Review Article

Uncertainty of Measurement: A Review of the Rules for Calculating Uncertainty Components through Functional Relationships

Ian Farrance,[1] Robert Frenkel[2]
[1]School of Medical Sciences, RMIT University, Bundoora, Victoria 3083, Australia. [2]National Measurement Institute Australia, Bradfield Road, West Lindfield, NSW 2070, Australia.
*For correspondence: Ian Farrance, e-mail: ian.farrance@bigpond.com

Abstract

The *Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement* (usually referred to as the GUM) provides general rules for evaluating and expressing uncertainty in measurement. When a measurand, *y*, is calculated from other measurements through a functional relationship, uncertainties in the input variables will *propagate* through the calculation to an uncertainty in the output *y*. The manner in which such uncertainties are propagated through a functional relationship provides much of the mathematical challenge to fully understanding the GUM.

The aim of this review is to provide a general overview of the GUM and to show how the calculation of uncertainty in the measurand may be achieved through a functional relationship. That is, starting with the general equation for combining uncertainty components as outlined in the GUM, we show how this general equation can be applied to various functional relationships in order to derive a combined standard uncertainty for the output value of the particular function (the measurand). The GUM equation may be applied to any mathematical form or functional relationship (the starting point for laboratory calculations) and describes the propagation of uncertainty from the input variable(s) to the output value of the function (the end point or outcome of the laboratory calculation). A rule-based approach is suggested with a number of the more common rules tabulated for the routine calculation of measurement uncertainty.

Introduction

With the adoption of the International Organization for Standardization (ISO) laboratory standard *Medical Laboratories – Particular Requirements for Quality and Competence* (ISO 15189, Australian Standard AS 4633), pathology laboratories in Australia (and elsewhere) have been required to provide estimates of measurement uncertainty for all quantitative test results.

Uncertainty of measurement (UM, also referred to as measurement uncertainty, MU), traceability and numerical significance are inter-related concepts that affect both the format and the information conveyed by a quantitative result. As every measurement is prone to error, it is often stated that a measurement result is complete only when accompanied by a quantitative statement of its uncertainty. This uncertainty assessment is required in order to decide if the result is adequate for its intended purpose (*fit for purpose*) and to ascertain if it is consistent with other similar or previous results. A detailed discussion which summarises the concepts of uncertainty of measurement in quantitative medical testing has been provided previously.[1] The development of strategies for setting quality goals in laboratory medicine and procedures for assessing fitness for purpose have been well covered in the clinical biochemistry literature.[2-5] In particular, quality specifications based on biological variation have been discussed in detail by Fraser.[6] The accuracy, precision and fitness for purpose of medical laboratory results rely on the basic metrological concepts of a common system of units, traceability of measured values, uncertainty of measurement and commutability of results within a calibration hierarchy. Metrological traceability in clinical biochemistry is the subject of a recent detailed review.[7]

Uncertainty of measurement is defined by ISO 15189 (item 3.17) as 'a parameter associated with the result of a measurement that characterises the dispersion of values' or, by the *International Vocabulary of Metrology (VIM) - Basic and General Concepts and Associated Terms* (VIM, item 2.26) as a 'non-negative parameter characterising the dispersion of the quantity values being attributed to a measurand'.[8] The *Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement* (usually referred to as the GUM) provides general rules for evaluating and expressing uncertainty in measurement.[9] For those in the medical sciences without a strong background in mathematics, the actual GUM procedures and associated mathematical descriptions can be a real challenge. In addition to the GUM itself, a number of supplements have been proposed which will assist with interpretation and enhance the scope and applicability of the GUM. The status of these additional documents is available on the Joint Committee for Guides in Metrology (JCGM) internet site.[10] The JCGM also provides an introduction to the GUM and proposed GUM supplements. This introduction summarises the main features of uncertainty in measurement and how implementation of the GUM may be achieved within a framework 'supported by mathematical statistics and probability'.[11]

The aim of this article is to provide a general review of the GUM and to provide a more pragmatic approach to the calculation of uncertainty, consistent with the operational procedures most often used within a medical testing laboratory. The suggested approach is rule-based and is defined by the functional relationship (equation) which specifies the input variables for calculating the particular measurand. That is, starting with the general equation for combining uncertainty components as outlined in the GUM, we show how this general equation can be applied to various functional relationships in order to derive a combined standard uncertainty for the output value of the particular function (the measurand). The GUM equation may be applied to any mathematical form or functional relationship (the starting point for laboratory calculations) and describes the propagation of uncertainty from the input variable(s) to the measurand (the end point or outcome of the laboratory calculation). The mathematical procedures used for evaluating the propagation of uncertainty in any particular situation will depend on the form of the functional relationship describing the input and output variables. These variations in the form of the functional relationship provide 'rules' for combining uncertainties in specified circumstances.

Measurand

Measurand is the term that denotes the quantity being measured. It replaces previous terms such as analyte or the name of the substance being measured which was often provided without further definition. VIM defines measurand as the 'quantity intended to be measured' but provides further definition by requiring the inclusion of the measuring system and the conditions under which the measurement is performed. These conditions are required to fully define the measurand, as different measurement procedures may determine different properties or attribute of a substance. For example, the measurement of serum sodium by a direct ion-selective electrode procedure provides a measurand which should be described as serum sodium activity, while serum sodium measured by flame photometry or an indirect ion-selective electrode procedure provides a measurand which should be described as serum sodium concentration.

Further examples and discussion of the term measurand are available in several articles[1,7,12,13] and recent text books on clinical biochemistry.

Uncertainty of Measurement and Measurement Error

The result of any quantitative measurement has two essential components:

- A numerical value (expressed in SI units as required by ISO 15189) which gives the best estimate of the quantity being measured (the measurand). This estimate may well be a single measurement or the mean value of a series of measurements.
- A measure of the uncertainty associated with this estimated value. In clinical biochemistry this may well be the variability or dispersion of a series of similar measurements (for example, a series of quality control specimens) expressed as a standard uncertainty (standard deviation) or combined standard uncertainty (see below).

By definition, the term error (or measurement error) is the difference between the true value and the measured value. The most likely or 'true' value may thus be considered as the measured value including a statement of uncertainty which characterises the dispersion of possible measured values. As the measured value and its uncertainty component are at best only estimates, it follows that the true value is indeterminate (VIM, GUM). Uncertainty is caused by the interplay of errors which create dispersion around the estimated value of the measurand; the smaller the dispersion, the smaller the uncertainty.

Even if the terms error and uncertainty are used somewhat interchangeably in everyday descriptions, they actually have different meanings according to the definitions provided by VIM and GUM. They should not be used as synonyms. The ±

(plus or minus) symbol that often follows the reported value of a measurand and the numerical quantity that follows this symbol, indicate the uncertainty associated with the particular measurand and not the error.

If repeated measurements are made of the same quantity, statistical procedures can be used to determine the uncertainties in the measurement process. This type of statistical analysis provides uncertainties which are determined from the data themselves without requiring further estimates. The important variables in such analyses are the *mean*, the *standard deviation* and the *standard uncertainty of the mean* (also referred to as the *standard deviation of the mean* or the *standard error of the mean*).

Systematic and Random Errors (Uncertainties)

Experimental errors may be divided into two classes: *systematic errors* and *random errors*. Three terms which are often used in association with laboratory errors are accuracy (inaccuracy), bias, and precision (imprecision). Both VIM and GUM define accuracy as a qualitative concept which describes the closeness of agreement between a measured quantity value and a true quantity value of a measurand. As such, accuracy includes the effects of systematic error even though it does not have a numerical value.[8,9] Bias is the term used to describe the magnitude of any systematic error, with VIM defining bias as 'an estimate of a systematic measurement error'.[8] Precision describes the unpredictable (random) variability of replicate measurements of a measurand.

The main distinctions with regard to systematic and random errors are that:

- Systematic error (bias) can, at least theoretically, be eliminated from the result by an appropriate correction.
- Random errors arise from unpredictable variations which influence the measurement procedure, are associated with the actual measurement (for example, failure to properly account for temperature fluctuations or measurement pipette variability), or possible imprecision in the definition of the measurand itself. VIM and GUM refer to such extraneous and environmental factors as *influence quantities*.
- Random errors may be analysed statistically while systematic errors are resistant to statistical analysis. Systematic errors are generally evaluated by non-statistical procedures.
- In clinical laboratories, random error (uncertainty) is usually evaluated through internal quality control procedures.
- The GUM procedures are based on the assumption that all systematic errors have been corrected and the only uncertainty relating to systematic error is the uncertainty of the correction itself. This correction uncertainty and its contribution to the uncertainty of the measurand may be either Type A or Type B depending on the evaluation procedure used (see Type A and Type B uncertainties below).
- The uncertainty in the reported value of the measurand comprises the uncertainty due to random errors and the uncertainty of any corrections for systematic errors.

Systematic error, often referred to as bias, can be identified as a fixed value for a discrepancy and should be corrected at the earliest practical opportunity in the measurement process. For example, if it is known that a particular set of clinical scales has a systematic error or bias of −1.0 gram at a weight of 100 grams (meaning that it reads a true weight of 100 grams as 99 grams), then 1.0 gram should be added to any reading when the scales measure a weight in the vicinity of 100 grams. But, because the bias of −1.0 gram itself has an *uncertainty,* (it might be anywhere between, say −0.9 grams and −1.1 grams, depending on the quality of the calibration procedure), the corrected weight will itself have an uncertainty (in this example, the uncertainty in the systematic error is ±0.1 grams, or a bias of −1.0 ± 0.1 grams).

Alternatively, random error indicates that the error fluctuates over the period of measurement, or from one set of measurements to the next. This variation may be caused by small continuous fluctuations in the environment, in the measuring instrument or at any point in the measurement process. For this reason, it is more appropriate to refer to random error in the plural, as random errors. For example, when an instrument provides a digital reading, random errors may manifest themselves by a fluctuation in the least (or least two or even more) significant digit(s) in the output display. The *range* of the fluctuations is a measure of the *uncertainty* created by the random errors.

Today, many clinical laboratories have more than one instrument (or analytical module) which can perform the same group of tests. Laboratories with automated systems which incorporate several analytical modules providing the same test capability may provide uncertainty estimates for each of the measurands on a system basis irrespective of which module actually produced the result. In this type of automated testing system, any differences which may actually be observed between modules will probably be considered as random effects, provided that all systematic errors have been correctly identified and appropriately corrected. On the other hand, if the uncertainty statement is intended to apply to one specific instrument or module, its imprecision relative to the group will require separate evaluation, provided again that systematic errors have been appropriately corrected.

In general, while error may refer to a single discrepancy (as for a systematic error) or to fluctuations which represent many errors (as for random errors), the effect of all these errors is the uncertainty. Even after all systematic errors have been appropriately corrected, an uncertainty still remains (as discussed above), since the correction itself must contain random errors and is therefore subject to uncertainty. This is why the GUM describes *uncertainty* as a parameter that characterises the *dispersion* of values that may be attributed to the measurand.

Uncertainty of Measurement and GUM Terminology

The definitions of general metrological terms used by the GUM are those provided by VIM, with basic statistical terms taken mainly from International Standard ISO 3534-1. Basic statistical terms and concepts are summarised in Annex C of the GUM, with a glossary of the principal symbols provided in Annex J.

Terms which may be specific to GUM but which now form part of the uncertainty of measurement vocabulary are defined separately within the GUM. The following definitions are those provided in section 2.3 of the GUM as 'terms specific to this guide':

- '*Standard uncertainty*: uncertainty of the result of a measurement expressed as a standard deviation.'
- '*Type A evaluation* (of uncertainty): any method for evaluating uncertainty using statistical analysis of a series of observations.'
- '*Type B evaluation* (of uncertainty): method for evaluating uncertainty by means other than the statistical analysis of a series of observations.'
- '*Combined standard uncertainty*: standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities.' Combined standard uncertainty may contain terms whose components are derived from Type A and/or Type B evaluations without discrimination between types (see below).
- '*Expanded uncertainty*: quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand.'
- '*Coverage factor* (*k*): numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty.' The coverage factor is essentially the same as the Z-score or Z-value in statistical terminology.

The coverage factor used in uncertainty of measurement calculations and the *Z-score or Z-value* used in statistics, are both terms which indicate the number of standard deviations that a particular value may be distant from the mean of the distribution. For clinical laboratory applications, a coverage factor is typically chosen to include approximately 95% of the distribution ($k = 1.96$) or 95.4% of the distribution ($k = 2.0$). For the Gaussian or normal distribution, a Z-score of ± 1.96 from the mean will encompass 95.0% of the values within the distribution, while a Z-score of ± 2.0 from the mean will encompass 95.4% of values. In clinical biochemistry, the mean with a Z-score of ± 2.0 (or ± 2.0 standard deviations) is often used to designate a range which contains 95% of values. From a purist statistical perspective, this is not technically correct as ± 2.0 actually represents 95.4% of values within the distribution. In a similar manner, there are many articles which use a Z-score of 1.96 (to correctly represent 95.0% of values). However, section 6.3.3 of the GUM suggests that for most measurement situations 'where the distribution characterised by y and $u(y)$ is approximately normal and the effective degrees of freedom of $u(y)$ is of significant size … one can assume that taking $k = 2$ produces an interval having a level of confidence of approximately 95%'.

Even though the term standard uncertainty has the same numerical value and mathematical form as a standard deviation, the statistical meaning of standard deviation is not the same as standard uncertainty. In statistics, there are many situations where the standard deviation does not imply the presence of errors or measurement uncertainty. Instead, the standard deviation simply describes the dispersion or spread of observations. Such examples would include the biological reference interval for a measurand or the measurement of the height of adult individuals of a particular ethnic background and gender. Thus, it would not be appropriate to associate the difference between an individual's height and the mean height of the sample as constituting an error and one would not normally regard the dispersion of heights as constituting some kind of uncertainty.

The distinction between *error* and *uncertainty* should again be noted. An error is the discrepancy between a measured value and the actual or true value. Uncertainty is the effect of many errors. This effect may manifest itself as the variability in replicate determinations of the measurand, or, as 'inherited' variability within a single component of the measurand. These two manifestations of uncertainty correspond to the Type A and Type B categories to be discussed below.

In addition to specific descriptive terms such as those outlined above, Annex J of the GUM provides a glossary of the principal symbols used in the various mathematical and

statistical expressions. It is important to fully understand the symbols used in the GUM, particularly those used to describe components of standard deviation, variance and uncertainty. For example:

- $s(x_1)$, $s(x_2)$, $s(x_i)$, $s(x_n)$; as a group of symbols, $s(x_1)$ should be interpreted as the standard deviation of the variable x_1, $s(x_2)$ as the standard deviation of the variable x_2, etc, not as two mathematical variables which should be multiplied together.
- $s^2(x_1)$, $s^2(x_2)$, $s^2(x_i)$, $s^2(x_n)$; as a group of symbols, $s^2(x_1)$ should be interpreted as the squared standard deviation of the variable x_1 or the variance of x_1, $s^2(x_2)$ as the squared standard deviation of the variable x_2 or the variance of x_2, etc, not as two mathematical variables which should be multiplied together.
- $u(x_1)$, $u(x_2)$, $u(x_i)$, $u(x_n)$; in a similar manner to the group of symbols which are used to represent standard deviation and variance (above), as a group, $u(x_1)$ should be interpreted as the standard uncertainty in the variable x_1, $u(x_2)$ as the standard uncertainty in the variable x_2, etc, not as two mathematical variables which should be multiplied together.
- $u^2(x_1)$, $u^2(x_2)$, $u^2(x_i)$, $u^2(x_n)$; as a group of symbols, $u^2(x_1)$ should be interpreted as the squared standard uncertainty of the variable x_1 or the variance of x_1, $u^2(x_2)$ as the squared standard uncertainty of the variable x_2 or the variance of x_2, etc, not as two mathematical variables which should be multiplied together.

When presented in equations such as those provided in the GUM and in Table 2, variance and uncertainty terms need to be considered as a descriptive group, not as individual characters or symbols. For consistency, the terminology used in this review is that used by the GUM, a summary of which is provided in Table 1.

Type A and Type B Uncertainties

The GUM describes two types of uncertainties, categorised respectively as Type A and Type B. A Type A uncertainty is based on the statistical analysis of a series of measurements (for example, statistical data obtained from quality control results). A Type B uncertainty has been obtained by non-statistical procedures and may include:

- Information associated with an authoritative published numerical quantity
- Information associated with the numerical quantity of a certified reference material
- Data obtained from a calibration certificate
- Information obtained from limits deduced through personal experience
- Scientific judgment

The information obtained from Type A and Type B uncertainty evaluations is essentially the same; they are given different names to emphasise that the data have been obtained by different procedures. Type A and Type B evaluations can apply to either random or systematic errors.

Both Type A and Type B uncertainties are characterised by their respective standard uncertainties (equivalent to standard deviations, as described previously) and degrees of freedom (see below). The final result of any measurement procedure should have an associated standard uncertainty obtained by combining the Type A and Type B components. In general, this combination is the root-sum-square of the Type A and Type B standard uncertainties. It should be noted however, that although Type A uncertainties are evaluated using statistical methods, once this evaluation is complete and the results of the measurements are reported, the Type A's effectively become Type B's from the perspective of subsequent users of the results. Type B uncertainties in many cases may therefore be regarded as 'inherited' or 'fossilised' Type A uncertainties. The combined standard uncertainty quoted in a calibration report, for example, is effectively a Type B uncertainty for subsequent use even though it may contain both Type A and Type B components. Very often the combined standard uncertainty is multiplied by a factor (the *coverage factor*) that indicates the level of confidence in the measurement result. This new value (combined standard uncertainty multiplied by a coverage factor) is the *expanded uncertainty*. The usual level of confidence is 95.4% (or 95.0%), in which case it is standard practice to use 2.0 (or 1.96) as the multiplying factor. In this situation, expanded uncertainty is equal to twice the standard uncertainty. For a higher level of confidence the coverage factor is obviously higher (for example, for a confidence level of 99% the factor is approximately 2.6).

The introduction to the GUM provides a clear statement as to the method for evaluating and combining Type A and Type B standard uncertainties. The following points have been summarised from the GUM introduction:

- Type A components are characterised by an estimate of their variances (s_i^2) or their estimated standard deviations (s_i) and the appropriate number of degrees of freedom (see below). A standard deviation (s_i) is numerically identical to a standard uncertainty (u_i). Covariances should be given where appropriate.
- Type B components should be characterised by uncertainty quantities (u_i), which may be considered as approximations to standard deviations. The quantities squared (u_i^2) may be treated like variances or squared standard uncertainties and the quantities themselves (u_i) like standard deviations or standard uncertainties. Covariances should be given where appropriate.

Table 1. Summary of mathematical terms and symbols.

Symbol or term	Description
y	Quantity to be determined, or measurand
x, x_1, x_2, x_i, x_n	Value for a given 'x' variable
s	Sample standard deviation
σ	Population standard deviation
$s(x_1)$, $s(x_2)$, $s(x_i)$, $s(x_n)$	Sample standard deviation for a given 'x' variable
s^2	Sample squared standard deviation or variance
σ^2	Population squared standard deviation or variance
$s^2(x_1)$, $s^2(x_2)$, $s^2(x_i)$, $s^2(x_n)$	Sample squared standard deviation for a given 'x' variable
u	Standard uncertainty
u^2	Squared standard uncertainty, equivalent to variance
$u(y)$	Standard uncertainty in y (measurand)
$u^2(y)$	Squared standard uncertainty in y, or variance of uncertainty in y
$u(x_1)$, $u(x_2)$, $u(x_i)$, $u(x_n)$	Standard uncertainty in a given 'x' variable
$u^2(x_1)$, $u^2(x_2)$, $u^2(x_i)$, $u^2(x_n)$	Standard uncertainty squared for a given 'x' variable
$y = f(x_i)$	Functional relationship between output 'y' and input variable 'x_i'
r	Correlation coefficient
r_{x_1,x_2} or $r_{1,2}$	Correlation coefficient for the variables x_1 and x_2
Z-score, Z-value	The number of standard deviations that an observation or value is from the mean- a standardised variable
k	Coverage factor, the number of standard deviations or Z-scores required to include a stated proportion of values
v	Degrees of freedom
Error	True value – measured value
Population covariance (x_1, x_2)	$\frac{1}{N}\left[\sum_{i=1}^{N}\left(x_{1,i}-\bar{x}_1\right)\left(x_{2,i}-\bar{x}_2\right)\right]$
$\partial y/\partial x_i$	Partial derivative of y with respect to x_i
$\log_{10}(y)$, $\log_{10}(x_i)$	Logarithm to the base 10
$\ln(y)$, $\ln(x_i)$	Logarithm to the base e

- The combined standard uncertainty should be characterised by the numerical value obtained by root-sum-squaring the Type A and Type B standard uncertainties. The combined standard uncertainty is statistically equivalent to a standard deviation.
- If, for a particular application, it is necessary to multiply the combined uncertainty by a factor to obtain an expanded uncertainty, the multiplying factor used must always be stated (for example × 2.0 or × 1.96).
- Once Type A and Type B uncertainties have been combined, they are treated in an identical manner and subsequently described as Type B uncertainties.

Type A and Type B uncertainties are **not** the same as type 1 and type 2 errors as described in many statistical text books. The latter have very specific statistical definitions which are outside the scope of this review.

Degrees of Freedom and Uncertainty in the Uncertainty

The concept of degrees of freedom is closely linked to the process of fitting population values, or *parameters*, to a sample of n observations. It is also linked to the reliability of the uncertainties associated with this fitting process. The general rule for obtaining the number of degrees of freedom, is the number of measurements with error less the number of fitted parameters. Both Type A and Type B uncertainties have an associated number of degrees of freedom (represented in the GUM by v, the Greek letter 'nu'). For Type A uncertainties, the calculation of v is usually straightforward and follows the examples described in standard statistical textbooks. The simplest case involves fitting a single number to a sample of n measurements which is used to derive an estimate of the population mean. This process yields the sample mean which is an estimate (a so-called unbiased estimate) of the population mean. For a sample of n measurements to which the mean has been fitted, the number of degrees of freedom (v) associated with the uncertainties of this fitting is n–1 (that is, $v = n-1$). The reduction from n (the sample size) to n-1 (v, the degrees of freedom) is a statistical consequence of using a sample of n measurements, all accompanied by random errors, to obtain just one estimate (of the population mean).

If points representing paired measurements are used to produce a scatter graph, their linear association may be summarised by fitting two parameters, namely the sample intercept and the sample slope. These are then the unbiased estimates of the population intercept and population slope. In this case, the uncertainties are based on $v = n - 2$ degrees of freedom when the independent (x axis) variable is assumed to be error free.
If v is not given explicitly for Type B uncertainties, it may be roughly estimated from two items of information:

1. The reported standard uncertainty u (if what has been reported is an expanded uncertainty, then this should be divided by the appropriate factor for conversion back to a standard uncertainty),
2. An estimate of the uncertainty Δu in this standard uncertainty itself.

The suggested formula for this approximate calculation is:

$$v = (1/2)\,(u/\Delta u)^2$$

The procedures for determining degrees of freedom are often complex statistical processes which vary according to the particular problem under consideration. Appendix G of the GUM discusses in detail a number of the procedures which may be used for calculating degrees of freedom. The equation suggested above for estimating v when considering Type B uncertainties, is essentially equation G3 as given in the GUM. This equation also illustrates a general characteristic of v and its relationship to uncertainty; that is, the smaller the v, the more uncertain we are of the uncertainty itself. This statement applies equally to both Type A and Type B uncertainties.

The equation suggested for estimating v (above), may also be used to provide an estimate of the uncertainty in the uncertainty. Surprisingly, this uncertainty (Δu) in an uncertainty (u) can be relatively large. For example, suppose a sample of n measurements, all subject to random error, are taken in order to estimate the mean of a particular population and the standard uncertainty in that mean. If s is the standard deviation for the particular sample, then the *standard uncertainty of the mean* is given by $s/\sqrt{n}$ (provided the individual results are uncorrelated as outlined in Appendix B). If n = 10, therefore implying 9 degrees of freedom, the formula can be used to estimate the relative uncertainty in the uncertainty (that is, the proportion $\Delta u/u$). In this example, the relative uncertainty in the standard uncertainty of the mean is 24%.

Functional Relationships, Input and Output Quantities

In many situations, the measurand is not measured directly but is calculated from other measurements through a functional relationship. From a mathematical perspective, a function can be regarded as a quantity whose value can be derived from one or more input quantities by applying a defined mathematical formula. In general, a functional relationship indicates that a mathematical relationship exists between the value of the function (output) and the input variables, without specifically identifying the exact mathematical form of the relationship.

The terminology used in mathematical and statistical texts to express a function or a functional relationship can vary considerably. A summary of the various mathematical terms and symbols used in this review is provided in Table 1.

The notation generally used to describe the combination of input variables through a functional relationship is $y = f(x_1, x_2, x_3, \ldots, x_n)$, where $x_1, x_2, x_3, \ldots, x_n$ are the input quantities upon which the output quantity y depends. In this case, the measurand y is not measured directly, but determined from n other quantities $x_1, x_2, x_3, \ldots, x_n$, through the functional relationship. These input quantities themselves can often be viewed as measurands, which may also depend on other quantities including corrections and correction factors for systematic error. Such interactions may well lead to a complicated functional relationship but the general formula described in the GUM is applicable to even this type of situation. The evaluation of functions which contain many input quantities, including correction factors, can be used to demonstrate how the combined uncertainty is proportioned across the individual input quantities.

Section 4.1.3 of the GUM describes how a set of input quantities may be characterised as:

- 'Quantities whose values and uncertainties are directly determined in the current measurement. These values and uncertainties may be obtained from, for example, a single observation, repeated observations or judgment based on experience, and may involve the determination of corrections to instrument readings and corrections for influence quantities, such as ambient temperature, barometric pressure, and humidity.'
- 'Quantities whose values and uncertainties are brought into the measurement from external sources, such as quantities associated with calibrated measurement standards, certified reference materials, and reference data obtained from handbooks.'

Propagation of Uncertainty

As indicated previously, a measurand may often be calculated using a functional relationship involving other measured quantities and their measured uncertainties. If measurand y is determined from input variables $x_1, x_2, \ldots, x_n$ through a functional relationship, then uncertainties in the x's will *propagate* through the calculation to an uncertainty in y.

There are essentially two types of approach for the estimation of uncertainty in a measurand. When a measurand can be determined by direct measurement (such as the measurement of serum sodium activity), uncertainty may be estimated from replicate analysis of the same or similar specimens. In contrast to direct measurement, the indirect approach requires the measurand to be calculated from other measured quantities (for example, the calculation of creatinine clearance or serum anion gap[1]). This latter approach relies on calculations derived from the propagation of uncertainty formula and may include the following steps:

- Multiple measurements of the various input variables in order to provide an estimate of their relevant uncertainties.
- Calculation of their mean and standard deviation.
- Combining the standard deviations for each of the input variables to give a combined standard deviation (combined standard uncertainty) using the appropriate rule as outlined in Table 2.

Uncertainty in Measurements With and Without Correlation

The manner in which errors (uncertainties) are propagated from measured values to a calculated quantity through a functional relationship provides much of the mathematical challenge to fully understanding the GUM. When the individual input quantities $x_1, x_2, x_3, \ldots, x_n$, are subject to random uncertainties, the variability in each term will usually be considered (for clinical laboratory applications) as following a Gaussian or 'normal' distribution, with a spread or dispersion characterised by standard deviations $s_1, s_2, s_3, \ldots, s_n$. These standard deviations are the standard uncertainties associated with each of the x's. The question now becomes: what is the distribution of values attributable to the measurand y (that is, standard uncertainty $u(y)$) given the known dispersion for each of the measured x values and the functional relationship $y = f(x_1, x_2, x_3, \ldots, x_n)$?

When considering the overall relationship between the various input variables, two situations require consideration. Usually, the input variables (that is the various x's) have no relationship with each other, except through their functional relationship which defines the measurand y. Under these circumstances, the input variables are described as having *zero* covariance or correlation. On the other hand, there are situations where a separate relationship may exist between two or more of the input variables. For example, if the measured value of x_1 is related to the measured value of x_2 or x_3, …, or x_n, the input variables in question are described as correlated or as measurements with covariance. Procedures for the propagation of uncertainty arising from both uncorrelated and correlated variables can be derived from the general expression describing the propagation of uncertainty as outlined in the GUM. These procedures are discussed in more detail below.

As the statistical notation for correlation involves correlated variables, but not necessarily correlated uncertainties, an

example may provide the best explanation:

- Assume that two identical mercury-in-glass thermometers have been calibrated and that each has a standard uncertainty of 0.5° C within the range 0° C to 50° C.
- They are placed in fixed positions in different rooms in a house. Their readings, T1 and T2, are taken once every 24 hours for a full year. There will thus be 365 pairs of values for T1 and T2.
- In this scenario, it is not unreasonable to expect T1 and T2 to be highly positively correlated as both will be high in summer and both will be low in winter, etc. The uncertainties u(T1) and u(T2) however, will not change but will remain at 0.5° C throughout the year.

Where correlated inputs require consideration, additional information is necessary to ascertain the actual numerical correlation between the particular input variables. This is usually obtained by calculation of the correlation coefficient. For most situations which occur in clinical biochemistry, functional relationships which rely on random uncorrelated variables and uncertainties are usually applicable.

From a statistical perspective, both covariance and correlation are measures of the interdependence between two or more random variables. Covariance ('co-variation') provides a measure of the interdependence of variables that may have different units of measurement, while correlation standardises the measure of interdependence by providing a dimensionless quantity, the correlation coefficient r, which facilitates the comparison of different data sets. The two concepts are themselves interrelated with the calculation of covariance forming an intermediate step in the calculation of correlation coefficient.

Correlation coefficient (r) = [covariance (x_1, x_2)] / [$s(x_1) \times s(x_2)$]

As a direct consequence of its definition, the correlation coefficient r is restricted to the range −1 through zero to +1. When $r = 0$, the correlation is zero, when $r = +1$, the correlation is perfect and positive, and when $r = -1$ the correlation is perfect and negative.

Several sections of the GUM discuss the effect which correlated variables may have on the calculation of uncertainty. In particular, section 5.2, Annex C (sections C3.4 and C3.6) and Annex F of the GUM provide detailed comments with regard to covariance and correlation. Uncertainty measurements with and without covariance or correlation are discussed in more detail below and in Appendix B of this review.

Uncertainty Components When the Inputs are Uncorrelated

A general equation which describes the propagation of uncertainty can be derived using standard statistical procedures. This general equation is independent of the exact form of the functional relationship. It can be simplified significantly for particular functional relationships to provide defined 'rules' for evaluating uncertainties in specified situations. This is the basis of the 'rules' approach provided in Table 2.

Given the general expression for the measurand y in terms of input quantities $x_1, x_2, x_3, \ldots, x_n$ (that is $y = f(x_1, x_2, x_3, \ldots, x_n)$), a corresponding expression which describes the combined standard uncertainty in y (that is $u(y)$), can be obtained by appropriately combining the standard uncertainties of the various input estimates (the various x's). There are two versions of this uncertainty expression:

- The 'simpler' version applies to input quantity uncertainties which are relatively small and the input variables $x_1, x_2, x_3, \ldots, x_n$ are uncorrelated.
- The second version applies to input quantity uncertainties which are relatively small but where the input variables are correlated.

The 'simpler' version, which requires the input quantities to be uncorrelated, is given by the following general expression,

$$u^2(y) = \left(\frac{\partial y}{\partial x_1}\right)^2 u^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 u^2(x_2) + \ldots + \left(\frac{\partial y}{\partial x_n}\right)^2 u^2(x_n)$$

(equation 1)

This is the same equation which describes the combined standard uncertainty provided by the GUM as equation 10, section 5.1.2.

$$u^2(y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i)$$

(equation 2)

These equations and their counterpart for correlated input quantities state that the squared combined standard uncertainty $u^2(y)$ of the measurand y, is a weighted sum of the squared standard uncertainties $u^2(x_i)$ of the input quantities x_i (where $i = 1, 2, 3, \ldots, n$). The partial derivatives are often referred to as *sensitivity coefficients*. They describe how the output estimate in y depends on the individual uncertainties

Table 2. Rules for the evaluation of standard uncertainty through functional relationships with uncorrelated variables.

Rule	Notes (below)	Function	Expression giving standard uncertainty
1		$y = x_1 + x_2$	$u^2(y) = u^2(x_1) + u^2(x_2)$
2		$y = x_1 - x_2$	$u^2(y) = u^2(x_1) + u^2(x_2)$
3	1	$y = A + Bx_1 + Cx_2 \ldots + Nx_n$	$u^2(y) = B^2u^2(x_1) + C^2u^2(x_2) + \ldots + N^2u^2(x_n)$
4		$y = x_1 / x_2$	$(u(y)/y)^2 = \left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2\right]$
5	1	$y = Ax_1 / Bx_2$	$(u(y)/y)^2 = \left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2\right]$
6		$y = x_1 \times x_2$	$(u(y)/y)^2 = \left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2\right]$
7	1	$y = Ax_1 \times Bx_2$	$(u(y)/y)^2 = \left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2\right]$
8		$y = (x_1 \times x_2) / (x_3 \times x_4)$	$(u(y)/y)^2 = \left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2 + (u(x_3)/x_3)^2 + (u(x_4)/x_4)^2\right]$
9	1	$y = x^A$	$(u(y)/y) = \lvert A \rvert (u(x)/x)$
10	1	$y = (x_1 / x_2)^A$	$(u(y)/y)^2 = A^2\left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2\right]$
11	1	$y = (x_1)^A \times (x_2)^B$	$(u(y)/y)^2 = A^2(u(x_1)/x_1)^2 + B^2(u(x_2)/x_2)^2$
12		$y = \ln x$	$u(y) = u(x)/x$
13	1	$y = A + \ln x$	$u(y) = u(x)/x$
14	1	$y = A + \ln Bx_1 + \ln Cx_2$	$u^2(y) = (u(x_1)/x_1)^2 + (u(x_2)/x_2)^2$
15	1, 2, 3	$y = \log_{10} x^A = A \log_{10} x$	$u(y) = \lvert A \rvert (u(x)/x) \log_{10} e$
16	1, 2, 3	$y = A + \log_{10} x$	$u(y) = (u(x)/x) \log_{10} e$
17	1, 2, 3	$y = A + \log_{10} Bx_1 + \log_{10} Cx_2$	$u^2(y) = (\log_{10} e)^2\left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2\right]$
18	1, 2	$y = Ae^{Bx}$	$u(y)/y = \lvert B \rvert u(x)$
19	4	$y = x^w$	$(u(y)/y)^2 = w^2\left(\left[\frac{u(x)}{x}\right]^2 + (\ln x)^2\left[\frac{u(w)}{w}\right]^2\right)$
20	4	$y = (x_1 / x_2)^w$ writing $q = x_1/x_2$ for brevity	$(u(y)/y)^2 = w^2\left(\left[\frac{u(x_1)}{x_1}\right]^2 + \left[\frac{u(x_2)}{x_2}\right]^2 + (\ln q)^2\left[\frac{u(w)}{w}\right]^2\right)$

Rule	Notes (below)	Function	Expression giving standard uncertainty
21	5	$y = f(x_1, x_2, ..., x_n)\ g(x_1, x_2, ..., x_n)$	$u^2(y) = (g\partial f / \partial x_1 + f\partial g / \partial x_1)^2 u^2(x_1) + (g\partial f / \partial x_2 + f\partial g / \partial x_2)^2 u^2(x_2) + ...$ $... + (g\partial f / \partial x_n + f\partial g / \partial x_n)^2 u^2(x_n)$
22	5	$y = f(x_1, x_2, ..., x_n) / g(x_1, x_2, ..., x_n)$	$u^2(y) = (1/g)^4 \left[(g\partial f / \partial x_1 - f\partial g / \partial x_1)^2 u^2(x_1) + (g\partial f / \partial x_2 - f\partial g / \partial x_2)^2 u^2(x_2) + ... \right.$ $\left. ... + (g\partial f / \partial x_n - f\partial g / \partial x_n)^2 u^2(x_n) \right]$
23		$y = (\ln x_1 - \ln x_2) / (x_3 - x_4)$	$u^2(y) = \left[1/(x_3 - x_4)^2 \right] \left[(u(x_1)/x_1)^2 + (u(x_2)/x_2)^2 \right] +$ $\left[(\ln (x_1 / x_2))^2 / (x_3 - x_4)^4 \right] \left[u^2(x_3) + u^2(x_4) \right]$
24	1, 4	$y = a(bx_1)^{D1}(cx_2)^{D2}$	$(u(y)/y)^2 = (D_1 u(x_1)/x_1)^2 + (D_2 u(x_2)/x_2)^2 + (u(a)/a)^2 + (D_1 u(b)/b)^2 + (D_2 u(c)/c)^2$
25	4	$y = a(bx_1)^{w1}(cx_2)^{w2}$	$(u(y)/y)^2 = (w_1 u(x_1)/x_1)^2 + (w_2 u(x_2)/x_2)^2 + (u(a)/a)^2 + (w_1 u(b)/b)^2 + (w_2 u(c)/c)^2$ $+ (\ln bx_1)^2 u^2(w_1) + (\ln cx_2)^2 u^2(w_2)$

Notes:

1. A, B, C, D, D_1, D_2 and N (all upper case) are constants with no uncertainty. They may be integers such as 2 or 3, a decimal number, a mathematical constant such as π, negative or positive.
2. e is Euler's number and e^x is the exponential function. The value of e is approximately 2.7183.
3. $\log_{10} e$ is approximately 0.4343.
4. a, b, c, w, w_1 and w_2 (all lower case) are uncorrelated (measured) variables with random uncertainty components.
5. x, x_1, x_2, x_3, … , x_n are all uncorrelated (measured) variables with random uncertainty components.

in the various x's, by determining the *sensitivity* of the output to uncertainty in each of the inputs. A brief overview of the relevance and use of differential calculus in determining uncertainty components and examples of sensitivity are provided in Appendix A. Details with regard to the derivation and application of the general equation for the propagation of uncertainty are provided in Appendix B.

The expression for the propagation of uncertainties for uncorrelated variables described above (equation 1), is a special form of the general *law for the propagation of uncertainty*. This law, in its *general* form, can be applied when covariance or correlation exists between the various inputs. In this latter situation however, a more detailed statistical approach is required. Appendix B provides a summary describing the derivation of the general law for the propagation of uncertainty for both uncorrelated and correlated input variables.

Uncertainty Components When the Inputs are Correlated

When two or more of the input quantities are correlated, an alternative form of the general law of uncertainty propagation is required. Section 5.2 of the GUM describes the approach when correlated input quantities must be considered.

An example of the 'correlated input equation' describing two input quantities (x_1 and x_2) is shown as equation 3. This expression is similar to that given as equation 1, but includes an additional term which incorporates the correlation coefficient r. It should also be apparent that if there is zero correlation (that is, $r = 0$), the last term in equation 3 becomes zero and the expression reverts to the same form as equation 1.

$$u^2(y) = \left(\frac{\partial y}{\partial x_1}\right)^2 u^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 u^2(x_2) + 2r_{x_1,x_2}\left(\frac{\partial y}{\partial x_1}\right)\left(\frac{\partial y}{\partial x_2}\right) u(x_1)\, u(x_2)$$

(equation 3)

An example is provided in Appendix B which shows the importance of both recognizing and accounting for correlated input variables when they occur. This example describes the relationship $y = x_1\, x_2$, where the measurand is the product of the two input variables x_1 and x_2. Initially, these are considered as uncorrelated inputs, so equation 1 is appropriate for calculating the uncertainty propagated from the two x's through to the measurand y. Then, x_1 and x_2 are assumed to be perfectly positively correlated, with $r = +1$, where the use of equation 3 is now the correct approach. This latter result is then compared with the result for $y = x_1^2$, evaluated using equation 1. These two results are and should be identical, since $y = x_1^2$, evaluated using equation 1, is equivalent to $y = x_1\, x_2$ with perfect positive correlation (that is, when $x_1 = x_2$ there is perfect correlation, as any quantity is perfectly correlated with itself).

Summary of Procedure for Determining the Propagation of Uncertainty

To summarise the steps for determining the propagation of uncertainty, assume again that there is a measurand y, and that y is a function of n inputs $x_1, x_2, x_3, \ldots, x_n$. The numerical values of these n inputs are assumed known, and their standard uncertainties $u(x_1), u(x_2), u(x_3), \ldots, u(x_n)$ are also assumed known. Since there is a functional relationship $y = f(x_1, x_2, x_3, \ldots, x_n)$ between the input variables and the measurand, the numerical value of the measurand (y) can be calculated by straightforward algebra without any calculus. The question now is how to determine the resulting standard uncertainty $u(y)$ in y, into which the component standard uncertainties $u(x_1), u(x_2), u(x_3), \ldots, u(x_n)$ have all propagated. It is this latter calculation where some calculus is necessary and this has the following steps:

1. Differentiate y with respect to x_1, taking all the other inputs $x_2, x_3, \ldots, x_n$ as constants. In the language of the calculus, we take the *partial derivative* of y with respect to x_1. This partial derivative is denoted by $\partial y/\partial x_1$. Because this partial derivative itself is a function of at least some of $x_1, x_2, x_3, \ldots, x_n$, and since these inputs all have known numerical values derived from the experimental data, the partial derivative itself has a known numerical value. As outlined in Appendix A, this partial derivative is a measure of the sensitivity of y to any changes in x_1 and can accordingly be called a *sensitivity coefficient*.
2. Take the square of $\partial y/\partial x_1$, obtaining $(\partial y/\partial x_1)^2$. Again, this expression has a known numerical value derived directly from the experimental data.
3. Multiply the squared partial derivative in step (2) by the squared standard uncertainty $u(x_1)$ *of* x_1, obtaining the product $(\partial y/\partial x_1)^2\, u^2(x_1)$.
4. Repeat steps (1) to (3) for input x_2, then for input x_3, and so on, up to input x_n.
5. The resulting n products $(\partial y/\partial x_1)^2\, u^2(x_1)$, $(\partial y/\partial x_2)^2\, u^2(x_2)$, $(\partial y/\partial x_3)^2\, u^2(x_3)$, … , $(\partial y/\partial x_n)^2\, u^2(x_n)$, are all summed. This sum is the squared standard uncertainty $u^2(y)$ in the measurand y and is identical to equation 1.
6. Equation 1 applies to the case of uncorrelated inputs as described above and in Appendix B. If some or all of the inputs are mutually correlated, equation 1 has additional terms involving correlation coefficients as shown in equation 3 and discussed in Appendix B.
7. Depending on the exact form of the functional relationship $y = f(x_1, x_2, x_3, \ldots, x_n)$, it may be more appropriate to use the relative uncertainties (that is, $u(y)/y$ and

$u(x_i)/x_i$) instead of the absolute uncertainty values ($u(y)$ and $u(x_i)$). As a general rule, use relative uncertainties for functions which include terms with multiplication and/or division but not addition or subtraction. Examples of expressions that involve relative uncertainties are given in Table 2. Relative uncertainty is also referred to as *coefficient of variation* (CV).

Precision Profile and Uncertainty of Measurement

The precision profile of an assay is a convenient way to describe the relationship between the concentration of a substance and its measured precision. It is usually presented as a plot of standard deviation or coefficient of variation against measurand value.[14]

In clinical laboratories, routine quality control is often evaluated at only two or three measurand values. However, uncertainty of measurement may vary over the analytical range of the method and outside of the range covered by the control material. A review of an assay's precision profile is a good way of determining the change in precision at values not directly covered by the routine control material. When deriving uncertainty of measurement values or using precision estimates in calculations, it is important to consider the use of the actual precision at the measurand value in question.

Rules for Calculating Uncertainty through Functional Relationships

Table 2 provides a selection of 'rules' for calculating uncertainty based on functional relationships which contain uncorrelated variables. They have been presented in this format to give specific examples which can be directly applied to many clinical laboratory calculations, without the requirement to be fully conversant with the GUM or the differential calculus.

In other scientific applications, many more types of calculation may be required. For example, calculations based on formulae which include trigonometric functions and/or more complicated logarithmic relationships would be common in the physical sciences. However, uncertainty associated with any measurement and its propagation through a defined functional relationship can be evaluated by differentiation (partial differentiation) and the application of the general equation for the propagation of uncertainty. The formulae for calculating the combined standard uncertainty provided in Table 2 have been derived using partial differentiation of the corresponding functional relationship as outlined in the previous section. Some of the simpler and more common functions which could well have been provided in a more general mathematical form, have been presented as separate expressions in order that visual comparison of the various rules may be more easily achieved. For example, rule 1 (for $y = x_1 + x_2$) and rule 2 (for $y = x_1 - x_2$) are essentially the same; they are specific cases of rule 3 with the uncertainty components being calculated in the same manner for all three rules.

Calculations involving other functional relationships and analysis of variance procedures may be found in Annex H of the GUM.

Examples: Application of the Rules for Calculating Uncertainty through Functional Relationships

There are many examples in laboratory medicine where the measurand is calculated from other measurements using a functional relationship. Table 3 provides examples of laboratory calculations and the rules to be used for evaluating uncertainty in their respective output values. These particular examples have been chosen to show the variety and form of the functional relationships which may be encountered and the rules which should be applied in order to calculate the output uncertainty in the measurand. In the examples provided in Table 3, it has been assumed that the variables in each of the equations are uncorrelated. As outlined in more detail below, this assumption may not actually be correct but the relevant variables are usually considered to be uncorrelated for most practical applications.

In some situations, uncertainty in one or more of the input variables may be considered small when compared with uncertainty in the other input variable(s). If this can indeed be shown, a particular term may be disregarded and a simpler form of the function may be applicable. However, any assessment of this type and the reason why some types of uncertainty may be excluded from a calculation should be recorded in the appropriate laboratory standard operating procedure in a manner consistent with ISO 15189 / AS 4633 (for example, ISO 15189 item 5.5.3). Several of the functions presented in Table 3 provide examples of both of these approaches. That is, it is appropriate to record both a formal assessment plus a simplified version where the uncertainty in one term may be considered to contribute a relatively small proportion to the combined standard uncertainty. The following comments with regard to some of the relationships outlined in Table 3 can be considered as general comments, which may be applied to all calculations involving the estimation of uncertainty using functional relationships.

Table 3, items 1, 2 and 3

These are common equations for calculating serum anion gap, serum osmolality and corrected serum calcium.[15-20] All three equations contain variables which are assumed to be uncorrelated. This is typical of their use within a clinical laboratory where the output values are calculated assuming

Table 3. Examples of laboratory calculations and the rules to be used for evaluating uncertainty in their respective output values as described in Table 2. As discussed in the text, the variables within each of the equations are assumed to be uncorrelated.

Item	Function	Example equation	Measurand	Rule
1	$y = x_1 + x_2 - x_3 - x_4$	$AG = Na^+ + K^+ - Cl^- - HCO_3^-$	AG; anion gap, unit mmol/L.	3
2	$y = Ax_1 + x_2 + x_3 + B$ where A and B are without uncertainty	C Osmol = 1.86 (Na^+) + urea + glucose + 9	C Osmol; calculated serum osmolality, unit mmol/Kg. [17,18]	3
3	$y = x_1 - A(B - x_2) = x_1 - AB + Ax_2$ where A and B are without uncertainty	C Ca^{++} = Ca^{++} − 0.02 (40 − S albumin)	C Ca^{++}; corrected serum calcium concentration, unit mmol/L. [19,20]	3
4	$y = (x_1 \times x_2) / (x_3 \times x_4)$	CrCl = (U × V) / (S × T)	CrCl; creatinine clearance, unit ml/min or ml/sec (depending on the units for T).	8
5	$y = (x_1 \times x_2) / (x_3 \times x_4)$	FEs = (Us × SCr) / (UCr × Ss)	FEs; fractional excretion of a substance, dimensionless quantity.[34]	8
6	$y = \log_{10} x$ and $y = \log_{10} x^A = A \log_{10} x$ where A is without uncertainty	$pH = -\log_{10} [H^+] = \log_{10} [H^+]^{-1}$	pH definition; dimensionless quantity.	15
7	$y = A + \log_{10} (x_1/x_2) = A + \log_{10} x_1 - \log_{10} x_2$ where A is without uncertainty	pH = pKa + $\log_{10}$ (base / acid)	pH; Henderson Hasselbalch equation assuming pKa has no uncertainty.	17
8	$y = x_1 + \log_{10} (x_2/x_3) = x_1 + \log_{10} x_2 - \log_{10} x_3$	pH = pKa + $\log_{10}$ (base / acid)	pH; Henderson Hasselbalch equation with an uncertainty estimate for pKa.	1,17
9	$y = (\ln x_1 - \ln x_2) / (x_3 - x_4)$	$K_d = (\ln C_1 - \ln C_2) / (T_2 - T_1)$	K_d = drug elimination rate constant, unit $time^{-1}$.[35]	23
10	$y = (x_1 / x_2)^A$ where A is without uncertainty	$INR = (P / N)^{ISI}$	INR; international normalised ratio, dimensionless quantity. ISI; international sensitivity index.	10
11	$y = a(b\,x_1)^{D1} \times (c\,x_2)^{D2}$ where $D1$ and $D2$ are without uncertainty	eGFR = 175 $(SCr \times 0.0113)^{-1.154}$ $(age)^{-0.203}$	eGFR; estimated GFR using MDRD/IDMS formula for males, units mL/min/1.73m^2.[30-33]	24
12	$y = a(b\,x_1)^{w1} \times (c\,x_2)^{w2}$ where w_1 and w_2 contribute uncertainty	eGFR = 175 $(SCr \times 0.0113)^{-1.154}$ $(age)^{-0.203}$	eGFR; estimated GFR using MDRD/IDMS formula for males, units mL/min/1.73m^2.[30-33]	25
13	$y = (x_1 / x_2)^w$ where w contributes uncertainty	$INR = (P / N)^{ISI}$	INR; international normalised ratio, dimensionless quantity. ISI; international sensitivity index.	20

the input measurands are uncorrelated. This assumption is not strictly correct however, as some variables which may be treated as uncorrelated for practical purposes may actually show a degree of correlation either physiologically or analytically. For example:

- In many physiological situations, a decreased serum sodium activity is associated with a decreased serum chloride activity and vice versa.
- In metabolic alkalosis, an increased serum bicarbonate concentration is usually associated with a decreased serum chloride concentration (activity).
- As serum calcium is bound to serum albumin, these two measurands can indeed be shown to have a positive correlation. To add further complication, the actual correlation of substances such as serum calcium and serum albumin is probably both patient and method-dependent. This is also exemplified by the numerous equations which have been proposed for calculating corrected serum calcium (that is, serum calcium adjusted for the actual patient's serum albumin concentration). A detailed discussion regarding the derivation and applicability of such equations is outside the scope of this review.[20-24]
- Analytical interference by bicarbonate in some serum chloride activity assays has been described as showing 'a clear linear increase in chloride with increasing bicarbonate …'.[25] Analytical interference of this nature can therefore be considered as showing a positive correlation between the particular measurands. In a similar manner, an increased serum sodium activity due to glucose interference (by a relatively high glucose concentration) has been reported.[26] These examples show that an observed correlation between two quantities that should be uncorrelated is often indicative of a potential defect in the particular analytical system.

When correlation is actually taken into account, an additional term must be included as described previously and in Appendix B. Taking the anion gap equation as a further example, if all correlations can be ignored the relevant formula for $u^2(y) = u^2(\mathrm{AG})$, the squared standard uncertainty of the anion gap, is given in Table 2, rule 3. If a correlation term is required to represent the analytical interference of bicarbonate in the chloride assay for example, the squared standard uncertainty of the anion gap now becomes:
$u^2(\mathrm{AG}) = u^2(\mathrm{Na^+}) + u^2(\mathrm{K^+}) + u^2(\mathrm{Cl^-}) + u^2(\mathrm{HCO_3^-}) + 2r(\mathrm{Cl^-}, \mathrm{HCO_3^-})\ u(\mathrm{Cl^-})\ u(\mathrm{HCO_3^-})$, where the term $r(\mathrm{Cl^-}, \mathrm{HCO_3^-})$ represents the correlation coefficient between chloride and bicarbonate for the analytical interference of bicarbonate in the chloride assay.

In addition, items 2 and 3 contain 'numerical constants' which are assumed (in the functional forms provided in Table 3) to be free of inherent uncertainty. Both of these equations however, have been derived by regression analysis using multiple paired measurements as the input variables. The 'numerical constants' are essentially coefficients derived by the regression analysis. Most comparisons of paired data, such as the comparison of measured versus calculated serum osmolality, will incorporate a degree of variability as shown by the scattering of data points in a '*y*' versus '*x*' data plot (scatter plot). Unless the standard uncertainty in the original regression line (curve) and any associated regression coefficients are also provided, the full impact on the output from the functional relationship which is now applied to new input variables cannot be assessed. That is, the 'numerical constants' shown in items 2 and 3 are not truly constant. They contribute an unknown degree of uncertainty which is not included in the expressions as stated in Table 3 or as usually applied within the clinical laboratory.

Even accepting that the equations for serum anion gap, calculated serum osmolality and corrected serum calcium may lack strict mathematical correctness, they do provide relevant examples of common laboratory calculations. As the output values from such calculations are often considered to provide only approximate clinical information, the associated uncertainty estimates which assume uncorrelated variables will also provide only an approximate range of possible values.

Table 3, items 7 and 8

An argument similar to that outlined above may well apply to the functional relationship given for the Henderson Hasselbalch equation. Item 7 provides an equation where it is assumed that the acidity constant (Ka) has no uncertainty. In practice, this may well have a very low uncertainty value which does not contribute to the uncertainty in the output measurand (pH) as determined by measurement of the concentrations of the two input variables 'base' and 'acid'. By contrast, item 8 provides a more general form in which no assumption has been made as to the degree of uncertainty in the Ka (or pKa).

Table 3, items 10 and 13

The universally utilised function which defines the International Normalised Ratio (INR) is a good example of an equation which includes a power term. The international sensitivity index (ISI), the power term in the INR function, provides an adjustment to align a given thromboplastin reagent and analytical device combination with a World Health Organization (WHO) international thromboplastin standard. If the given thromboplastin reagent is identical to the WHO standard, the ISI equals 1.0. However, depending

on the particular reagent preparation and the measurement equipment employed, ISI values can range from 0.9 to approximately 1.8.[27,28]

There are several possible ways to evaluate uncertainty in the INR functional relationship. One method could be to rely on multiple quality control results obtained over many operational situations. For this method of assessment, the standard deviation of internal quality control results which include multiple values from several reagent batches (with different ISI values which have been appropriately applied for each particular batch of reagent), may well be considered as providing a direct measure of INR uncertainty.

A second method would be to calculate the propagation of uncertainty through the INR function using uncertainties in the three input variables (measured patient prothrombin time (P), average normal prothrombin time (N), and ISI) as described in Table 3. In this method, the uncertainty inherent in the ISI should be obtained from the reagent certification documentation provided by the reagent manufacturer, or derived 'in-house' using appropriate regression statistics which include an assessment of the standard uncertainty of the regression plus the uncertainty inherent in the derived regression coefficients.[27,29]

Table 3, items 11 and 12

Items 11 and 12 are the Modification of Diet in Renal Disease (MDRD) formula for estimating glomerular filtration rate (eGFR) in male subjects, with serum creatinine measurement standardised against isotope dilution mass spectrometry (IDMS).[30] A number of similar formulae for the calculation of eGFR have been proposed and recently reviewed.[31,32] Specific differences have been observed for males, females and various racial groups. As these differences all add multiplier terms to the basic form of the equation, the approach used to evaluate uncertainty in the output variable (eGFR) is similar for all forms of this equation.

The equations shown in items 11 and 12 also contain 'numerical constants' as both multipliers and power terms. Presented in this manner, it may well be assumed that these 'numerical constants' are free of uncertainty even though the original reports clearly indicate this assumption cannot be correct. Like many empirical equations which have useful applications in medicine, the eGFR equations have been derived by stepwise multiple regression procedures to determine a set of variables that jointly predict eGFR.[30,33] As the regression coefficients have been derived from data with a wide inherent variability, the actual uncertainty associated with each of the numerical terms is also required in order to provide a more complete uncertainty estimate of the output variable (eGFR). Item 12 provides the rule for estimating eGFR uncertainty when all measurements, including the empirically derived coefficients and power values have associated uncertainty estimates included. As the factor 0.0113 in the serum creatinine term provides a conversion of units from μmol/L, its value is known with sufficient precision to be treated as a true numerical constant.

Item 11 provides a simplified version of item 12, where the empirically derived coefficient and power values are 'incorrectly' assumed to be uncertainty free.

Summary

Accreditation of Australian medical testing laboratories under ISO 15189 (AS 4633) has been required since July 2005.[1] As part of this requirement, the uncertainty of quantitative results must also be available to demonstrate how well a measured value actually represents the quantity being determined. The GUM is generally accepted 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 general *law for the propagation of uncertainty*. The manner in which uncertainties are propagated from measured values to a calculated quantity through a functional relationship provides much of the mathematical challenge to fully understanding the GUM.

This review provides a summary of the major features of the GUM with specific application to medical testing (clinical) laboratories. The application of the GUM formula to various types of functional relationship provides a rule-based approach for the calculation of uncertainty estimates. A selection of the principal 'rules' and their routine application is provided, in order that other examples with similar mathematical forms may be evaluated by direct comparison.

Even though the aim of this article is to provide an overview of the GUM, there are many learned texts which provide appropriate information on the theory, derivation and application of the general law for the propagation of uncertainty. Many assume a good grasp of mathematics and statistical theory, but less complex versions are available and thus can provide a good starting point for those wishing to investigate this all-important topic in more detail.[38,39,40]

Competing Interests: None declared.

References

1. White GH, Farrance I. Uncertainty of measurement in quantitative medical testing - A laboratory implementation guide. Clin Biochem Rev 2004;25:

S1-24.
2. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 1997;34: 8-12.
3. Fraser CG, Hyltoft Petersen P. Analytical performance characteristics should be judged against objective quality specifications. Clin Chem 1999;45:321-3.
4. Fraser CG. General strategies to set quality specifications for reliability performance characteristics. Scand J Clin Lab Invest 1999;59:487-90.
5. Klee GG. Establishment of outcome-related analytic performance goals. Clin Chem 2010;56:714-22.
6. Fraser CG. Quality specifications. In: Biological Variation: From Principles to Practice. Washington DC: AACC Press; 2001. p. 29-66.
7. White GH. Metrological traceability in clinical biochemistry. Ann Clin Biochem 2011;48:393-409.
8. Joint Committee for Guides in Metrology. International vocabulary of metrology - Basic and general concepts and associated terms (VIM). JCGM 200:2012. http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf (Accessed 8 March 2012).
9. Joint Committee for Guides in Metrology. Evaluation of measurement data - Guide to the expression of uncertainty in measurement (GUM). JCGM 100:2008. http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf (Accessed 8 March 2012).
10. Bureau International des Poids et Mesures. Joint Committee for Guides in Metrology (JCGM). http://www.bipm.org/en/publications/guides/gum.html (Accessed 8 March 2012).
11. Joint Committee for Guides in Metrology. Evaluation of measurement data – An introduction to the "Guide to the expression of uncertainty in measurement" and related documents. JCGM 104:2009. http://www.bipm.org/utils/common/documents/jcgm/JCGM_104_2009_E.pdf (Accessed 8 March 2012).
12. Siekmann L. Requirements for reference (calibration) laboratories in laboratory medicine. Clin Biochem Rev 2007;28:149-54.
13. White GH. Basics of estimating measurement uncertainty. Clin Biochem Rev 2008;29(Suppl 1):S53-60.
14. Linnet K, Boyd JC. Selection and analytical evaluation of methods – with statistical techniques. In: Burtis CA, Ashwood ER, Bruns DE, editors. Tietz Fundamentals of Clinical Chemistry. Sixth edition. St Louis, MO: Elsevier Saunders; 2008. p. 208.
15. Lolekha PH, Lolekha S. Value of the anion gap in clinical diagnosis and laboratory evaluation. Clin Chem 1983;29:279-83.
16. Kraut JA, Madias NE. Serum anion gap: its uses and limitations in clinical medicine. Clin J Am Soc Nephrol 2007;2:162-74.
17. Dorwart WV, Chalmers L. Comparison of methods for calculating serum osmolality form chemical concentrations, and the prognostic value of such calculations. Clin Chem 1975;21:190-4.
18. Bhagat CI, Garcia-Webb P, Fletcher E, Beilby JP. Calculated vs measured plasma osmolalities revisited. Clin Chem 1984;30:1703-5.
19. Endres DB, Rude RK. Disorders of bone. In: Burtis CA, Ashwood ER, Bruns DE, editors. Tietz Fundamentals of Clinical Chemistry. Sixth edition. St Louis, MO: Elsevier Saunders; 2008. p. 711-34.
20. Barth JH, Fiddy JB, Payne RB. Adjustment of serum total calcium for albumin concentration: effects of non-linearity and of regression differences between laboratories. Ann Clin Biochem 1996;33:55-8.
21. Pain RW, Rowland KM, Phillips PJ, Duncan BM. Current "corrected" calcium concept challenged. Br Med J 1975;4:617-9.
22. Ramsay LE, Shelton JR. "Corrected" calcium concept. Br Med J 1976;1:219.
23. Payne RB, Carver ME, Morgan DB. Interpretation of serum total calcium: effects of adjustment for albumin concentration on frequency of abnormal values and on detection of change in the individual. J Clin Pathol 1979;32:56-60.
24. Walker BE, Payne RB. Adjusted calcium conflict resolved? Differing effects on plasma total calcium of changes in plasma albumin after venous stasis or myocardial infarction. J Clin Pathol 1979;32:488-91.
25. Dimeski G, Clague AE. Bicarbonate interference with chloride-ion-selective electrodes. Clin Chem 2004;50:1106-7.
26. Al-Musheifri A, Jones GRD. Glucose interference in direct ISE sodium measurements. Clin Biochem Rev 2007;28(Suppl):S18.
27. Favaloro EJ, Hamdam S, McDonald J, McVicker W, Ule V. Time to think outside the box? Prothrombin time, international normalised ratio, international sensitivity index, mean normal prothrombin time and measurement of uncertainty: a novel approach to standardisation. Pathology 2008;40:277-87.
28. Harris NS, Winter WE. The international normalized ratio. Clin Lab News 2010;36(11) http://www.aacc.org/publications/cln/2010/november/Pages/TheInternationalNormalizedRatio.aspx (Accessed 8 March 2012).
29. Poller L, Ibrahim S, Keown M, Pattison A, Jespersen J. Simplified method for international normalized ratio (INR) derivation based on the prothrombin time/INR line: an international study. Clin Chem 2010;56: 1608-17.
30. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al; Chronic Kidney Disease Epidemiology Collaboration. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med 2006;145:247-54.
31. Florkowski CM, Chew-Harris JS. Methods of estimating

GFR – different equations including CKD-EPI. Clin Biochem Rev 2011;32:75-9.
32. Jones GR. Estimating renal function for drug dosing decisions. Clin Biochem Rev 2011;32:81-8.
33. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D; Modification of Diet in Renal Disease Study Group. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Ann Intern Med 1999;130:461-70.
34. Deacon A. Renal function. In: Calculations in Laboratory Science. London: ACB Venture Publications; 2009. p. 94.
35. Deacon A. Basic pharmacokinetics. In: Calculations in laboratory science. London: ACB Venture Publications; 2009. p. 119.
36. Kirkup L, Frenkel B. Correlated inputs. In: An Introduction to Uncertainty in Measurement. Cambridge: Cambridge University Press; 2006. p. 109.
37. Zhang NF. Calculation of the uncertainty of the mean of autocorrelated measurements. Metrologia 2006;43:S276-81.
38. Kirkup L, Frenkel B. An introduction to uncertainty in measurement. Cambridge: Cambridge University Press; 2006.
39. Taylor JR. An Introduction to Error Analysis, the study of uncertainties in physical measurements. Sausalito, California: University Science Books; second edition, 1997.
40. European co-operation for accreditation. EA4/02:1999. Expression of the uncertainty of measurement in calibration. http://www.european-accreditation.org/n1/doc/EA-4-02.pdf (accessed 08 March 2012).

Appendix

Appendix A. Differential Calculus and Uncertainty of Measurement

Overview

Differential calculus provides a mathematical description of the ways in which related quantities change. This description relies on the concept that a *small change* in a quantity such as x, symbolised by δx, corresponds to a small change in the output variable y, symbolized by δy. The Greek letter delta (δ) is used to represent this change and does *not* mean δ multiplied by x or δ multiplied y, but is mathematical shorthand for 'a small change in the specified quantity'. For example, if $x = 10$, then δx might be 0.1; that is, δx represents a 1% change in x and this change may be either positive or negative. This concept of a small change also underlies its relevance to the analysis of uncertainties.

Uncertainties in measurement are created by errors, where an error as previously defined may also be described as a *small departure* in a quantity from its true value. An error is the discrepancy between a measured value and the actual or true value, while uncertainty is the effect of several errors or a very large ensemble of errors acting conjointly. As it may only be possible to specifically identify and numerically estimate a few of these errors, overall statistical variability can be used to provide an estimate of the range into which a measured value is expected to fall.

Higher Order Terms in Differential Equations

An important theme that runs throughout the differential calculus is that so called *higher-order* terms are regarded as negligible, in a manner to be described more fully in the next section. Such higher-order terms are the squares, cubes and higher-order powers of $\delta x/x$ such as $(\delta x/x)^2$. Consider again the example where $x = 10$ and $\delta x = 0.1$. The ratio $\delta x/x$ has the relatively small value of 0.01 or 1% . Higher-order terms such as $(\delta x/x)^2$ or $(\delta x/x)^3$ have even smaller relative values. Continuing with the example where $x = 10$ and $\delta x = 0.1$, $(\delta x/x)^2 = 0.0001$ or 0.01% and $(\delta x/x)^3 = 0.000001$ or one part in a million. Thus, when considering $\delta x/x$ as an indicator of relative change, higher-order terms generally contribute negligible additional information.

Suppose we have a *linear relationship* described by $y = 7x + 3$. What defines this type of equation as a straight line is the absence of squares or other powers in either x or y, and the absence of functional relationships such as square root, sine, logarithm, etc. A small change δx in x, will then take the value of x to a new value of $x + \delta x$ and create a small change δy in y, taking y to the new value of $y + \delta y$. The new functional relationship will be $y + \delta y = 7(x + \delta x) + 3$. However, as $y = 7x + 3$, we can cancel it from both sides of the equation to give $\delta y = 7(\delta x)$. Because of the linear relationship, there are no higher-order terms such as $(\delta x/x)^2$. This last equation ($\delta y = 7(\delta x)$), may also be written as $\delta y/\delta x = 7$.

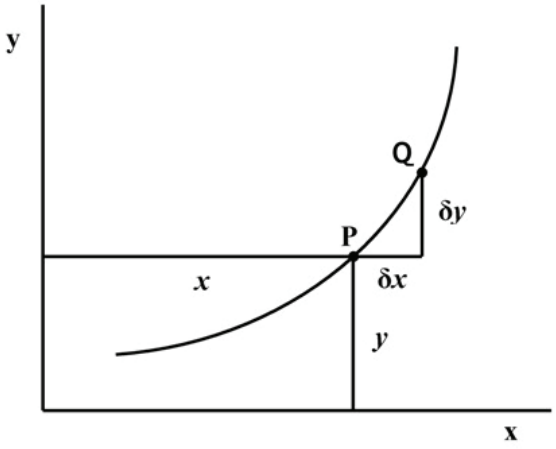
On the other hand, suppose that the relationship was $y = 7x^2 + 3$. This is a non-linear relationship as it contains an x^2 term. Then, if x changes to $x + \delta x$, y will change to $y + \delta y = 7(x + \delta x)^2 + 3 = 7(x^2 + 2x\delta x + (\delta x)^2) + 3$. Cancelling $y = 7x^2 + 3$ from both sides of the equation in a similar manner to that used above, leaves $\delta y = 14x\delta x + 7(\delta x)^2$. Now, dividing both sides of the equation by x^2 gives $\delta y/x^2 = 14(\delta x/x) + 7(\delta x/x)^2$. Using the fact that $(\delta x/x)^2$ is much smaller than $(\delta x/x)$ and can be ignored, gives $\delta y/x^2 = 14(\delta x/x)$. Multiplying by x^2 leaves $\delta y = 14x\delta x$, which may also be written as $\delta y/\delta x = 14x$.

In both the linear and non-linear case, the final relationship between δy and δx has been written as the ratio $\delta y/\delta x$. The reason for this will be discussed in more detail below.

Differentiation

Figure 1(a) shows a curve of output y against a single input x. The point P on the curve represents input x and the corresponding output (measurand value) y. The point Q on the curve is near P, and at Q the corresponding values of input and output are $x + \delta x$ and $y + \delta y$ respectively. We may imagine Q as approaching P more and more closely (but still remaining as a point on the curve), so that δx decreases and so of course does the corresponding δy. Thus, *in the limit* (a popular expression in the differential calculus), when Q is only an infinitesimally small distance from P (but still remaining as a point on the curve), the higher powers (squared, cubed, etc) of δx and δy become even *more and more justifiably negligible*. Then, as shown in Figure 1(b), the *ratio* $\delta y/\delta x$, effectively becomes the tangent to the curve at the point P with a slope of α (alpha). The universally accepted symbol for the ratio $\delta y/\delta x$ in this limiting case is dy/dx. The differential calculus provides the mathematical procedures for calculating dy/dx. Such a calculation is called 'differentiating y with respect to x', or 'taking the derivative of y with respect to x'.

1a



1b

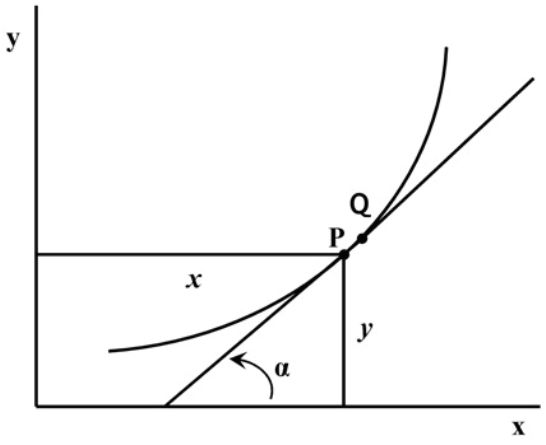


Figure 1(a). Points P and Q on the curve $y = f(x)$. Coordinates of P = (x, y), coordinates of Q = $(x+\delta x, y+\delta y)$; where δx represents a small change in x and δy the corresponding small change in y. **Figure 1(b).** As Q approaches P, the distance P to Q along the curve approximates a straight line. When Q is infinitesimally close to P, the tangent line to P and Q makes an angle α (alpha) to the x-axis, where tan α equals the derivative dy/dx. This is the derivative of y with respect to x.

The actual result obtained for dy/dx will naturally depend on the particular functional form which describes the dependence of y on x. Continuing with the example of a single input x, the functional form which relates output y to input x is $y = f(x)$. For example, if the curve in Figures 1(a) and 1(b) is the curve $y = x^2$, with the input x and output y located at P, then at Q we have:

$$y + \delta y = (x + \delta x)^2 = x^2 + 2x\delta x + (\delta x)^2 \qquad \text{(A 1)}$$

If we replace y on the left hand side of A1 with its equivalent x^2, the x^2 terms cancel and we are left with:

$$\delta y = 2x\delta x + (\delta x)^2, \text{ or } \delta y / \delta x = 2x + \delta x \qquad \text{(A 2)}$$

As Q moves closer to P, δx and δy both approach zero but the ratio $\delta y/\delta x$ remains finite (that is, one small number divided by another small number remains a finite or non-zero number). On the other hand, δx on the right hand side of A2 will eventually become zero, thus:

$$dy / dx = 2x \qquad \text{(A 3)}$$

which is now the derivative of y with respect to x.

An alternative approach for obtaining the derivative described in A3 relies on the arguments outlined above for setting higher order terms of the ratio $\delta x/x$ to zero. Starting again with A1 and cancelling $y = x^2$ from both sides of the equation leaves:

$$\delta y = 2x\delta x + (\delta x)^2 \qquad \text{(A 4)}$$

and dividing both sides by x^2 gives:

$$\frac{\delta y}{x^2} = 2\frac{\delta x}{x} + \left(\frac{\delta x}{x}\right)^2 \qquad \text{(A 5)}$$

As discussed previously, the second term on the right hand side of A6 is negligible relative to the first term on the right hand side, so removing the second term (that is, $(\delta x/x)^2$) and rearranging the remaining terms, A3 can again be obtained.

In the simplest case of a *linear relationship* between y and the single input x (that is, $y = x$), $y + \delta y = x + \delta x$, implying

immediately that $\delta y = \delta x$, or in the limit, $dy/dx = 1$. In a similar manner to the derivation of A3, it can be shown that if $y = x^3$, then $dy/dx = 3x^2$; if $y = x^4$ then $dy/dx = 4x^3$. In general, if $y = Kx^m$, where m is any number and not necessarily an integer and K is a constant, then $dy/dx = Kmx^{m-1}$. If $m = -1$, so that $y = K/x$, then $dy/dx = -K(1/x^2)$.

Examples of Differentiation and 'Sensitivity'

Suppose that the measurand y is the *circumference of a circle* with radius r. Here the symbol x is replaced by r and:

$$y = 2\pi r \qquad \text{(A 6)}$$

Thus $dy/dr = 2\pi$ (in this case, $K = 2\pi$ and $m = 1$), and from A6 it also follows that:

$$\delta y = 2\pi \delta r \qquad \text{(A 7)}$$

Combining A7 and A6 as a ratio (that is, dividing both sides of A7 by y), it follows that:

$$\frac{\delta y}{y} = \frac{2\pi \delta r}{2\pi r} = \frac{\delta r}{r} \qquad \text{(A 8)}$$

From A8 it can now be seen that a 1% change in r (for example) also creates a 1% change in y. Still keeping to the case of a single input, suppose that the measurand y is now the *area of a circle* with radius r. The input again is r and the area is:

$$y = \pi r^2 \qquad \text{(A 9)}$$

In this example $K = \pi$, $m = 2$ and $dy/dr = 2\pi r$, giving:

$$\delta y = 2\pi r \delta r \qquad \text{(A 10)}$$

Combining A10 and A9 in a similar manner to that described previously:

$$\frac{\delta y}{y} = \frac{2\pi r \delta r}{\pi r^2} = \frac{2\delta r}{r} \qquad \text{(A 11)}$$

In this example, a 1% change in the radius now creates a 2% change in the measurand y, the area of the circle. This example also shows a more *sensitive dependence* of y on the input as compared to the previous example of the circumference of a circle. The factor 2 in A11 is derived from the squared term (r^2) in A9.

As a final example of the one input case, if we take y as the volume of a sphere with radius r, then:

$$y = (4/3)\pi r^3 \qquad \text{(A 12)}$$

In this case $K = (4/3)\pi$, $m = 3$ and $dy/dr = 4\pi r^2$, giving:

$$\delta y = 4\pi r^2 \delta r \qquad \text{(A 13)}$$

As before, combining A13 and A12 gives:

$$\frac{\delta y}{y} = \frac{4\pi r^2 \delta r}{(4/3)\pi r^3} = \frac{3\delta r}{r} \qquad \text{(A 14)}$$

As might be expected, this example now shows a highly sensitive dependence of the volume on the radius. A 1% change in the radius creates a 3% change in the volume. Summarising these three examples, we see that the increasing *sensitivity* of the measurand to the radius is explained by the successive increase in the power which defines the dependence of the measurand on the radius: linear for the circumference, quadratic for the area and cubic for the volume.

When there are several terms involving the same input in the expression for the measurand y, dy/dx is found by differentiating each term separately. For example, if $y = Ax + Bx^3$ where A and B are constants, then $dy/dx = A + 3Bx^2$.

When there is more than one input, the ordinary derivatives become *partial derivatives*. We now differentiate with respect to each input $x_1, x_2, \ldots, x_n$ in turn, where all but the variable of interest is held fixed during the differentiation and temporarily regarded as constant. The symbol for a partial derivative is ∂. For example, if:

$$y = C_1 x_1 - C_2 x_2^3 + C_3 x_1 x_2 \qquad \text{(A 15)}$$

where all the *C*s are constants, we have:

$$\frac{\partial y}{\partial x_1} = C_1 + C_3 x_2 \quad \text{and} \quad \frac{\partial y}{\partial x_2} = -3C_2 x_2^2 + C_3 x_1 \qquad \text{(A 16)}$$

Some Useful Derivatives of *y* with Respect to *x*

(a) As stated above, if $y = Kx^m$, where K is a constant and m can be positive, negative or zero and is not necessarily an integer, then $dy/dx = Kmx^{m-1}$.

(b) If $y = \ln x$, where ln is the natural logarithm (to the base e), then $dy/dx = 1/x$. If $y = \log_{10} x$, where $\log_{10}$ denotes log to the base 10, then since $\log_{10} x = (\ln x)(\log_{10} e)$, $dy/dx = (\log_{10} e)(1/x) = 0.4343\ (1/x)$.

(c) If y is given by the product $y = f(x) \cdot g(x)$, where $f(x)$ and $g(x)$ are themselves both functions of x, then $dy/dx = f(x)\ dg(x)/dx + g(x)\ df(x)/dx$.

For example, if $f(x) = x^2$ and $g(x) = x+x^3$, then $y = f(x) \cdot g(x) = x^2\ (x+x^3) = x^3 + x^5$.

Differentiating this expression using rule (a) above, gives $dy/dx = 3x^2 + 5x^4$.

On the other hand, using rule (c) we have:

$$\frac{dg(x)}{dx} = 1 + 3x^2, \text{ so } f(x)\frac{dg(x)}{dx} = x^2 + 3x^4$$

and

$$\frac{df(x)}{dx} = 2x, \text{ so } g(x)\frac{df(x)}{dx} = 2x^2 + 2x^4$$

Adding these two expressions gives the same result as obtained with direct differentiation.

(d) If y is given by the ratio $y = \dfrac{f(x)}{g(x)}$, then:

$$\frac{dy}{dx} = \frac{g(x)\dfrac{df(x)}{dx} - f(x)\dfrac{dg(x)}{dx}}{[g(x)]^2}$$

Appendix B. Determining Standard Uncertainty in the Measurand

Uncertainty in Measurements With and Without Correlation

Given the general expression for the measurand y in terms of input quantities $x_1, x_2, x_3, \ldots, x_n$ (that is, $y = f(x_1, x_2, x_3, \ldots, x_n)$), a corresponding expression which describes the combined standard uncertainty in y ($u(y)$) can be obtained by appropriately combining the standard uncertainties of the various input estimates (the various $u(x)$'s). The general equation for the propagation of uncertainties for uncorrelated inputs as outlined in the GUM and described previously (equation 1), is given by the expression:

$$u^2(y) = \left(\frac{\partial y}{\partial x_1}\right)^2 u^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 u^2(x_2) + \ldots + \left(\frac{\partial y}{\partial x_n}\right)^2 u^2(x_n) \qquad \text{(B1)}$$

Equation B1 in effect states that the squared standard uncertainty $u^2(y)$ of the measurand y is a weighted sum of the squared standard uncertainties $u^2(x_i)$ of the inputs x_i (where $i = 1, 2, 3, \ldots, n$), the weights being the squared *sensitivity coefficients* . The input variables (the x_i) can be completely different physical or chemical quantities and need not be measured in the same units.

Starting with the functional relationship $y = f(x_1, x_2, x_3, \ldots, x_n)$, a small change δx_1 in x_1, δx_2 in δx_2, … , δx_n in x_n, all propagate to produce a small change δy in y in the following manner:

$$\delta y = \frac{\partial y}{\partial x_1}\delta x_1 + \frac{\partial y}{\partial x_2}\delta x_2 + \ldots + \frac{\partial y}{\partial x_n}\delta x_n \qquad \text{(B2)}$$

Equation B2 is actually an approximation, since any higher order terms such as $(\delta x_1)^2$ have been set to zero as described in Appendix A. However, this approximate form is sufficiently accurate for most purposes. It is also important to note that the derivatives $\partial y/\partial x_1$, $\partial y/\partial x_2$, etc, are given numerical values when evaluated at the actual experimental values of x_1, x_2, etc, and so these derivatives are effectively constants.

Equation B2 may also be recognised as a generalisation of the single-input examples given in Appendix A. A further example using multiple inputs may help demonstrate that this equation is indeed plausible. Suppose that $y = 3x_1 + x_1 x_2$. Then, if x_1 changes by a small amount δx_1 and x_2 by a small

amount δx_2, the resulting change δy in y is given by:

$$y + \delta y = 3(x_1 + \delta x_1) + (x_1 + \delta x_1)(x_2 + \delta x_2) \quad \text{(B3)}$$

In a similar manner to that described in Appendix A, subtracting $y = 3x_1 + x_1 x_2$ from B3 leaves:

$$\delta y = 3\delta x_1 + x_1\delta x_2 + x_2\delta x_1 + \delta x_1\delta x_2 \quad \text{(B4)}$$

Rearranging B4 and ignoring the small second-order term ($\delta x_1\delta x_2$) at the end gives:

$$\delta y = (3 + x_2)\delta x_1 + x_1\delta x_2 \quad \text{(B5)}$$

Using some of the concepts outlined in Appendix A and differentiating the original equation ($y = 3x_1 + x_1 x_2$) gives:

$$\frac{\partial y}{\partial x_1} = 3 + x_2 \quad \text{and} \quad \frac{\partial y}{\partial x_2} = x_1 \quad \text{(B6)}$$

Substituting the partial derivatives from B6 into B5 then gives the more general expression:

$$\delta y = \frac{\partial y}{\partial x_1}\delta x_1 + \frac{\partial y}{\partial x_2}\delta x_2 \quad \text{(B7)}$$

The procedures used in obtaining B7 demonstrate the manner in which the general expression for the propagation of uncertainty given in B2 has been derived.

We now identify each small change δx_1, δx_2, … , δx_n, as *random errors* in the inputs x_1, x_2, … , x_n. This implies that repeated measurements of the particular input will give slightly different results. These random errors can be any one of a very large, even infinite, set of similar random errors in the particular input and may arise from inherent fluctuations in the measuring device, imperfect quality control or lack of uniformity in any of the associated procedures. If there is any bias or systematic error in the measurement of any input, it should already have been corrected as described previously. This process allows us to assume a *zero mean value* for all the random errors associated with each of the inputs (that is, the mean of a large set of random deviations is zero). Even if bias does exists or has not been adequately corrected, it will not affect the equation for the propagation of uncertainties.

Each input now has a set of associated random errors each with zero mean. Then, since B2 is linear in each of the small changes δy, δx_1, δx_2, ... , δx_n, we can replace the deviations (δy, δx_i) with their mean values (with the *mean value* of all the random errors associated with an input x_i (i = 1, 2, 3, … , n) being zero) and obtain the correct, but uninteresting result that,

$$0 = \partial y / \partial x_1 \times 0 + \partial y / \partial x_2 \times 0 + ... + \partial y / \partial x_n \times 0$$

$$or \quad 0 = 0 + 0 + ... + 0 \quad \text{(B8)}$$

If instead of averaging the absolute values of the deviations however, we use the common statistical technique of taking the *squared deviations* which converts both positive and negative numbers into quantities which are all positive, a more meaningful outcome is achieved. Squaring both sides of B2 we obtain:

$$(\delta y)^2 = \left(\frac{\partial y}{\partial x_1}\right)^2 (\delta x_1)^2 + \left(\frac{\partial y}{\partial x_2}\right)^2 (\delta x_2)^2 + ... + \left(\frac{\partial y}{\partial x_n}\right)^2 (\delta x_n)^2$$

$$+2\left(\frac{\partial y}{\partial x_1}\right)\left(\frac{\partial y}{\partial x_2}\right)\delta x_1\delta x_2 + 2\left(\frac{\partial y}{\partial x_1}\right)\left(\frac{\partial y}{\partial x_3}\right)\delta x_1\delta x_3 + ..$$

$$+ 2\left(\frac{\partial y}{\partial x_2}\right)\left(\frac{\partial y}{\partial x_3}\right)\delta x_2\delta x_3 ... \quad \text{(B9)}$$

Equation B9 shows the terms which have been directly squared, followed by the *cross-product* terms which are preceded by the factor 2. This expression is simply the generalisation of the algebraic identity $(A+B)^2 = A^2 + B^2 + 2AB$.

Considering for a moment some basic statistics, we recall that if a population has a variance σ^2, we can estimate this variance using a sample of size M which is drawn from the population. If the individual sample measurements are represented by z_i (that is, z_1, z_2, … , z_M) the sample mean is given by:

$$\bar{z} = \frac{1}{M}\sum_{i=1}^{M} z_i \quad \text{(B10)}$$

The *unbiased estimate* of the variance (σ^2) for the population of z's is given by the *variance* (s^2) as defined in B11:

$$\text{estimate of } \sigma^2 = s^2 = \frac{\sum_{i=1}^{M}(z_i - \bar{z})^2}{M-1} = \frac{(z_1 - \bar{z})^2 + (z_2 - \bar{z})^2 + \ldots + (z_M - \bar{z})^2}{M-1} \quad \text{(B11)}$$

Equation B11 is a general statistical result for estimating the unbiased variance of a population by using sample statistics. B11 shows a summation over M terms, but then a division by $M-1$. It is this division by $M-1$ (the *degrees of freedom* associated with the sample mean) rather than by M, that gives the unbiased estimate of the population variance. B11 is applicable for both small and large M.

If we now identify the z's as the random errors δx in any of the inputs $x_1, x_2, \ldots, x_n$ (with a different collection of z's for each input), each with a mean of zero ($\bar{z} = 0$) as outlined above, then M is the number of random errors in each of the inputs and B11 becomes:

$$\text{estimate of } \sigma^2 = s^2 = \frac{\sum_{i=1}^{M} z_i^2}{M-1} = \frac{z_1^2 + z_2^2 + \ldots + z_M^2}{M-1} \quad \text{(B12)}$$

The operation described in B12 of summing over M and then dividing the sum by $M-1$, is now applied to each term on the right side of B9 and to the single term on the left side of B9. We assume for the moment that the mean of each term with a cross-product such as $\delta x_1 \delta x_2$, $\delta x_1 \delta x_3$, etc, is also zero (as this amounts to asserting that the various inputs are mutually uncorrelated; see below). Equation B9 now becomes:

$$\text{estimate of population variance of } y = \left(\frac{\partial y}{\partial x_1}\right)^2 \times \text{estimate of population variance of } x_1 + \left(\frac{\partial y}{\partial x_2}\right)^2 \times \text{estimate of population variance of } x_2 + \ldots + \left(\frac{\partial y}{\partial x_n}\right)^2 \times \text{estimate of population variance of } x_n \quad \text{(B13)}$$

Another term for the 'estimate of population variance' which is associated with an uncertainty is u^2, the 'squared standard uncertainty'. Equation B13 now becomes the basic statement for the *law for the propagation of uncertainty* with independent (uncorrelated) inputs $x_1, x_2, \ldots, x_n$, and output (measurand) y:

$$u^2(y) = \left(\frac{\partial y}{\partial x_1}\right)^2 u^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 u^2(x_2) + \ldots + \left(\frac{\partial y}{\partial x_n}\right)^2 u^2(x_n) \quad \text{(B14)}$$

Strictly speaking, the general equation for the propagation of uncertainty (B9) has been derived for population variances expressed as squared standard uncertainties $u^2(y)$, $u^2(x_i)$, with $i = 1, 2, 3, \ldots, n$. In practice, a population variance is usually an estimate derived from a 'small' sample of relatively few laboratory measurements. Depending on the particular situation, M could well be as low as 10 or greater than 100. For routine high throughput laboratory tests such as serum sodium activity or blood haemoglobin concentration, M, which now represents the number of internal quality control specimens, may well be greater that 1000. Whether M is large or small, the population variance is obtained as an unbiased estimate using the standard formula given in B12 for a sample with M items and a divisor of $M-1$ degrees of freedom.

Uncertainty Components with Covariance or Correlation

The procedure described above of taking the mean of the cross-product terms as zero when input values are uncorrelated requires additional comment. As discussed further below, the mean value of $\delta x_1 \delta x_2$ is equivalent to the covariance and can be written as $r_{1,2}\, u(x_1)\, u(x_2)$, where $r_{1,2}$ is the *correlation coefficient* between x_1 and x_2. Thus, when there is significant correlation between any two or more of the inputs, say between x_1 and x_2, the product $\delta x_1 \delta x_2$ will *not* be zero even though both δx_1 and δx_2 have zero mean values. Conversely, if the mean of the product $\delta x_1 \delta x_2$ is zero, the input variables x_1 and x_2 are uncorrelated (see below, 'A further note on correlation').

When there is correlation therefore, B14 has additional terms in which the means of the cross-products (the covariances) have been replaced by their equivalent expressions with correlation coefficients and uncertainties. That is:

$$u^2(y) = \left(\frac{\partial y}{\partial x_1}\right)^2 u^2(x_1) + \left(\frac{\partial y}{\partial x_2}\right)^2 u^2(x_2) + \ldots + \left(\frac{\partial y}{\partial x_n}\right)^2 u^2(x_n) +$$

$$2r_{1,2}\left(\frac{\partial y}{\partial x_1}\frac{\partial y}{\partial x_2}\right)u(x_1)\,u(x_2) + 2r_{1,3}\left(\frac{\partial y}{\partial x_1}\frac{\partial y}{\partial x_3}\right)u(x_1)\,u(x_3) + \ldots +$$

$$2r_{2,3}\left(\frac{\partial y}{\partial x_2}\frac{\partial y}{\partial x_3}\right)u(x_2)\,u(x_3) + \ldots \qquad \text{(B15)}$$

Equation B15 is the general form of the law for the propagation of uncertainty.

Example – the Propagation of Uncertainty Equation

To illustrate a particular use of the full propagation equation (B15) where correlations are taken into account, suppose that the measurand y is given simply by the product of two inputs x_1 and x_2 which have standard uncertainties $u(x_1)$ and $u(x_2)$ respectively. That is:

$$y = x_1 x_2 \qquad \text{(B16)}$$

Suppose first that x_1 and x_2 are uncorrelated ($r_{1,2} = 0$). Then, since $\partial y/\partial x_1 = x_2$ and $\partial y/\partial x_2 = x_1$, we obtain using B14 for uncorrelated inputs:

$$u^2(y) = x_2^2 u^2(x_1) + x_1^2 u^2(x_2) \qquad \text{(B17)}$$

Now suppose that $x_1 = x_2$, the same single input. Input x_1 is now perfectly correlated with input x_2, as any quantity is perfectly correlated with itself. Then B16 is simply:

$$y = x_1^2 \quad \text{and} \quad dy/dx_1 = 2x_1 \qquad \text{(B18)}$$

Combining equations B14 and B18 gives:

$$u^2(y) = 4x_1^2 u^2(x_1) \qquad \text{(B19)}$$

This same result (B19) is also obtained if we keep the notation for x_1 and x_2 separate, but use the propagation equation which involves correlation. If we start with B17, an additional general correlation term is now required. That is:

$$u^2(y) = x_2^2 u^2(x_1) + x_1^2 u^2(x_2)$$

$$+ 2r_{1,2}(\partial y/\partial x_1)(\partial y/\partial x_2)u(x_1)u(x_2) \qquad \text{(B20)}$$

$$\text{or} \quad u^2(y) = x_2^2 u^2(x_1) + x_1^2 u^2(x_2) + 2r_{1,2}\,x_1 x_2\,u(x_1)\,u(x_2)$$

In B20 we now put $x_1 = x_2$ and therefore $r = +1$. This gives:

$$u^2(y) = 2x_1^2\,u^2(x_1) + 2x_1^2\,u^2(x_1)$$
$$= 4x_1^2\,u^2(x_1) \qquad \text{(B21)}$$

This is the same result as obtained using B19 and demonstrates the consistency of the full propagation equation.

Standard Uncertainty of the Mean

Another interesting application of the full propagation equation (B15) provides a proof of the well known formula for calculating the standard uncertainty of the mean (also referred to as the standard deviation of the mean or standard error of the mean). Suppose that the inputs $x_1, x_2, x_3, \ldots, x_n$ all have the same units and are simply repetitions or replications of the same measurement. In this situation, we may assume the inputs as drawn from the same population with variance $u^2(x)$, where now there is no need for the subscript i on the x of $u^2(x)$. The usual purpose of replicated measurements is to obtain the mean y of the n inputs, so the measurand y is simply the mean of the inputs. The functional relationship $y = f(x_1, x_2, x_3, \ldots, x_n)$ becomes:

$$y = \frac{x_1 + x_2 + \ldots + x_n}{n} \qquad \text{(B22)}$$

In this particular case, the partial derivatives are given by

$$\frac{\partial y}{\partial x_1} = \frac{1}{n}, \quad \frac{\partial y}{\partial x_2} = \frac{1}{n}, \quad \ldots \quad \frac{\partial y}{\partial x_n} = \frac{1}{n} \qquad \text{(B23)}$$

All the partial derivatives are equal at the value of $1/n$, so B15 gives the standard uncertainty $u^2(y)$ of the measurand as:

$$u^2(y) = u^2(x)\left(\frac{1}{n^2} + \frac{1}{n^2} + \ldots + \frac{1}{n^2}\right) +$$

$$u^2(x)\left(r_{1,2}\frac{2}{n^2} + r_{1,3}\frac{2}{n^2} + r_{2,3}\frac{2}{n^2} + \ldots\right) \qquad \text{(B24)}$$

In the first bracket of B24, n terms are summed. If all the inputs are uncorrelated, then all the r's in the second bracket are zero and B24 then becomes:

$$u^2(y) = u^2(x)\frac{n}{n^2} = \frac{u^2(x)}{n} \qquad \text{(B25)}$$

Thus, the standard uncertainty $u(y)$ in y is given by:

$$u(y) = \frac{u(x)}{\sqrt{n}} \qquad \text{(B26)}$$

Equation B26 may be recognised as the formula for the *standard uncertainty of the mean*. As expected, when the input measurements are uncorrelated the standard uncertainty of the mean is less than the standard uncertainty $u(x)$ of the n inputs by a factor $1/\sqrt{n}$. This is indeed why the mean is often used for summarising measurements: the mean is better 'buffered' against random errors than are any of the individual component measurements. However, the often overlooked proviso is that the measurements must be uncorrelated, otherwise B14 (and B26) will be invalid. For example, the presence of a drift or trend in the measurements is an indication that some correlation may be present. In such cases, $u(y)$ will be greater than $u(x)/\sqrt{n}$ and may even approach $u(x)$ itself. As might be expected in most of these more complex situations, techniques do exist for estimating the relationship between $u(y)$ and $u(x)$. When drift does appear to be present in a long sequence of measurements, more specialised techniques for estimating uncertainty have been described.[36,37] If implementing such specialised procedures is not practical, a justifiable approach would be to take $u(y) \approx u(x)$.

A Further Note on Correlation

Section 5.2 of the GUM describes the approach required when correlated input quantities must be considered. The correlation coefficient $r_{1,2}$ between two variables, x_1 and x_2, is generally defined as:

$$r_{1,2} = \frac{\text{covariance}\,(x_1, x_2)}{\sqrt{\text{variance}\,(x_1)\text{variance}\,(x_2)}} \qquad \text{(B27)}$$

Depending on the particular statistical textbook, B27 may be presented in a number of different forms. The various terms which contribute to B27 are:

$$\text{covariance}\,(x_1, x_2) = \frac{\sum_{i=1}^{M}(x_{1,i} - \bar{x}_1)(x_{2,i} - \bar{x}_2)}{M} \qquad \text{(B28)}$$

$$\text{variance of } x_1 = \frac{\sum_{i=1}^{M}(x_{1,i} - \bar{x}_1)^2}{M}$$

$$\text{variance of } x_2 = \frac{\sum_{i=1}^{M}(x_{2,i} - \bar{x}_2)^2}{M} \qquad \text{(B29)}$$

where M is the sample size and $\bar{x}_1$ and $\bar{x}_2$ are the sample means of the variables x_1 and x_2 respectively. Combining the expressions in B28 and B29 gives the more usual form of B27:

$$r_{1,2} = \frac{\sum_{i=1}^{M}(x_{1,i} - \bar{x}_1)(x_{2,i} - \bar{x}_2)}{\sqrt{\left[\sum_{i=1}^{M}(x_{1,i} - \bar{x}_1)^2\right]\left[\sum_{i=1}^{M}(x_{2,i} - \bar{x}_2)^2\right]}} \qquad \text{(B30)}$$

There are four important points to be noted with respect to B30:

- The range of r is −1 through zero to +1. These values correspond to perfect negative correlation, complete absence of correlation and to perfect positive correlation respectively.
- The correlation between x_1 and x_2 is exactly the same as between x_2 and x_1; that is $r_{1,2} = r_{2,1}$.
- The same equation would be obtained if either M or $M-1$ were to be used in the denominators of equations B28 (the covariance) and B29 (the variances). That is, the correlation coefficient is the same whether population parameters or their estimates, sample statistics, are being used.
- The correlation coefficient is a dimensionless quantity. This can be seen from B30, where the dimensions of r are $x_1 \times x_2$ divided by $\sqrt{x_1^2 \times x_2^2}$.

As outlined in Appendix A, uncertainties in measurement are created by errors, where an error may be described as a *small departure* in a quantity from its 'true' or actual value. If we assume that an estimate of the actual value is provided by the sample mean obtained from a set of replicate measurements x_i (i = 1, 2, 3, … , n), the error δx_i is given by:

$$\delta x_i = (x_i - \bar{x}) \qquad \text{(B31)}$$

This implies that for variable x_1 with mean $\bar{x}_1$, the small error or departure $\delta x_{1,i}$ from the mean may be written as:

$$\delta x_{1,i} = (x_{1,i} - \bar{x}_1) \quad \text{(B32)}$$

and similarly for variable x_2:

$$\delta x_{2,i} = (x_{2,i} - \bar{x}_2) \quad \text{(B33)}$$

Equation B30 may therefore be written:

$$r_{1,2} = \frac{\sum_{i=1}^{M} \delta_{1,i}\, \delta_{2,i}}{\sqrt{(\sum_{i=1}^{M} \delta_{1,i}^2)(\sum_{i=1}^{M} \delta_{2,i}^2)}} \quad \text{(B34)}$$

or

$$r_{1,2} = \frac{\frac{1}{M}\sum_{i=1}^{M} \delta_{1,i}\, \delta_{2,i}}{\sqrt{\left(\sum_{i=1}^{M} \frac{\delta_{1,i}^2}{M}\right)\left(\sum_{i=1}^{M} \frac{\delta_{2,i}^2}{M}\right)}} \quad \text{(B35)}$$

From their definitions in B32 and B33, the deviations (errors) $\delta_{1,i}$ and $\delta_{2,i}$ have zero means. If the mean value of their product ($\delta_{1,i}\, \delta_{2,i}$) which appears in the numerator of B35 is also zero, then $r_{1,2} = 0$. That is, if the covariance is zero, the correlation coefficient is zero. Conversely, if $r_{1,2} = 0$, then the mean value of the product $\delta_{1,i}\, \delta_{2,i}$ is zero. This now verifies the statement in the first paragraph under the heading 'Uncertainty components with covariance or correlation.'

When a small number of data items have been used to determine a correlation, it is not expected that $r_{1,2}$ will be exactly zero for uncorrelated variables, due to inherent sampling variation. It is expected however that $r_{1,2}$ will be small. From a practical perspective, any assessment of the correlation between two variables should only be attempted with a substantial number of measurements. A suggested minimum would be 30 measurement pairs.