the obtained values, expressed as a standard deviation, is the imprecision. Total error is usually expressed as average bias + 1.65 SD (~95 % confidence). A limitation of the approach is that by adding bias and imprecision total error calculates the worst case error for the procedure, and does not recognise that individual results may have less error.

Concept of Measurement Uncertainty (MU)

In the 1990s it was recognised that measurement comparability between laboratories and methods required an internationally agreed approach to estimating and expressing measurement uncertainty, which is described in the 'Guide to the Expression of Uncertainty in Measurement' (GUM). Although GUM is primarily for measurements in physics, the principles are applicable to biological and chemical measurements. The terminology of the science of measurement (Metrology) is defined in the 'International Vocabulary of Basic and General Terms in Metrology' (VIM).

The basic parameter of MU is the SD, and the symbol for uncertainty is *u*. Some additional useful terms and definitions are given in the table.

In practice, bias correction and replicate measurements can reduce, but not completely eliminate systematic and random errors, and therefore total error cannot be exactly known. It follows that the true value of a measured quantity cannot be exactly known either. This assumption is fundamental to the MU approach. The MU concept also assumes that if the bias of a procedure is known, then steps are taken to minimise it, e.g. by re-calibration. However, because the bias value cannot be known exactly, an uncertainty will be associated with such a correction. Thus, in the MU concept, a measurement result can comprise two uncertainties (i) that associated with a bias correction (u_{Bias}), and (ii) the uncertainty due to random effects (imprecision, u_{Imp}). Both these uncertainties are expressed as SDs which, when combined together, provide the combined standard uncertainty for the procedure (u_{Proc}).

Table. Common metrological terms and their definition (VIM).

Term	Definition
Coefficient of variation (CV); also termed relative standard measurement uncertainty	Standard measurement uncertainty (SD) divided by the absolute value of the measured quantity value. CV = SD/x or SD/mean value.
Combined standard measurement uncertainty (u_c)	Standard measurement uncertainty that is obtained using the individual standard measurement uncertainties associated with the input quantities in a measurement model.
Coverage factor (k)	Number larger than one by which a combined standard measurement uncertainty is multiplied to obtain an expanded measurement uncertainty.
Coverage interval	Interval containing the set of true values of a measurand with a stated probability, based on the information available.
Expanded measurement uncertainty (U)	Product of a combined standard measurement uncertainty and a coverage factor larger than the number one.
Measurand	Quantity intended to be measured.
Measurement uncertainty	Non-negative parameter characterising the dispersion of the quantity values being attributed to a measurand, based on the information used.
Quantity	Property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference.
Standard measurement uncertainty (u)	Measurement uncertainty expressed as a standard deviation.
Trueness	Closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value.
True value	Quantity value consistent with the definition of a quantity.

These and other relevant definitions can be found at the BIPM website http://www.bipm.org/en/publications/guides/vim.html

Thus the MU approach considers a single measurement result to be the best estimate of the measured quantity, and centred on this the combined standard uncertainty provides an interval of values within which the true value of the measured quantity is believed to lie, with a stated coverage probability. For example, a plasma glucose concentration of 5.4 ± 0.1 mmol/L (95% coverage probability), means the true value for the glucose concentration is believed to lie within the interval 5.3–5.5 mmol/L with a probability of 95%.

The GUM describes a bottom-up approach to estimating MU, by which an uncertainty budget for a given measurement procedure is assembled by identifying all potential sources of uncertainty (e.g. calibration, weighing, pipettings, temperature and instrument fluctuations) and attributing to each an uncertainty estimate as an SD obtained by experiment or from available information. The contributing uncertainties are combined in a mathematical model that best represents their interactions in the measurement process. The combining calculation yields the estimated combined standard measurement uncertainty for the whole procedure (u_c). A simple example is the estimation of the concentration of glucose (c) by weighing it (w) into a known volume of water (v). The weighing and volume uncertainties, estimated as SDs, are then combined in the model c = w/v to calculate the combined standard uncertainty of c.

The GUM bottom-up approach can quickly become unwieldy and mathematically complex. Fortunately, clinical biochemistry measurement methods employ quality control (QC) materials to estimate and monitor whole procedure imprecision, so QC data can be used to estimate the contribution of random effects (u_{Imp}) to the measurement uncertainty of the whole procedure (u_{Proc}), with the assumption that the measurand behaves identically in both patient samples and quality control material.

If a procedure has been adjusted for bias, then the uncertainty associated with the correction (u_{Bias}) may need to be combined with u_{Imp} to estimate u_{Proc}. This decision depends

on the magnitude of u_{bias} relative to u_{Imp}. The *t* test can be used to objectively assess the relative significance of u_{bias}, or sometimes a subjective decision is made, e.g. ignore u_{Bias} if it is <10% of u_{Imp}. If u_{Bias} is ignored, then $u_{Proc} = u_{Imp}$. Similarly, if bias itself is assessed as not being significant or it is not evaluated, then again $u_{Proc} = u_{Imp.}$ If u_{Imp} varies significantly across the reportable range, more than one estimate may be required. This top-down approach is generally recognised as a direct estimate of the combined standard uncertainty of the whole procedure (u_c) using the GUM approach.

Estimating MU in the routine biochemistry laboratory

The effort and cost of estimating MU should be commensurate with the clinical quality of measurement required.

MU Goals

It is important that the MU for a given measurement procedure falls within clinically acceptable limits, so that results are of appropriate quality and reliability for patient management. Depending on analyte physiology, specimen type and clinical use of results, MU goals may be based on biological variation, expert group recommendations, or professional opinion. It should be noted that MU goals cannot always be met due to performance limitations of available routine technology.

Step 1: Defining the Measurand

A measurand is the quantity intended to be measured and should be well defined. This is straightforward for analytes that are chemically well characterised, such as sodium or urea, but may be difficult when the measured quantity is method dependent e.g. enzyme activity at a specified pH or temperature, or peptide hormone immunoassays where antibody and epitope specificity may differ between manufacturers. In such cases the measurand is procedure-dependent and measurand definition should include sufficient detail e.g. X Diagnostics two site sandwich β human chorionic gonadotropin immunoassay.

Measurand definition includes the system containing the component (analyte) of interest, e.g. whole blood, plasma, cerebrospinal fluid, and may require further specification (e.g. urine, early morning first void). The measurand description must also identify the kind-of-quantity being examined (e.g. amount-of-substance concentration, amount-of-substance activity, number concentration).

Uncertainty can be associated with the measurand due to:

- Incomplete definition, e.g. antibody interaction with peptide fragments
- Imperfect realisation of the measurand, e.g.
  - incomplete hormone release from natural binding proteins
  - inadequate knowledge of environmental effects on measurements
  - inexact values for standards, calibrators, constants, calculations
  - approximations, assumptions of method
  - unaccounted systematic and random effects.

Step 2: Imprecision of Measurement

Because patient results are compared over time with clinical decision limits or previous results, it is essential to estimate imprecision across as many unavoidable standard operating procedure variables as possible, e.g. calibrator and reagent batch changes, instrument maintenance, different operators, environment (intermediate conditions). The intermediate imprecision (u_{Imp}), expressed as SD, is an estimate of the uncertainty due to the random effects of the whole procedure over time. u_{Imp} may be required at more than one analyte level across the reportable range. If the procedure has not been adjusted for bias, then

$$u_{Proc} = u_{Imp}$$

Step 3: Bias of Measurement

The GUM approach assumes that if the bias of a procedure has been estimated, usually by replicate measurements of a reference material, then it is minimised by re-calibration or by a correction factor. However, bias cannot be known exactly, and so the value used for re-calibration or result correction will have an associated uncertainty (u_{Bias}) due to the uncertainty of the:

- analyte value assigned to the reference material used to assess the bias (u_{Ref})
- mean value of the analyte in the reference material when measured in replicate by the routine procedure (u_{Rep})

u_{Ref} is obtained from the reference material certificate, and u_{Rep} is the standard error of the mean (SEM) of the replicate measurements.

u_{Bias} is calculated by combining the two uncertainties:

$$u_{Bias} = (u_{Ref}^2 + u_{Rep}^2)^{1/2}$$

Hence, the bias of a procedure = Bias value ± u_{Bias}

u_{Bias} should be assessed for significance relative to the procedure imprecision (u_{Imp}) as described earlier. If an appropriate reference material or reference procedure is unavailable, then alternative approaches may be used, e.g. external quality assessment data or inter-laboratory comparisons.

Step 4: Calculation of MU

N.B. Independent uncertainties are combined as variances:

$$u_C = (u_A^2 + u_B^2 + \ldots u_X^2)^{1/2}$$

Scenario 1. Bias ignored or not evaluated (Appendix 1)

If the procedure has not been adjusted for bias, then the estimated MU of the procedure is the intermediate imprecision expressed as 1 SD i.e.:

$$u_{Proc} = u_{Imp}$$

Scenario 2. Bias evaluated (Appendix 2)

If u_{Bias} is assessed as being significant relative to u_{Imp} (e.g. >10%), then it is combined with u_{Imp}, so that:

$$u_{Proc} = (u_{Bias}^2 + u_{Imp}^2)^{1/2}$$

Step 5: MU of a Measurement Result Calculated from Other Measurement Results

If a test result is calculated from the results of other measurements (inputs), then the MU of the final result is obtained by combining the uncertainties of the independent inputs with the same rule used above (square root of the sum of the squares), but the choice of using SD or coefficient of variation (CV) depends on whether the inputs interact by addition or subtraction (use SD), or by multiplication or division (use CV).

Scenario 3. MU of calculated plasma anion gap (Appendix 3)

Anion gap (AG) = (Na + K) – (Cl + HCO_3)

The MU for plasma anion gap (u_{AG}) is calculated by combining the MUs of the analyte results used in the anion gap (AG) equation.

$$u_{AG} = (u_{Na}^2 + u_K^2 + u_{Cl}^2 + u_{HCO3}^2)^{1/2}$$

Because the AG is calculated by adding and subtracting the four contributing results, SDs must be used for combining their uncertainties when calculating u_{AG}.

Scenario 4. MU of calculated creatinine clearance (Appendix 4)

$$\text{Creatinine clearance} = \frac{\text{Urine creatinine (µmol/L) x Urine volume (mL)}}{\text{Plasma creatinine (µmol/L) x Collection time (min)}}$$

The MU for creatinine clearance (u_{CrCl}) is similarly calculated by combining the MUs of the values used in the creatinine clearance equation.

$$u_{CrCl} = (u_{UCreat}^2 + u_{UVol}^2 + u_{PCreat}^2 + u_{Time}^2)^{1/2}$$

However, because the creatinine clearance calculation uses multiplication and division, the contributing MUs must be expressed as CVs when calculating u_{CrCl}.

Step 6: Expression of MU

MU is essentially a normal probability distribution of values within which the true value is believed to lie with a stated probability. The MU parameter is 1 SD. The number of significant digits given for an MU should be the same as that used for reported results. For practical purposes, u is usually expanded (U) by a coverage factor (k), commonly two, to provide an increased coverage probability, e.g.

Plasma glucose (random) = 5.7 ± 0.3 mmol/L (95% coverage probability).

Competing Interests: None declared.

Appendix 1. Scenario 1. Estimation of MU for a glucose measurement procedure ignoring bias uncertainty (artificial data for illustrative purposes).

Parameter	Data	Calculations	Comment
Measurand	Amount-of-substance concentration of glucose in plasma or serum		
Units	mmol/L		
Procedure	Hexokinase/G-6-P dehydrogenase		
Traceability	Isotope Dilution-Mass Spectrometry method; high purity preparation, NIST SRM 917b		From reagent manufacturer information sheet.
Imprecision			
QC Level 1: Mean SD (u_{Imp})	 4.80 mmol/L 0.11 mmol/L	04/02/07-08/05/07 n=317 data points	Imprecision for intermediate conditions included one batch change of reagent(s) and calibrator, several operator changes, one instrument routine maintenance.
QC level 2: Mean SD (u_{Imp})	 15.70 mmol/L 0.38 mmol/L	04/02/07-08/05/07 n=320 data points	
Bias *	Not assessed		EQA program did not indicate bias problem
Measurement Uncertainty u_{Proc}			u_{Proc} considered = u_c
Plasma glucose: 4.8 mmol/L 15.7 mmol/L	 $u_{Proc} = u_{Imp} = 0.11$ mmol/L $u_{Proc} = u_{Imp} = 0.38$ mmol/L		
Expanded uncertainty: 4.8 mmol/L 15.7 mmol/L	 $U = 0.22$ mmol/L $U = 0.76$ mmol/L	$k = 2$	Multiplication of u_{Proc} by a coverage factor of 2 provides an interval of values that is believed to include the true value with a coverage probability of 95%.
Coverage intervals: 5.0-10.0 mmol/L 10.0-20.0 mmol/L	 ± 0.2 mmol/L ± 0.8 mmol/L		For ease of general use by lab staff and for lab 'handbooks', data is rounded. Coverage probability 95%

* Bias was not evaluated or was evaluated and assessed as not significant.

Appendix 2. Scenario 2. Estimation of MU for a glucose measurement procedure including bias uncertainty (artificial data for illustrative purposes).

Parameter	Data	Calculations	Comment
Imprecision (u_{Imp})	0.11 mmol/L		Data from scenario 1. QC Level 1: plasma glucose 4.80 mmol/L.
Bias	Evaluated		Level 3 CRM (certified reference material) 965a measured under replicate conditions (n=10).
CRM 965a: Glucose. Level 3 value: *U:* u_{CRM}:	 6.777 mmol/L 0.073 mmol/L 0.0365 mmol/L	 Calculated as *U*/2	Data from CRM certificate includes assigned value and expanded uncertainty *U* (*k* = 2). Calculate u_{CRM} from expanded uncertainty.
CRM 965a value using routine procedure (n=10) Mean value SD SEM	6.9, 7.0, 6.9, 6.8, 7.2, 7.1, 6.9, 6.8, 7.2, 6.9 6.97 mmol/L 0.149 mmol/L 0.047 mmol/L	 SEM = 0.149/√10 = u_{Rep}	10 measurements of SRM glucose using routine procedure under repeatability conditions (N.B. artificial data for illustration only). Standard error of mean.
Bias *	0.193 mmol/L	6.970-6.777 mmol/L	Mean of replicate measurements by routine procedure minus CRM assigned value.
u_{Bias}	u_{CRM}: 0.0365 mmol/L u_{Rep}: 0.047 mmol/L u_{Bias}: 0.0595 mmol/L	 $u_{Bias} = (0.0365^2 + 0.047^2)^{1/2}$	u_{Bias} is uncertainty of value used for bias correction (0.193 mmol/L). u_{Bias} calculated by combining u_{CRM} and SEM of mean CRM value estimated by routine procedure (u_{Rep}).
Is Bias significant?	 *t* = 3.24 0.01<P<0.02	*t* = Bias/ u_{Bias} 0.193/0.0595 df = 10-1 = 9	*t* test; use distribution table for *t* (2 tailed) values df = degrees of freedom. Bias is significant; hence, procedure re-calibration should include offset of -0.193 mmol/L.
Is u_{Bias} significant?	u_{Bias} / u_{Imp} = 0.54	0.0595/0.11	u_{Bias} is of significant magnitude relative to imprecision, therefore include in MU estimate.
Measurement Uncertainty u_{Proc}	 0.125 mmol/L 0.1 mmol/L (rounded)	 $u_{Proc} = (u_{Bias}^2 + u_{Imp}^2)^{1/2}$ $(0.0595^2 + 0.11^2)^{1/2} = 0.125$	Combined standard uncertainty for whole procedure by combining u_{Bias} and u_{Imp}. u_{Proc} is rounded to same number of decimal places as reported results (unless calculating *U*).
Expanded Uncertainty: U	0.25 mmol/L 0.3 mmol/L (rounded)	*k* = 2; 2 x 0.125 = 0.250	Coverage factor 2 to calculate *U*. *U* is rounded to same number of significant figs as reported results.
Coverage interval: 5.0 -10.0 mmol/L	 ± 0.3 mmol/L		Rounded data for general use by lab staff and 'handbook'. Coverage probability 95%.

* Bias evaluated and procedure adjusted to minimise estimated bias.

Appendix 3. Scenario 3. MU of plasma anion gap (u_{AG}).

Calculation of anion gap	$(Na + K) – (Cl + HCO_3)$
Patient X	Na: 137; K: 4.0; Cl: 106; HCO_3: 10 (mmol/L) AG = 25 mmol/L
MU of AG calculated as square root of sum of squares of MUs of contributing results.	$u_{AG} = (u_{Na}^2 + u_K^2 + u_{Cl}^2 + u_{HCO3}^2)^{1/2}$
Calculation of AG uses addition and subtraction; hence, calculate u_{AG} using SDs.	SD_{Na} = 1.48 mmol/L; SD_K = 0.04 mmol/L; SD_{Cl} = 0.72 mmol/L; SD_{HCO3} = 0.84 mmol/L
Combining u_{Na}, u_K, u_{Cl}, $u_{HCO3.}$	$u_{AG} = (1.48^2 + 0.04^2 + 0.72^2 + 0.84^2)^{1/2} = 1.85$ mmol/L
Coverage factor $k = 2$.	U_{AG} = 3.7 mmol/L (95% coverage probability)
Result rounded for clinical use.	u_{AG} = 4 mmol/L
Patient result.	AG = 25 ± 4 mmol/L (95% coverage probability)

Apendix 4. Scenario 4. MU of calculated creatinine clearance (u_{CrCl}).

Calculation of creatinine clearance (Cl_{cr})	$\frac{\text{Urine creatinine (µmol/L) x Urine volume (mL)}}{\text{Plasma creatinine (µmol/L) x Collection time (min)}}$
Results for 68-year-old woman.	Plasma creatinine: 92 µmol/L, SD: 2.26 (QC), CV: 0.0246 Urine creatinine: 2560 µmol/L, SD: 340 (QC), CV: 0.1328 Urine volume 2683 mL, SD: 25 (estimate), CV: 0.0093 Collection time: 24 h (1440 min), SD: 30 (estimate), CV: 0.0208
Creatinine clearance.	$\frac{2560 \times 2683}{92 \times 1440} = 51.8$ mL/min
MU of Cl_{cr} calculated as square root of sum of squares of MUs of contributing results.	$u_{CrCl} = (u^2_{UCreat} + u^2_{UVol} + u^2_{PCreat} + u^2_{Time})^{1/2}$
Creatinine clearance calculated using division and multiplication; therefore must calculate u_{CrCl} using CVs.	$CV_{CrCl} = (CV^2_{UCreat} + CV^2_{UVol} + CV^2_{PCreat} + CV^2_{Time})^{1/2}$ $= (0.0246^2 + 0.1328^2 + 0.0093^2 + 0.0208^2)^{1/2} = 0.137$
SD = measurement result x CV.	SD = 51.8 mL/min x 0.137 = 7.096 mL/min = u_{CrCl}
Coverage factor $k = 2$.	U_{CrCl} = 14.192 mL/min
Result rounded for clinical use.	U_{CrCl} = 14.2 mL/min
Patient result.	51.8 ± 14.2 mL/min (95% coverage probability)

Further Reading

Useful for Medical Laboratories

Department of Health and Ageing. Requirements for the estimation of measurement uncertainty. National Pathology Accreditation Advisory Council. Canberra, Australia, 2007.

Fraser CG. Biological Variation: From Principles to Practice. Washington, DC, USA: AACC Press; 2001.

Westgard JO. Desirable specifications for total error, imprecision and bias, derived from biologic variation - databases for goal setting. http://www.westgard.com (Accessed 28 February 2008)

White GH, Farrance I. Uncertainty of measurement in quantitative testing. Clin Biochem Rev 2004;25:S1-24.

MU and Analytical Chemistry

EA Guideline A-4/16: Expression of uncertainty in quantitative testing. EA 2003. http://www.european-accreditation.org (Accessed 28 February 2008)

Guide to the evaluation of measurement uncertainty for quantitative test results. Eurolab Technical Report No 1/2006, August 2006. http://www.eurolab.org (Accessed 28 February 2008)

Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation. Eurolab Technical Report No 1/2007. http://www.eurolab.org (Accessed 28 February 2008)

Quantifying uncertainty in analytical measurement. Eurachem/CITAC Guide: 2000.http://www.measurementuncertainty.org (Accessed 28 February 2008)

ISO Normative References, Technical Specification

ISO/IEC Guide 99:2007, International vocabulary of metrology – Basic and general concepts and associated terms (VIM). Geneva Switzerland: ISO; 2007.

ISO/IEC Guide 98:1995. Guide to the expression of uncertainty in measurement. Geneva, Switzerland: ISO; 1995.

ISO/TS 21748:2004. Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation. Geneva, Switzerland: ISO; 2004.