

# A novel practical approach to calculate measurement uncertainty in clinical pathology laboratories using quality control data with the use of biological variation where applicable

## Abstract

**Background:** Measurement Uncertainty (MU) can assist the interpretation and comparison of the laboratory results against international diagnostic protocols, facilitate a reduction in health care costs and also help protect laboratories against legal challenges. Determination of MU for quantitative testing in clinical pathology laboratories is also a requirement for ISO 15189.

**Methods:** A practical and simple to use statistical model has been designed to make use of data readily available in a clinical laboratory to assess and establish MU for quantitative assays based on internal quality control data to calculate Random Error and external quality assurance scheme results to calculate Systematic Error. The model explained in this article has also been compared and verified against quality specifications based on Biological Variation.

**Results:** Examples that explain and detail MU calculations for the proposed model are given where different components of MU are calculated with tabulated results.

**Conclusions:** The designed model is cost-effective because it utilises readily available data in a clinical pathology laboratory. Data obtained from internal quality control programs and external quality assurance schemes are used to calculate the MU using a practical and convenient approach that will not require resources beyond what is available. Such information can additionally be useful not only in establishing limits for MU to satisfy ISO 15189 but also in selecting and/or improving methods and instruments in use. MU can as well play an important role in reducing health care costs as shown by examples in the article.

**Keywords:** measurement uncertainty, MU, clinical pathology testing, quantitative testing, random error, systematic error

Volume 10 Issue 5 - 2021

Ashraf Mina,<sup>1,2</sup> Shanmugam Banukumar,<sup>1</sup> Santiago Vazquez<sup>1</sup>

<sup>1</sup>NSW Health Pathology, Forensic & Analytical Science Service (FASS), Toxicology Unit, Macquarie Hospital, NSW, Australia

<sup>2</sup>Affiliated Senior Clinical Lecturer, Faculty of Medicine and Health, Sydney University, NSW, Australia

**Correspondence:** Ashraf Mina, NSW Health Pathology, Forensic & Analytical Science Service (FASS), Toxicology Unit, Macquarie Hospital, NSW, Australia, Tel +61 02 9887 5666, +61 0403 108 682, Fax +61 02 9805 1259, Email ashraf.mina@health.nsw.gov.au

**Received:** September 27, 2021 | **Published:** October 20, 2021

**Abbreviations:** AACB, Australasian association of clinical biochemists; BV, biological variation; CU, combined uncertainty; CSIRO, commonwealth scientific and industrial research organization; EFLM, European federation of clinical chemistry and laboratory medicine; ESDM, experimental standard deviation of the mean; EQAS, external quality assurance scheme; IQC, internal quality control; MU, measurement uncertainty; NML, national measurement laboratory; RE, random error; RCPA, royal college of pathologists of Australasia; SD, standard deviation; SE, systematic error; TE, total error

## Introduction

Measurement Uncertainty (MU) is defined as the parameter associated with the result of a measurement that characterises the dispersion of the values of the measurand.<sup>1-4</sup> In practical clinical laboratory terms, uncertainty is referred to as total error, which is a combination of both random and systematic error. The MU concept is not incongruent with the daily practice in a clinical laboratory. For example, calibrated pipettes that are used to make up External Quality Assurance Scheme (EQAS) materials and also thermometers have stated MU values. For example, a 2mL glass graduated pipette may state it measures 2mL±0.01mL and thermometer specifications may state it measures a certain temperature range with ±0.1 °C.

Some articles recognized the importance of the MU concept but also acknowledged lacking an agreed practical method to calculate it in a clinical laboratory.<sup>5</sup> The problem faced in some of the proposed models to measure MU is the assumption of bias can be eliminated, corrected, or ignored. Resolving the differences in these concepts and approaches is currently a global issue.<sup>6,7</sup> There are many valid reasons for measuring MU in clinical laboratory practice. MU can expose an increased cost of certain services in healthcare. One study estimated that errors in measurements of calcium levels in the blood might add US\$ 60-199 million in annual healthcare costs. Analysing data for more than 89,000 patients receiving serum calcium tests, calibration errors were found to add US\$ 8-89 per patient to the healthcare of about 3.5 million patients. High calcium levels can be a symptom of cancer or thyroid disorder. Accurate measurements are critical because calcium levels in healthy people fall within a narrow range (2.13-2.63 mmol/L). Results for up to 15% of calcium laboratory tests were estimated to contain calibration errors of 0.03-0.13 mmol/L. Thus, some results that fall in the centre of the normal range, such as 2.43 mmol/L, might be for patients with elevated calcium, defined as 2.55 mmol/L and above, who will therefore receive no follow-up. On the other hand, patients with values above the threshold might derive from those who have normal calcium levels, and they may receive unnecessary follow up such as hormone measurements and

chest X-rays. Increasing the use of international or external diagnostic protocols and the need to correlate targets to local laboratory results also drives the requirement for MU. Some examples here include lipids, pediatric bilirubin, endocrine dynamic function tests, diagnosis of diabetes mellitus, and glycohaemoglobin monitoring. Different laboratories use different methodologies, instrumentation, and calibrators. A statement of MU will help make better sense of test results, particularly when comparing results between laboratories.<sup>8</sup> A legal challenge is another imperative driving MU. For example, if an athlete is banned after being tested for illicit doping because the cut-off limit was 5 ng/mL and the athlete's test result was 6 ng/mL and then it was subsequently determined that the MU for that test was  $\pm 2$  ng/mL. This means in practice that the original test result of 6 ng/mL could have been within normal limits. In legal terms, the laboratory could be held liable for damage or potential loss of earnings caused by the initial decision to ban the athlete since it was based on the initial laboratory finding. In Australia, legal measurements made under the National Measurement Act 1960 and calibration reports issued by Commonwealth Scientific and Industrial Research Organization (CSIRO) National Measurement Laboratory (NML) contain an uncertainty statement based on a 95% confidence level.

MU is a requirement under the ISO guidelines 15189. MU should not be considered just a parameter to fulfil the accreditation standards, but it should be used as a key quality indicator to describe both the performance of a measuring system and the laboratory itself.<sup>9</sup> Additionally, the designing of a protocol to estimate MU is important to avoid overestimating MU, for example running the samples in duplicate to measure the imprecision will result in double counting the imprecision contribution to MU.<sup>10</sup> Other issues have been reported regarding the way MU is calculated.<sup>11</sup> Another article used MU to correct for interferences (bias).<sup>12</sup>

## Materials

Sources of laboratory errors should be identified. They could be preanalytical, analytical, and post-analytical.<sup>13</sup> Major sources of errors include differences in analytical methods used by different laboratory instruments, lot-to-lot variation in calibration materials, and lack of "traceability" between secondary reference materials and primary standards. Secondary reference materials are substances that do not have the same level of purity as primary standards but each one has been characterised for certain chemical or physical properties that can then be used in clinical chemistry. Primary reference materials (also referred to as primary standards) are highly purified materials that can be measured directly to produce a substance of 'exact' known concentration. Sources of laboratory error contribute to both Random Error (RE) and Systematic Error (SE). RE can be defined as any random deviation from the laboratory mean result. There is a level of "expected" or "acceptable" RE which generally lies between  $\pm 3$  SD of the mean. It may follow that any deviation greater than  $\pm 3$  SD from the mean is considered "unacceptable" RE. Because of its random nature, this type of error is unpredictable but can be estimated by repeated measurements. SE can be defined as a trend or shift from the assay mean (system bias). Small amounts of SE are tolerable. SE will remain until corrective action is taken. Repeated measurements will not reduce the size of SE or improve the knowledge of its magnitude or direction. Some other articles addressed MU different types of errors.<sup>14,15</sup> It is not uncommon that different methods may have different reference intervals to mask the inherited SE in that method.

## Methods

In MU calculations, it is important to consider a worst-case scenario. Accordingly, MU calculations that do not take both RE and SE into account are likely to provide underestimates of the actual MU. The proposed model described here is simple, cheap, and practical to calculate MU from existing analytical measurements. This is particularly important given the limited resources laboratories currently have in terms of budget pressure, time and staff.

Firstly, the measurement process should be under control before any assessment of uncertainty is attempted, and secondly, MU should express a "worst-case scenario". That does not mean that the process should overestimate MU, it simply means establishing a total error range, rather than either a RE or SE range. There are essentially six steps to determine MU.<sup>16</sup>

Step 1: Make a model of the measurement system

Modelling is a critical part of this process. If the model is not correct, it means that it will either underestimate or overestimate MU. In clinical laboratories, Total Error (TE) for a quantitative result consists of at least one RE component and at least one SE component. Thus, the model adopted is:

$$MU \text{ in a clinical setting} = (\text{Total Error, TE}) = RE + SE$$

Step 2: List all the sources of uncertainties

Sources of uncertainties should be identified and listed, but those contributing less than 1/5 to 1/3 of the total MU will not have much impact on the combined uncertainty and can usually be disregarded unless several sources fall into this category. Other potential sources of error such as pre-analytical factors including collection and transportation should not be an issue because specimens that do not comply with test prerequisites should be rejected, or in some situations, the result is released with a comment addressing the problem that might have been caused. Nevertheless, these events should not represent the norm that reflects the day-to-day operation and Good Laboratory Practice in place. The sources of errors should be compiled under the two categories of RE and SE. It is acceptable to combine the factors that can affect an analyser and/or tests performed on it together, instead of doing separate analyses for each test. Examples of Random and Systematic Errors and some of the factors that can cause either Random or Systematic Errors are summarized in Table 1.

Step 3: Calculate the standard uncertainties for each component using "Type A" analysis for those with repeated measurements and "Type B" for others.

"Type A" evaluation of standard uncertainty component is an evaluation of standard uncertainty by statistical processing. This is only possible when a series of observations such as repeated measurements are feasible. The most common reason for taking repeated measurements is to minimise the potential effect of RE or 'noise'. This type of calculation is used to calculate RE. The standard uncertainty is calculated as follows:

$$ESDM = \frac{S.D.}{\sqrt{n}} \text{ Standard Uncertainty or}$$

Standard uncertainty or ESDM is the experimental standard deviation of the mean. It is the uncertainty component due to the random variation in resultant means obtained from repeated sets of measurements, each with "n" (i.e. number of) measured values.

**Table 1** Examples of Random and Systematic Errors. Some factors can cause either Random or Systematic Errors

Type of Error	Causes
Random Error (RE)	1. Power supply.
	2. Misplacement of control specimen in the run.
	3. Air bubble in water supply, where applicable.
	4. Random air bubbles in reagent or sample pipette system.
	5. Double pipetting of a control sample.
	6. Incorrect reconstitution of the control product.
	7. Inappropriate storage of control in a frost-free freezer.
	8. Use of non-reagent grade water in the test system.
	9. Different operation techniques by different technologists.
Systematic Error (SE)	1. Improper alignment of sample or reagent pipettes.
	2. Drift or shift in incubator chamber temperature, where applicable.
	3. Inappropriate temperature/humidity levels in the testing area.
	4. Change of reagent or calibration lot.
	5. Deterioration of reagent while in use, storage or shipment.
	6. Deterioration of calibrator while in use, storage or shipment.
	7. Deterioration of control product while in use, storage or shipment.
	8. Incorrect handling of control product (e.g., freezing when not recommended).
	9. Recent calibration.
	10. A dirty filter, where applicable.
	11. Failing light source.
	12. Inappropriate storage of control in a frost-free freezer.
	13. Use of non-reagent grade water in the test system.
	14. Different operation techniques by different technologists.

Standard Deviation (SD) may be obtained from any relevant Internal Quality Control (IQC) program. Such a commutative standard deviation should satisfy the “Type A” measurement at the corresponding commutative level (commutative mean) of the measurand (analyte). Ideally, the IQC levels should represent or be close to the medical decisions limit; usually the upper limit or lower limit of the reference intervals. If the medical decision limit is not covered by IQC, then the alternatives are to choose another QC

material that covers that limit, to choose a Certified Reference Material (CRM) at that limit, or if possible to obtain a patient specimen that is close to the required limit. Table 3 gives an example of the “Type A” calculation.

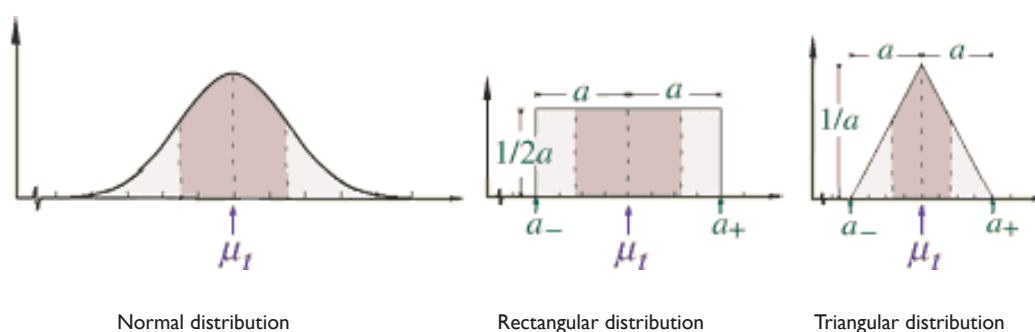
“Type B” evaluation of uncertainty component is an evaluation by means other than by established statistical methods and is used for estimation of SE. It involves using data provided for uncertainties on calibration certificates or from the manufacturer’s specifications, reference tables, and books. It involves sound scientific and engineering estimates of limits by experts and scientific societies. Lack of “traceability” between secondary reference materials and primary standards is one of the main factors that may contribute to SE or bias. As an example, if such bias is not discovered or recognised by a chemical kit manufacturer, a recall of the affected product is usually in place. In some cases, the manufacturer may compensate for that bias by assigning a value, if workable.

It is important to recognise that for many SEs the magnitude of the error is not known, only the limits within which it lies. The next step, therefore, involves deciding on the distribution that best describes the dispersion of the errors. The most common methods used to estimate that sort of error are shown in Figure 1 (with permission from the National Institute of Standards and Technology, USA)<sup>15</sup> and in Table 2.

The proposed model uses the rectangular distribution. It is the most convenient for clinical assays because there is no assumption that the data collected follows a normal distribution. The EQAS could be used as a source of information to calculate SE (assay bias). To calculate bias using EQAS, select the closest result to the internal QC analyte level from an end-of-cycle report Table 3. A laboratory can pool the results of more than one cycle of EQAS and exclude problematic results that do not accurately represent its performance. Some programs perform regression analysis and provide slope and intercept. If EQAS does not provide these parameters, regression analysis could be performed by putting the EQAS results on the *X*-axis and the results obtained by the laboratory for that EQAS on the *Y*-axis. It is beyond the scope of this report to discuss different types of regression, but if the laboratory is going to perform the regression, it is recommended they use weighted Deming regression analysis to calculate the slope and intercept. The laboratory can still use simple regression analysis, in “Type B”, but weighted Deming regression is superior as it does not assume the reference method, in this case, EQAS results, is free from errors and it also compensates for proportional SEs either negative or positive. It is worth mentioning that *average bias* should not be used as it will overestimate the bias at a lower level and underestimate it at higher levels, as shown in Table 4. If bias is constant about the mean, divide it by two to calculate the semi-range (*a*). If there is a systematic proportional error, make it equal to the larger value which means do not divide it by 2. By using the equation used to calculate rectangular distribution standard uncertainty, the standard uncertainty for bias is then calculated (Table 5).

$$\text{Rectangular Distribution Standard Uncertainty} = \frac{a}{\sqrt{3}}$$

If bias cannot be calculated from EQAS due to problems with the sample matrix for example, then the manufacturer of the instrument and/or calibrator should be contacted to provide the calibration certificate that provides a value for MU. If the certificate provides “expanded uncertainty” at 95% confidence intervals, divide it by 2 to calculate the standard uncertainty.



**Figure 1** Normal, rectangular and triangular distribution. (With permission from the National Institute of Standards and Technology, USA).

**Table 2** The most commonly used distribution for Type “B”

Distribution	Characteristic	Equation
Normal distribution	If we know the uncertainty component to be normally distributed and can estimate approximate 100% limits $\pm$ semi-range, denoted by “a”	Standard uncertainty = $\frac{a}{3}$ This is based on 99.7% of the area under the curve being within $\pm 3$ SD
Rectangular distribution	It is the distribution of minimal knowledge because all we know are the limits. The maximum and minimum limits are assumed to be equal and called the semi-range, denoted by “a”.	Standard uncertainty = $\frac{a}{\sqrt{3}}$ If bias is symmetric about the mean, divide it by 2 to calculate semi-range (a). If asymmetric, make it equal to the larger value (i.e. do not divide by 2)
Triangular distribution	It is used where there is evidence that the values near the mean are the most probable and, as the limits are approached, the probability decreases to zero.	Standard uncertainty = $\frac{a}{\sqrt{6}}$ It represents less knowledge than in the normal distribution case, but more knowledge than the rectangular distribution case

**Table 3** Using internal quality control data to calculate imprecision or “Type A” measurement

Analyte Name	Internal QC Analyte Level	Random Error (Internal QC)			
		S.D.	Specimen No(n)	$\sqrt{n}$	$ESDM = \frac{S.D.}{\sqrt{n}}$
Cortisol	88	8	171	13.1	0.612
	578	40	170	13	3.068
	884	64	170	13	4.909

**Table 4** Using external quality control schemes data to calculate bias or “Type B” measurement

Analyte name	Internal QC Analyte Level	QAP Target or Median	Result obtained by Lab	(QAP Target or Median X slope) $\pm$ Intercept	Bias	Average Bias
Cortisol (Low)	88	134	143	$[(134 \times 1.2) - 42.8] = 118$	$134-118 = 16$	77
Cortisol (Medium)	578	553	588	$[(553 \times 1.2) - 42.8] = 621$	$553-621 = -68$	
Cortisol (High)	884	950	1187	$[(950 \times 1.2) - 42.8] = 1097$	$950-1097 = -147$	

**RCPA-AACB QAP, Cycle 20 (Cortisol: slope = 1.2, intercept = - 42.8)**



**Table 5** Using external quality control schemes data to calculate bias or “Type B” measurement for Cortisol

Analyte Name	Systematic Error (EQAS)		Rectangular Distribution Standard Uncertainty = $\frac{a}{\sqrt{3}}$
	Average Bias	Corresponding Bias* Semi-range (a)	
Cortisol	77	16	9.24
		-68	-39.26
		-147	-84.87

### Standard uncertainty = Expanded Uncertainty/ Coverage factor

#### Step 4: Calculate the sensitivity coefficients

Sensitivity coefficients convert all uncertainty components to the same unit as the measurand, which allows different components to be combined. Because “Type A” and “Type B” both use the same unit (the analyte unit), the sensitivity coefficient is equal to 1. Therefore, this would not affect the calculation in the proposed model. For certain measurements, if the units of the uncertainty components differ from the units of the measurand, a conversion factor (also known as sensitivity coefficients) will be required. For example, the measured values of temperature and the associated uncertainty will have units of °C, whereas alternative factors such as pressure may influence assay results, so a conversion factor will be required to obtain the effect of uncertain temperature in terms of pressure units. In a clinical setting, because we use the same unit, the sensitivity coefficient is 1. Sensitivity coefficients are not only related to the unit of the measurand. They may also be used when all uncertainty components are given in the same units.

**If all units are the same, then Sensitivity coefficient = 1**

Step 5: Calculate the combined uncertainty and, if appropriate, its effective degree of freedom

$$\text{Combined uncertainty} = \sqrt{\sum_{i=1}^n \left[ c \left( \frac{\text{standard uncertainty}}{\text{uncertainty}} \right) \right]^2}$$

- c: Sensitivity coefficient.

$\sum_{i=1}^n$  : The summation of all the terms, of which there are n

This means that the uncertainty components are converted to the same units as the measurand using sensitivity coefficients, then these products (“Type A” and “Type B”) are squared and summed up. So for this model, the results for “ESDM” that represent RE and “Standard Uncertainty” that represent SE are squared up, added together and the square root is taken for the total because the Combined Uncertainty (CU) is the square root of the sum.

$$2\sqrt{(ESDM)^2 + (Standard Uncertainty)^2}$$

Effective degrees of freedom are used to determine the coverage factor for a 95% confidence interval that is used to calculate expanded uncertainty (Step 6). In the suggested model, there is no need to calculate “effective degrees of freedom” as all the components that contribute to imprecision are combined under one main source (IQC), as is also done for assay bias (EQAS). This process should cover the

majority of assays. In special cases, if there are many components to be considered and a different number of specimens (n) used to determine the effect of each component, then “effective degrees of freedom” should be calculated as every component of uncertainty can have an appropriate number of degrees of freedom (v). This is to allow the correct selection of the value of the *Student’s t- table* that is to be used to determine the coverage factor for a 95% confidence interval. It also indicates how well a component may be relied on. A high number of (v) is associated with low variance or low dispersion. A low number of (v) corresponds to a large dispersion or poor confidence in the value.

In such cases, “Type A” calculation is straightforward, and the Standard Deviation (SD) can be used as the uncertainty due to the scatter of the measured values. For “Type B”, if the limits are determined with high confidence in their value, then,  $v = \infty$ . (Where  $t = 1.96$  for a 95% confidence interval). If the limits themselves have some uncertainty, a lesser number of v must be assigned by the following formula:

$$v = \frac{1}{2} (\Delta \text{ uncertainty} / \text{Total uncertainty})^{-2}$$

$$v = \frac{1}{2} (\text{relative uncertainty})^{-2}$$

This is a number <1, but for convenience can be thought of as a percentage or a fraction. The smaller the number, the better defined the magnitude of the uncertainty.

e.g. If Cortisol = 88 nmol/L, bias = 16 nmol/L. The relative uncertainty for bias is  $16/88 = 0.18$  (i.e. 18%).

Accordingly,  $v = \frac{1}{2} (0.18)^{-2} = 15$  (where  $t = 2.13$  for a 95% confidence interval)

So 2.13 can be used as a coverage factor for a 95% confidence interval.

Once the uncertainty components have been combined, the “v” in the combined uncertainty is calculated by the Welch-Satterthwaite equation<sup>16</sup> as follows:

$$v_{\text{eff}} = \frac{(\text{The combined standard uncertainty})^4}{\sum [( \text{standard uncertainty of a component})^4 / v]}$$

Where:  $v_{\text{eff}}$  = the effective degrees of freedom and  $\sum$  = the sum of all the terms, where each component is raised to the power of 4 and divided by its v

#### Step 6: Calculate the expanded uncertainty

Use a nominal or a calculated coverage factor. Round the measured value and the uncertainty to obtain the reported values. To have an

adequate probability that the value of the measurand lies within the range given by the uncertainty, the combined uncertainty is multiplied by a coverage factor. A factor of 2 gives an expanded uncertainty with an approximate 95% confidence level. The reason for that is that the values for the *t*-test for various degrees of freedom “v” for a 95 % confidence interval is around 2.

MU is usually expressed as a percentage, multiples of standard deviation, or in a range of the analyte units, as shown in Table 6.

#### Quality Specifications Based on Biological Variation (BV)

Biological variation refers to the variance in test results within a single individual over a specified period of time, and between individuals of a defined population. Biological variation (BV) could be used, if desired, to compare both the outcome of the calculation

and the performance of the method in use.<sup>17</sup> BV comprises the natural fluctuation of body fluid constituents around the homeostatic setting point and has two components: within and between-subject variation. Many articles have been published estimating the components of BV. The BV material was compiled from published articles referenced in BIOS, CURRENT CONTENTS, EMBASE, MEDLINE, books, and doctoral theses provided by the authors into a database. BV in pathological states was higher than in the healthy state for various quantities and, in these, quality specifications were derived only from the healthy population “most stringent”.<sup>17</sup> BV is a preferred model to set quality specifications because it is firmly based upon medical requirements, is generated using simple to understand models, is widely accepted by professionals in the field, is higher in the hierarchy of strategies to set quality specifications<sup>18,19</sup> and it is useable in all laboratories irrespective of size.

**Table 6** Calculating combined and expanded uncertainty for Cortisol with comparison against quality specifications based on Biological Variation (Total Error %)

Analyte Name	Combined Uncertainty (CU) in test units	Expanded Uncertainty in test units (CU x 2)	Percentage (%)	Quality Specifications Based on Biological Variation (%) (Total Error)
Cortisol	9.26	18.52	21	Minimum = 43.9
	39.38	78.76	13.6	Desirable = 28.9
	85.01	170.02	19.2	Optimum = 14.4

The within ( $CV_w$ ) and between-subject ( $CV_g$ ) components of variation, expressed in coefficients of variation and the corresponding quality specifications for analytical imprecision or variation ( $CV_A$ ), analytical bias ( $B_A$ ), and total error ( $TE_A$ ), expressed in percentages of the assay value. The full derivation of these equations is explained in other references.<sup>19</sup> Published  $CV_w$  and  $CV_g$  values are also available on the internet, such as The European Federation of Clinical Chemistry & Laboratory Medicine.<sup>20</sup> The minimum, desirable and optimum parameters for assay performance are summarised below.

#### 1) Minimum quality specification based on Biological Variation:

$$CV_A < 0.75 CV_w$$

$$B_A < 0.375 (CV_w^2 + CV_g^2)^{1/2}$$

$$TE_A < 0.375 (CV_w^2 + CV_g^2)^{1/2} + 1.65(0.75 CV_w)$$

#### 2) Desirable quality specification based on Biological Variation:

$$CV_A < 0.5 CV_w$$

$$B_A < 0.25 (CV_w^2 + CV_g^2)^{1/2}$$

$$TE_A < 0.25 (CV_w^2 + CV_g^2)^{1/2} + 1.65(0.5 CV_w)$$

#### 3) Optimum quality specification based on Biological Variation:

$$CV_A < 0.25 CV_w$$

$$B_A < 0.125 (CV_w^2 + CV_g^2)^{1/2}$$

$$TE_A < 0.125 (CV_w^2 + CV_g^2)^{1/2} + 1.65(0.25 CV_w)$$

## Results

Table 3 shows the calculation of EDSM using IQC to calculate “Type A” measurement. In this example, the results were obtained

from the end-of-cycle report of EQAS provided by the Royal College of Pathologists of Australasia (RCPA) and Australasian Association of Clinical Biochemists (AACB). The IQC levels used for Cortisol in our laboratory are, level 1 = 88 nmol/L, level 2 = 578 nmol/L and level 3 = 884 nmol/L.

To calculate SE, the EQAS end-of-cycle (number 20) was consulted to locate values closest to these; these were 134 nmol/L, 553 nmol/L, and 950 nmol/L respectively. The levels do not have to be the same, as the idea is to target the same general area from the line of regression. The corresponding EQAS target or median is multiplied by the obtained slope  $\pm$  intercept depending on the sign obtained. Then, the outcome is subtracted from the EQAS target or median to calculate the bias. The sign is an indication of positive or negative bias, which is also obvious from the comparison between the line of regression with the line of best fit. Using external quality control schemes data to calculate bias or “Type B” measurement are shown in Table 4.

It is worth mentioning that *average bias* should not be used as it will overestimate the bias at a lower level and underestimate it at higher levels (see Table 6 as an example). If bias is constant about the mean, divide it by two to calculate the semi-range (a). If there is a systematic proportional error, make it equal to the larger value (i.e. do not divide by 2). By using the equation used to calculate rectangular distribution standard uncertainty, the standard uncertainty for bias is then calculated. Using external quality control schemes data to calculate bias or “Type B” measurement for Cortisol as an example is explained in Table 5.

$$\text{Rectangular Distribution Standard Uncertainty} = \frac{a}{\sqrt{3}}$$

Table 6 shows the calculating combined and expanded uncertainty for Cortisol with comparison against quality specifications based on

Biological Variation (Total Error %). Table 7 Using internal quality control data to calculate imprecision or “Type A” measurement. Table 8 shows examples of using IQC data to calculate imprecision or “Type B” measurement. Table 9 shows examples of minimum, desirable and optimum criteria based on BV.  $TE_A$  values derived from BV are

calculated and tabulated (where possible against MU calculated using the proposed model) to compare MU against minimum, desirable and optimum criteria of assay performance by using a quality specification based on BV. Tables 10 show examples of calculating combined and expanded uncertainties and comparing the outcome with BV.

**Table 7** Shows examples from the clinical practice of using IQC data to calculate imprecision or “Type A” measurement

Analyte Name	Internal QC Analyte Level	Random Error (Internal QC)			
		S.D.	Specimen No(n)	$\sqrt{n}$	$ESDM = \frac{S.D.}{\sqrt{n}}$
DHEAS	0.5	0.1	148	12.2	0.008
	4.3	0.4	148	12.2	0.033
	17.1	1.3	148	12.2	0.107
Gastrin	30	5	37	6.1	0.822
	69	9	37	6.1	1.48
	171	17	37	6.1	2.795
hGH	7.6	0.4	79	8.9	0.045
	23.8	1.2	80	8.9	0.134
	58.3	3.7	80	8.9	0.414
17OHP	3.3	0.5	40	6.3	0.079
	13.9	2.4	40	6.3	0.379
	32.9	6	39	6.2	0.961
IgE	120	8	79	8.9	0.9
	164	11	79	8.9	1.238
	313	28	80	8.9	3.13
E2	297	47	144	12	3.917
	978	91	144	12	7.583
	2175	151	143	12	12.627
Progesterone	2	0.4	144	12	0.033
	25.7	2.5	144	12	0.208
	61.2	6.6	144	12	0.55
SHBG	32.7	2.6	29	5.4	0.483
	79.7	5.2	29	5.4	0.966
Testosterone	1.3	0.5	145	12	0.042
	19.2	2.3	145	12	0.191
	33.8	3.3	145	12	0.274
Vit.D3	53	5.6	27	5.2	1.078
	103.9	13.4	27	5.2	2.579
Aldosterone	160	14	36	6	2.333
	752	49	37	6.1	8.056
	1652	96	37	6.1	15.782
PTH	7	0.4	17	4.1	0.097
	38.5	1.9	16	4	0.475
Renin	378	50	17	4.1	12.127

Table Continued

Analyte Name	Internal QC Analyte Level	Random Error (Internal QC)			
		S.D.	Specimen No(n)	$\sqrt{n}$	$ESDM = \frac{S.D.}{\sqrt{n}}$
IGF-I	889	123	17	4.1	29.832
	4625	2107	16	4	526.75
	21.4	2.3	39	6.2	0.368
	52.4	5.2	39	6.2	0.833
C-Peptide	0.3	0.04	83	9.1	0.004
	1.21	0.11	84	9.2	0.012
	2.32	0.07	83	9.1	0.008
CA 15-3	18.9	1.5	51	7.1	0.21
	45.1	2.8	51	7.1	0.392
Ca 19-9	5.9	1.2	83	9.1	0.132
	39.2	3.4	81	9	0.378
Adrenaline	101	11	19	4.4	2.524
	283	19	19	4.4	4.359
Noradrenaline	297	27	23	4.8	5.63
	1167	119	23	4.8	24.813
Dopamine	1	0.2	20	4.5	0.045
	3.1	0.5	20	4.5	0.112
Metanephrine	0.8	0.1	23	4.8	0.021
	1.9	0.2	23	4.8	0.042
Normetanephrine	2.4	0.3	24	4.9	0.061
	7.1	0.5	24	4.9	0.102
5HIAA	21	2	8	2.8	0.707
	173	14	15	3.9	3.615
Amphetamine	224	19	150	12.2	1.557
	356	18	150	12.2	1.47
Benzodiazepine	231	15	150	12.2	1.225
	335	17	150	12.2	1.388
Cocaine	252	9	150	12.2	0.735
	338	10	150	12.2	0.816
Opiates	222	11	150	12.2	0.898
	378	17	150	12.2	1.388
Cannabinoids	33	3	150	12.2	0.245
	66	4	150	12.2	0.327



**Table 8** Using external quality control schemes data to calculate bias or “Type B” measurement

Analyte Name	Internal QC Analyte Level	Systematic Error (EQAS)		Rectangular Distribution Standard Uncertainty = $\frac{a}{\sqrt{3}}$
		Average Bias	Corresponding Bias* Semi-range (a)	
Cortisol	88	77	16	9.24
	578		-68	-39.26
	884		-147	-84.87
DHEAS	0.5	0.47	-0.6	-0.35
	4.3		-0.2	-0.12
	17.1		0.6	0.35
Gastrin	30	16.6	-12	-6.93
	69		7.7	4.45
	171		16	9.24
hGH	7.6	2.27	-3.9	-2.25
	23.8		-1.1	-0.64
	58.3		1.8	1.04
17OHP	3.3	0.93	0.8	0.46
	13.9		0	0
	32.9		-2	-1.15
IgE	120	7.5	-7	-4.04
	164		-7	-4.04
	313		-8.5	-4.91
E2	297	347.7	-31	-17.9
	978		-468	-270.2
	2175		-544	-314.08
Progesterone	2	3.7	-5	-2.89
	25.7		-1	-0.58
	61.2		5	2.89
SHBG	32.7	2	-3	-1.73
	79.7		-3	-1.73
Testosterone	1.3	6	1.3	0.75
	19.2		6.3	3.64
	33.8		10.3	5.95
Aldosterone	160	49.3	14	8.08
	752		-36	-20.78
	1652		-98	-56.58
PTH	7	1.2	-0.9	-0.52
	38.5		-2.1	-1.21

Table Continued

Analyte Name	Internal QC Analyte Level	Systematic Error (EQAS)		Rectangular Distribution Standard Uncertainty = $\frac{a}{\sqrt{3}}$
		Average Bias	Corresponding Bias* Semi-range (a)	
Renin	378	257	86	49.65
	889		193	111.43
	4625		471	271.93
IGF-I	21.4	3.7	3	1.73
	52.4		3	1.73
C-Peptide	0.3	0.1	0	0
	1.21		0.1	0.06
	2.32		0.2	0.12
CA 15-3	18.9	3.7	-3	-1.73
	45.1		-1	-0.58
Ca 19-9	5.9	21.3	-3	-1.73
	39.2		8	4.62
Adrenaline	101	23	13	7.51
	283		41	23.67
Noradrenaline	297	19	-8	-4.62
	1167		21	12.12
Dopamine	1	0.17	0.11	0.06
	3.1		0.34	0.2
Metanephrine	0.8	0.53	0	0
	1.9		0.5	0.29
Normetanephrine	2.4	2.1	-0.1	-0.06
	7.1		1.7	0.98
5HIAA	21	1	0	0
	173		-2	-1.15
Amphetamine	224	19	18	10.39
	356		20	11.55
Benzodiazepine	231	19.5	19	10.97
	335		20	11.55
Cocaine	252	12.5	10	5.77
	338		15	8.66
Opiates	222	16	14	8.08
	378		18	10.39
Cannabinoids	33	3.5	3	1.73
	66		4	2.31

\*If bias is symmetric about the mean, divide it by 2. If asymmetric, make it equal to the larger value

**Table 9** Examples of minimum, desirable and optimum criteria based on Biological Variation

Analyte	Biological Variation										
	CVW	CVg	Minimum			Desirable			Optimum		
			CVA	BA	TEA	CVA	BA	TEA	CVA	BA	TEA
CA-125	13.6	46.5	10.2	17.9	34.7	6.8	11.7	23	3.4	5.8	11.4
Cortisol	20.9	45.6	15.7	18.1	43.9	10.5	11.7	28.9	5.2	5.7	14.4
DHEAS	3.4	30	2.6	11.3	15.5	1.7	7.5	10.3	0.9	3.8	5.2
Insulin	21.1	58.3	15.8	22.7	48.8	10.6	14.8	32.2	5.3	7.3	16
E2	22.6	24.4	17	11.1	39.1	11.3	6.7	25.4	5.7	3.1	12.5
Testosterone	8.8	21.3	6.6	8.4	19.3	4.4	5.4	12.7	2.2	2.7	6.3
Aldosterone	29.4	40.1	22.1	17.2	53.5	14.7	10.7	34.9	7.4	5.1	17.2
CA 19-9	24.5	93	18.4	35.5	65.9	12.3	23.5	43.7	6.1	11.7	21.8
B2 Micro	5.9	15.5	4.4	6	13.3	3	3.9	8.8	1.5	1.9	4.4
CA 15-3	5.7	42.9	4.3	16.2	23.2	2.9	10.7	15.5	1.4	5.4	7.7

**Table 10** Calculating combined and expanded uncertainty with some examples to compare against quality specifications based on Biological Variation (Total Error %)

Analyte Name	Internal QC Analyte Level	Combined Uncertainty (CU) in test units	Expanded Uncertainty in test units (CU x 2)	Percentage (%)	Quality Specifications Based on Biological Variation (%) (Total Error)
Cortisol	88	9.26	18.52	21	Minimum = 44.7
	578	39.38	78.76	13.6	Desirable = 29.8
	884	85.01	170.02	19.2	Optimum = 14.9
DHEAS	0.5	0.35	0.69	138.6	Minimum = 15.5
	4.3	0.12	0.24	5.6	Desirable = 10.4
	17.1	0.36	0.73	4.2	Optimum = 5.2
Gastrin	30	6.98	13.95	46.5	
	69	4.69	9.37	13.6	
	171	9.65	19.3	11.3	
hGH	7.6	2.25	4.5	59.3	
	23.8	0.65	1.3	5.5	
	58.3	1.12	2.24	3.8	
17OHP	3.3	0.47	0.94	28.4	
	13.9	0.38	0.76	5.5	

Table Continued

Analyte Name	Internal QC Analyte Level	Combined Uncertainty (CU) in test units	Expanded Uncertainty in test units(CU x 2}	Percentage (%)	Quality Specifications Based on Biological Variation (%)(Total Error)
IgE	32.9	1.5	3	9.1	
	120	4.14	8.28	6.9	
	164	4.23	8.45	5.2	
E2	313	5.82	11.64	3.7	
	297	18.32	36.64	12.3	Minimum = 40.4
	978	270.31	540.61	55.3	Desirable = 27.0
Progesterone	2175	314.33	628.66	28.9	Optimum = 13.5
	2	2.89	5.77	288.7	
	25.7	0.61	1.23	4.8	
SHBG	61.2	2.94	5.88	9.6	
	32.7	1.8	3.6	11	Minimum = 31.6
	79.7	1.98	3.97	5	Desirable = 21.1
Testosterone					Optimum = 10.5
	1.3	0.75	1.5	115.6	Minimum = 19.5
	19.2	3.64	7.28	37.9	Desirable = 13.0
Vit.D3	33.8	5.95	11.91	35.2	Optimum = 6.5
	53	1.08	2.16	4.1	
	103.9	2.58	5.16	5	
Aldosterone	160	8.41	16.83	10.5	Minimum = 55.0
	752	22.29	44.58	5.9	Desirable = 36.7
	1652	58.74	117.48	7.1	Optimum = 18.3
PTH	7	0.53	1.06	15.1	
	38.5	1.3	2.6	6.8	
Renin	378	51.11	102.22	27	
	889	115.35	230.71	26	
	4625	592.8	1185.6	25.6	
IGF-I	21.4	1.77	3.54	16.5	

Table Continued

Analyte Name	Internal QC Analyte Level	Combined Uncertainty (CU) in test units	Expanded Uncertainty in test units(CU x 2}	Percentage (%)	Quality Specifications Based on Biological Variation (%)(Total Error)
	52.4	1.92	3.84	7.3	
C-Peptide	0.3	0	0.01	2.9	Minimum =17.6
	1.21	0.06	0.12	9.7	Desirable = 11.7
	2.32	0.12	0.23	10	Optimum =5.9
CA 15-3	18.9	1.74	3.49	18.5	Minimum = 23.3
	45.1	0.7	1.4	3.1	Desirable = 15.5
					Optimum = 7.8
CA 19-9	5.9	1.74	3.47	58.9	Minimum = 66.4
	39.2	4.63	9.27	23.6	Desirable = 44.4
					Optimum = 22.1
GHb - Variant	5.4	0.12	0.23	4.3	
	9.5	0.29	0.58	6.1	
Adrenaline	101	7.92	15.84	15.7	
	283	24.07	48.14	17	
Noradrenaline	297	7.28	14.56	4.9	
	1167	27.62	55.23	4.7	
Dopamine	1	0.08	0.16	15.5	
	3.1	0.23	0.45	14.6	
Metanephrine	0.8	0.02	0.04	5.2	
	1.9	0.29	0.58	30.7	
Normetanephrine	2.4	0.08	0.17	7	
	7.1	0.99	1.97	27.8	
5HIAA	21	0.71	1.41	6.7	
	173	3.79	7.59	4.4	
Amphetamine	224	10.51	21.01	9.4	
	356	11.64	23.28	6.5	
Benzodiazepine	231	11.04	22.08	9.6	



Table Continued

Analyte Name	Internal QC Analyte Level	Combined Uncertainty (CU) in test units	Expanded Uncertainty in test units(CU x 2)	Percentage (%)	Quality Specifications Based on Biological Variation (%) (Total Error)
Cocaine	335	11.63	23.26	6.9	
	252	5.82	11.64	4.6	
	338	8.7	17.4	5.1	
Opiates	222	8.13	16.27	7.3	
	378	10.48	20.97	5.5	
Cannabinoids	33	1.75	3.5	10.6	
	66	2.33	4.66	7.1	

## Discussion

In the proposed model, two types of assessments are used; “Type A” measures REs by using internal quality control data and is expressed in terms of EDSM. This type of error identifies with RE in a clinical setting as shown in Table 3. “Type B” measures Bias by using external quality assurance schemes data and is expressed in terms of Rectangular Distribution. This type of error identifies with SE in a clinical setting as shown in Table 4 and 5. The combined and extended uncertainty is then calculated as shown in Table 6. The equations used in the suggested model can be built into a spreadsheet that facilitates the calculations as shown in the Result section. Tables 7, 8 and 10 show how the model is applied on different analytes to calculate both “Type A” and “Type B” errors and then calculate the combined and extended MU with a comparison against the minimum, desirable and optimum Biological Variation criteria for assay performance. There are many other available MU guidelines. Some of the recommendations appear to be contradictory and some may underestimate MU. For example, whilst the RCPA guideline<sup>21</sup> recommends the use of EQAS as a source to determine CV% while the AACB Laboratory Implementation Guide recommends against this.<sup>22</sup> Moreover, the RCPA document recommends that factors such as BV and age-related values are best considered and included as a part of the reference interval, while the AACB document uses BV to calculate TE. Although BV data is a good source of information, it does not cover non-naturally occurring substances (e.g. drugs and antibiotics) and in-house assays. Furthermore, establishing BV data is very time-consuming and expensive, and generally beyond the capacity of most clinical laboratories. Others have suggested a way to add MU to reports to facilitate the interpretation of laboratory results.<sup>23-25</sup>

The RCPA document essentially recommends obtaining the MU measurement by multiplying CV% by 2. This appears to oversimplify (or under-estimate) the MU and also appears to conflict with ISO15189 guidelines which indicates that MU should at least have one component of RE and one component of SE. The AACB document calculates Bias in some examples by using “Calibrator Certificate”; however, in practice, such documents are not always readily available. Also, in the AACB document, the example for drugs

(Digoxin) does not have MU value, instead of providing a comment to address possible interference. Regardless of MU, assay limitations such as haemolysis, lipaemia, bilirubinaemia, and drug interference should be commented on if clinically significant. MU for quantitative assays requires a numerical value, not simply a comment to address interference, which should be in place already regardless of M.

## References

- Williams SLREMRA. *Quantifying Uncertainty in Analytical Measurement*. 2<sup>nd</sup> Edition ed: EURACHEM/CITAC; 2000.
- ISO/IEC. *ISO/IEC Guide 99:2007 International vocabulary of metrology — Basic and general concepts and associated terms*. 2007.
- ISO/IEC. *ISO/IEC Guide 98-3:2008 Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*. 2008.
- Meyer VR. Measurement uncertainty. *J Chromatogr A*. 2007;1158(1-2):15–24.
- Cerioti F. Deriving proper measurement uncertainty from Internal Quality Control data: An impossible mission? *Clin Biochem* 2018;57:37-40.
- Westgard JO, Westgard SA. Measuring Analytical Quality: Total Analytical Error Versus Measurement Uncertainty. *Clin Lab Med*. 2017;37(1):1–13.
- Westgard JO. Managing quality vs. measuring uncertainty in the medical laboratory. *Clin Chem Lab Med*. 2010;48(1):31–40.
- Gallaher MP ML, Klee GG, Schryver P. *The Impact of Calibration Error in Medical Decision Making*. National Institute of Standards and Technology (NIST); 2004.
- Infusino I, Panteghini M. Measurement uncertainty: Friend or foe? *Clin Biochem*. 2018;57:3–6.
- Kadis R. Evaluation of the measurement uncertainty: Some common mistakes with a focus on the uncertainty from linear calibration. *J Chromatogr A*. 2017;1499:226–229.
- Dallas Jones GR. Measurement uncertainty for clinical laboratories - a revision of the concept. *Clin Chem Lab Med*. 2016;54(8):1303–1307.
- Jones GR, Hawkins RC. Correcting laboratory results for the effects of

- interferences: an approach incorporating uncertainty of measurement. *Ann Clin Biochem.* 2015;52(Pt 2):226–231.
13. Padoan A, Sciacovelli L, Zhou R, et al. Extra-analytical sources of uncertainty: which ones really matter? *Clin Chem Lab Med.* 2019;57(10):1488–1493.
14. Frenkel RB, Farrance I. Uncertainty in Measurement: Procedures for Determining Uncertainty With Application to Clinical Laboratory Calculations. *Adv Clin Chem.* 2018;84:125–207.
15. NIST. *Evaluating Uncertainty Components*. National Institute of Standards and Technology, USA. 2017.
16. RR C. *Assessment of Uncertainties of Measurement for Calibration and testing Laboratories*. 2<sup>nd</sup> edn: National Association of Testing Authorities (NATA), Australia. 2002.
17. Ricós C, Alvarez V, Cava F, et al. Current databases on biological variation: pros, cons and progress. *Scand J Clin Lab Invest.* 1999;59(7):491–500.
18. Fraser CG, Kallner A, Kenny D, et al. Introduction: strategies to set global quality specifications in laboratory medicine. *Scand J Clin Lab Invest.* 1999;59(7):477–478.
19. CG F. *Biological Variation from Principles to Practice*. Press A, editor. 2001.
20. EFLM. *EFLM Biological Variation Database: The European Federation of Clinical Chemistry & Laboratory Medicine*. 2021.
21. RCPA. *Uncertainty of Measurement Guideline (Guideline No 2)*. Royal College of Pathologists of Australasia. 2004.
22. White GH, Farrance I. Uncertainty of measurement in quantitative medical testing: a laboratory implementation guide. *Clin Biochem Rev.* 2004;25(4):S1–24.
23. Plebani M, Sciacovelli L, Bernardi D, et al. What information on measurement uncertainty should be communicated to clinicians, and how? *Clin Biochem.* 2018;57:18–22.
24. Padoan A, Sciacovelli L, Aita A, et al. Measurement uncertainty in laboratory reports: A tool for improving the interpretation of test results. *Clin Biochem.* 2018;57:41–47.
25. Giroud C, Arnaud J, Vassault A. Uncertainty of measurement. *Ann Biol Clin (Paris)*. 2010;68:237–245.