Measurement Uncertainty in Forensic Toxicology: Its Estimation, Reporting and Interpretation

Rod G. Gullberg

Clearview Statistical Consulting, Snohomish, WA, USA

1. Introduction

All measurements, regardless of their purpose, context or quality, possess uncertainty. No measurement is performed with absolute perfection since all are approximations. Uncertainty, however, does not mean there is anything wrong or inappropriate with the results. Uncertainty is simply a measure of the confidence we have in our best estimate and results from limitations in our technology, our methods, our standards and our limited understanding of the property being measured. [Drosg] Uncertainty is a fundamental property of the natural world in which we live and work. Moreover, no measurement is fully interpretable within a given context until the full process generating the result is understood. The general additive measurement function observed in equation 1 illustrates this basic limitation of all measurements:

$$Y = \mu + \beta + \varepsilon \tag{1}$$

where: Y = the measurement result
μ = the true value of the measurand
β = measurement error due to bias
ε = random measurement error

Our measurement is an imperfect representation of the measurand due to bias and random error components. Bias may be corrected for when reliably determined with traceable controls. Random error, on the other hand, cannot be corrected for but can be minimized to an acceptable level. Figure 1 illustrates how these two contributors to uncertainty influence measurement results - where we have assumed a normal distribution. Bias is simply the difference between the mean and the reference value while random error, determined by the variance or standard deviation, defines the width of the distribution. Figure 1 also illustrates another important property of measurement - all results are random variables that arise from a specified distribution. As a result they have a fixed mean and variance from which confidence intervals can be determined – an useful metric for defining uncertainty. The fact that uncertainty exists in our measurements, however, should not alarm us. We simply need to understand it, acknowledge it, estimate it in a statistically valid way, report it and ensure that it is fit-for-purpose.
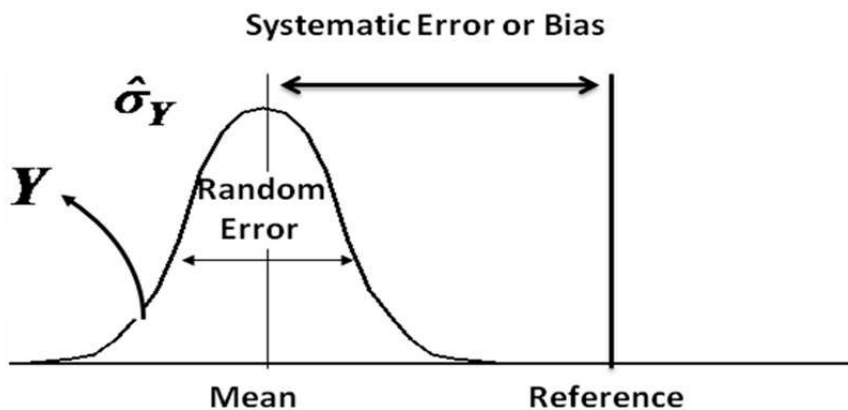
Systematic Error or Bias

Fig. 1. Measurement results, Y, are random representations from a distribution having a fixed mean and variance. The variance defines the random error while the mean relative to a reference defines their bias

Forensic toxicologists have a conceptual understanding of measurement uncertainty. However, most would probably find it difficult to actually compute a statistically valid estimate of the uncertainty, accounting for all relevant factors, and report it in an intuitive and comprehendible fashion for a jury to understand. For most analytical measurements performed by forensic toxicologists, both quantitative and qualitative, the formalization of measurement uncertainty is not generally considered or provided. This is due, in large part, to the lack of customer demand. The primary customers of forensic toxicologists are the courts and members of the legal community. They do not understand measurement uncertainty and are not aware of its relevance or importance. This, however, is changing. The legal community is becoming more aware of the concept and is now demanding it in several jurisdictions. The uncertainty allows the user to judge the quality and validity of the measurement results for a given application. Several factors have contributed to this renewed interest in measurement uncertainty. One is a recent report from the National Academy of Sciences in 2009. The NAS report states, "All results for every forensic science method should indicate the uncertainty in the measurements that are made,...". (NAS, 2009) The report was largely critical of the forensic sciences arguing the lack of a strong scientific foundation for their claims and practices. Another influencing factor has been the US Supreme Court decision in 1993 of Daubert vs. Merrell Dow Pharmaceuticals. The court required one of four criteria for admissibility to be "...the technique's known or potential rate of error...". (Daubert vs. Merrell Dow, 1993) The ruling requires that uncertainty be considered and accompany the introduction of measurement results in court. Finally, accrediting agencies are now requiring that forensic laboratories perform and report measurement uncertainty as part of their analytical protocol. The ASCLD/LAB-International accreditation program, for example, has adopted the ISO/IEC 17025 program and requires in part that, "...the laboratory estimate the measurement uncertainty for any area of testing or calibration where the customer makes the request or the jurisdiction or statute requires such". (ASCLD/LAB, 2011) These and other factors have now brought attention on this issue to measurement uncertainty. Forensic toxicologists need to address the issue and be prepared to compute, report and explain measurement uncertainty.

Moreover, providing the uncertainty along with measurement results is one important step in ensuring evidence-based inference. (Mnookin, et.al., 2011) We intend to illustrate and explain here several practical ways this can be accomplished.

Very basically, measurement uncertainty is best described by an interval, symmetric about the measurement result and within which we claim that the true value (the measurand) exists with some level of probability. The end points of this interval are called uncertainty or confidence limits. This interval quantifies the precision of the measurement result. Figure 2 illustrates this concept of uncertainty. The classical statistical view would state that the measurand (μ) is a fixed quantity and the measurement result along with the interval limits are random variables. The probability, therefore, relates to the random interval actually encompassing the fixed true value (μ). This involves some subtle distinctions between classical and Bayesian statistics which will not be discussed further here. Suffice it to say, our general approach regarding the estimation of measurement uncertainty will be classical in nature.
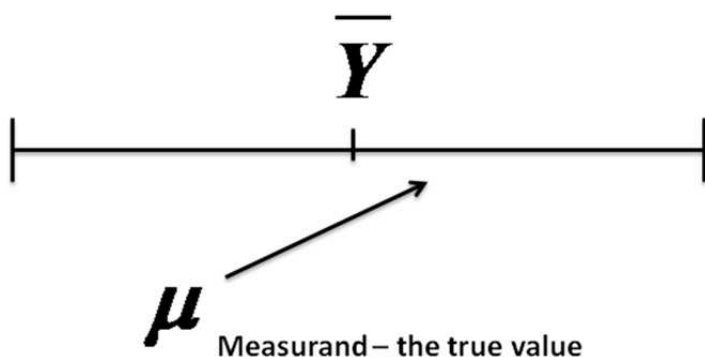

Fig. 2. Measurement uncertainty is best viewed as an interval symmetric about the mean and within which we claim the measureand lies with some stated level of probability

Not all measurement processes are capable of providing a rigorous and statistically valid estimate of uncertainty. This fact is acknowledged by metrologists and by the ISO 17025 document in particular. (IEC/ISO 17025, 2000) For these situations, ISO 17025 requires that the analyst or laboratory at least identify the uncertainty components and make a reasonable effort to express the uncertainty. All of the published guides on measurement uncertainty recognize that every measurement context is different and there are multiple ways for estimation. Accordingly, forensic toxicologists should develop a well reasoned documented approach that can be justified to both the legal and accrediting communities.

Consider the following two separate blood alcohol concentrations measured on samples from two different individuals: **0.086 g/dL, 0.104 g/dL.** Which result presents the stronger inference that the subject's true blood alcohol concentration exceeds 0.080 g/dL? Very simply, we do not know. We have no information regarding the measurement process or the uncertainty for each. Now consider the same two results along with their two standard deviation uncertainty estimates: **0.086 ± 0.005 g/dL, 0.104 ± 0.027 g/dL.** From this we now see that the first results (0.086 ± 0.005 g/dL) provide the stronger evidence that the individual's true blood alcohol concentration exceeds 0.080 g/dL. Figure 3 illustrates this as well. The

result of 0.104 g/dL actually has a significant probability that the true value is below 0.080 g/dL. This illustrates the additional value provided by measurement uncertainty, particularly in the cases near critical prohibited limits. Such information would be important for a court to consider.
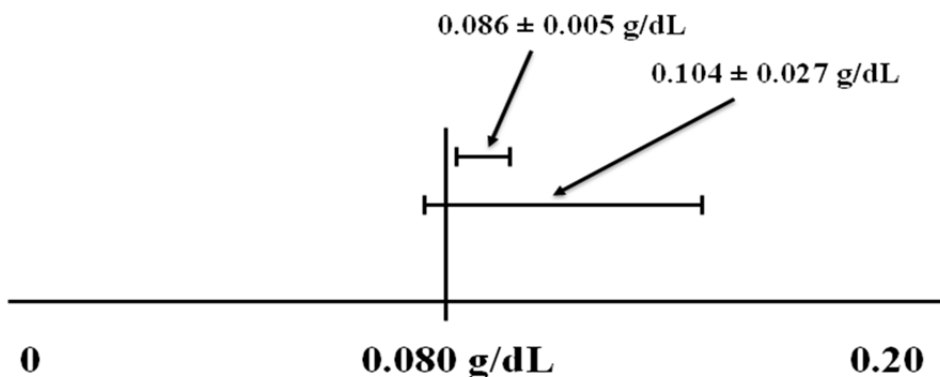

Fig. 3. Including measurement uncertainty adds considerable information when interpreting measurement results near critical concentrations

1.1 The meaning of Fit-for-purpose

Fitness-for-purpose (FFP) is a very important concept in analytical measurements designed to be used in important decision making contexts. FFP is the assurance that a measurement result will be suitable or appropriate for its intended applications. FFP is closely associated with uncertainty and the confidence that is necessary for a measurement result in a particular application. Measurement results in forensic toxicology have significant implications for the rights and property of individuals. Major consequences result from their interpretation in a legal context. For this reason, measurement results generated by forensic toxicologists must have a high level of confidence with minimum uncertainty to ensure their FFP. Determining the FFP in forensic toxicology can be challenging. (Thompson and Fearn, 1996) Toxicologists and customers should both contribute to establishing the appropriate FFP in a forensic context. Forensic toxicologists should continually strive to optimize their process and enhance the quality.

1.2 Published resources

There are a few important resource documents regarding measurement uncertainty that should be read and kept as references by the forensic toxicologist. These represent standards in the field of metrology. They are rigorous and well grounded theoretically. However, this does not mean there is uniform acceptance of these documents. There is a great deal of literature debating their application and interpretation. (Bich and Harris, 2006, Deldossi and Zappa 2009, Kacker,et.al. 2007, Kacker,et.al. 2010, Krouwer, 2003, Kristiansen, 2003) Three references of significant importance are:

1. *Guide to the Expression of Uncertainty in Measurement (GUM):* (ISO, 2008) This is commonly referred to as the *GUM* document and is published by ISO along with several other international standards organizations. The *GUM* provides primarily a

"bottom-up" approach to uncertainty estimation. They generally begin with an assumed measurement model and then proceed to employ the general method of error propagation.

2. EURACHEM/CITAC Guide, *Quantifying Uncertainty in Analytical Measurement:* (EURACHEM/CITAC, 2000) This document is similar to the *GUM* and provides all of the basic terminology and computations. The illustrated examples are more relevant to chemistry and may be more helpful to toxicologists.

3. NIST Technical Note 1297, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results:* (NIST, 1994) This document is brief but includes the key concepts and definitions. There are very few illustrated examples.

All of these documents are available on the internet and can be downloaded free of charge. There are also a large number of other documents and guidelines regarding measurement uncertainty available on the internet. As one begins to read this large body of literature it soon becomes apparent that there is no consensus in the analytical sciences on the best approach to estimating measurement uncertainty.

2. The measurement model

The measurement model is a mathematical function where the measurement result (the response variable) is expressed explicitly as a function of several input (predictor) variables. Equation 2 shows the general form:

$$Y = f(X_1, X_2, ..., X_n) \tag{2}$$

where: Y = the measurement result
X_i = the predictor or input variables

The values of X in equation 2 may represent quality control results, bias estimates, traceability components, a total measurement method component, calibrant materials, etc. Moreover, the values of X may themselves be functions of other input variables. The function f may be additive as illustrated in equation 3:

$$Y = X_1 + X_2 + ... + X_n \tag{3}$$

For additive models with independent input variables, the uncertainty is found from the root sum square (RSS) of the variance terms for each component as illustrated in equation 4:

$$u_Y = \sqrt{u_{X_1}^2 + u_{X_2}^2 + ... + u_{X_n}^2} \tag{4}$$

where: $u_{X_i}^2$ = the variance estimate for the ith variable

The function f may, on the other hand, be multiplicative as in equation 5:

$$Y = X_1 \cdot X_2 \cdot ... \cdot X_n \tag{5}$$

For the multiplicative model with independent variables the uncertainty is found by employing the RSS of the coefficients of variation squared as in equation 6:

$$\frac{u_{\overline{Y}}}{\overline{Y}} = \sqrt{CV_{X_1}^2 + CV_{X_2}^2 + ... + CV_{X_n}^2} \tag{6}$$

Notice also that equation 6 incorporates the mean $\overline{Y}$ and yields the standard deviation of the mean. This will result when we incorporate the appropriate sample sizes (values of n) for each term within the radical sign of equation 6. The function f may even be a combination of additive and multiplicative terms as in equation 7:

$$Y = \frac{X_1 \cdot X_2}{X_3 + X_4} - X_5 \tag{7}$$

In this case the uncertainty must be estimated by employing the general method of error propagation. The equation for this estimation is derived from the first-order (linear term) of the Taylor series expansion: (Ku, 1966)

$$\frac{u_{\overline{Y}}}{\overline{Y}} = \sqrt{\left[\frac{\partial Y}{\partial X_1}\right]^2 u_{X_1}^2 + \left[\frac{\partial Y}{\partial X_2}\right]^2 u_{X_2}^2 + \ldots + \left[\frac{\partial Y}{\partial X_n}\right]^2 u_{X_n}^2} \tag{8}$$

Equation 8 also assumes that all of the input variables are independent. When this is not the case, a covariance term must be added as seen in equation 9:

$$\frac{u_{\overline{Y}}}{\overline{Y}} = \sqrt{\sum_{i=1}^{n}\left[\frac{\partial Y}{\partial X_i}\right]^2 u_{X_i}^2 + 2\left[\frac{\partial Y}{\partial X_i}\right]\left[\frac{\partial Y}{\partial X_j}\right] Cov\left(X_i, X_j\right)} \tag{9}$$

 where:

$$Cov\left(X_i, X_j\right) = r_{\left(X_i, X_j\right)} S_{X_i} S_{X_j}$$

The value of r in equation 9 is the correlation coefficient between the two input variables. For each pair of input variables that are correlated an additional covariance term would need to be added. A simple example of a concentration measurement function that could apply to either blood or breath alcohol measurement is shown in equation 10:

$$C_{Corr} = \frac{C_0 R}{\overline{X}} \tag{10}$$

where: C_{Corr} = the corrected measurement concentration result
C_0 = the raw measurement results (either a mean or a single observation)
R = the traceable reference control value
$\overline{X}$ = the mean results from measuring the control reference standard (R)

Since equation 10 is multiplicative and we assume all three variables are independent we could employ the RSS for the CV's squared according to equation 11. Notice that we have incorporated the values of n, which may vary for each term, where this information is known. This will result in $u_{C_{Corr}}$ representing the standard deviation (or standard error) of the mean. Equation 12 illustrates a more complicated model that may represent the measurement of breath alcohol concentration. Bias in the breath test instrument is adjusted for by measuring controls which have been measured by gas chromatography and which in turn has had its bias accounted for by measuring other traceable controls.

$$\frac{u_{C_{Corr}}}{C_{Corr}} = \sqrt{CV_{C_0}^2 + CV_R^2 + CV_{\overline{X}}^2} = \sqrt{\left[\frac{\frac{u_{C_0}}{\sqrt{n_{C_0}}}}{C_0}\right]^2 + \left[\frac{u_R}{R}\right]^2 + \left[\frac{\frac{u_X}{\sqrt{n_X}}}{\overline{X}}\right]^2} \qquad (11)$$

$$\overline{Y}_{Corr} = \frac{\overline{Y}_0 \cdot GC_{Sol} \cdot R}{\overline{X} \cdot K \cdot GC_{Cont}} \qquad (12)$$

where: $\overline{Y}_0$ = the mean of the original n measurements

GC_{Sol} = the mean of the simulator solution measurements by gas chromatography

R = the traceable reference value of alcohol in water solutions purchased from a commercial vendor

$\overline{X}$ = the mean of the breath test instrument measuring the simulator solution heated to 34⁰C

K = 1.23 the ratio of partition coefficients relating to the simulator heated to 34⁰C

GC_{Cont} = the mean results from measuring the traceable controls on the gas chromatograph

Notice also that equation 12 is simply a set of correction factors that adjust for bias in the gas chromatograph as well as in the breath test instrument:

$$\overline{Y}_{Corr} = \frac{\overline{Y}_0 \cdot GC_{Sol} \cdot R}{\overline{X} \cdot K \cdot GC_{Cont}} = \overline{Y}_0 \cdot \left[\frac{GC_{Sol}}{\overline{X} \cdot K}\right] \cdot \left[\frac{R}{GC_{Cont}}\right] = \overline{Y}_0 \cdot f_{Inst} \cdot f_{GC} \qquad (13)$$

where: f_{Inst} = correction factor for the breath test instrument

f_{GC} = correction factor for the gas chromatograph

The uncertainty estimates for R and K will generally be Type B estimates available from certificates of analysis or other documentation. The other four factors will be Type A estimates since they are based on actual experimental results. The uncertainty computation for equation 13 can be determined from employing either the RSS method of equation 6 (since the function is multiplicative) or the error propagation method of equation 8. Both will yield the same estimate. We have illustrated only a few of the many measurement functions that may be relevant for forensic toxicologists. More examples are found in the *EURACHEM/CITAC Guide* as well as other literature sources. (Kristiansen and Peterson, 2004) The important point is to try and develop a model best describing the measurement process which will facilitate selecting the most appropriate uncertainty computation to perform. Where the measurement model is unknown it is common to assume a multiplicative form. The justification for this is the fact that variation generally increases with concentration, a property of a multiplicative model. (Kristiansen, 2001)

3. Traceability

Traceability is defined within the *VIM* document as a "...property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of

calibrations, each contributing to the measurement uncertainty". (ISO/VIM, 2008) Figure 3 illustrates this concept of traceability which links a measurement result (breath alcohol) to a national metrological authority with each link propagating its own uncertainty. The magnitude of uncertainty will increase with each additional level of the metrological chain. Since standards are imperfect there is the associated uncertainty that must be included as part of the final combined measurement uncertainty. The ultimate reference is usually a property maintained and defined by some metrological authority such as a National Metrological Institute (NMI). Chemical analytes are generally considered traceable to a method or standard reference material (SRM) such as NIST 1828b. There are other intermediate standards often used between the measurement result and the NMI. These are referred to as Certified Reference Materials (CRM) or simply Reference Materials (RM). (Thompson, 1997) Traceability is important for establishing the property of comparability and to determine and correct for bias. Uncertainty information regarding traceable standards are found on the certificates of analysis (COA).
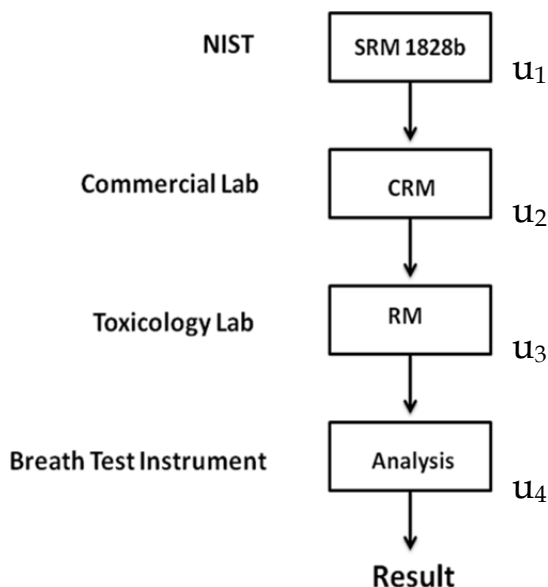

Fig. 4. Illustrating traceability where a measurement result is linked through an unbroken chain of comparisons to the national metrological authority

4. Practical steps for estimating measurement uncertainty

There are several valid approaches to estimating and quantifying measurement uncertainty. For our present purposes, we will present a very general "bottom-up" corresponding to the *GUM* document. Later, we will discuss other approaches as well. We will assume the

following eight basic steps for estimating measurement uncertainty that should generally apply for most quantitative measurements in forensic toxicology:
1. Clearly define the property to be measured (the measurand)
2. Identify the measurement function
3. Identify the components contributing to the measurement uncertainty
4. Quantify the standard uncertainty for each component
5. Combine the standard uncertainties for each component and compute the combined uncertainty
6. Compute the expanded uncertainty and the confidence interval
7. Produce the uncertainty budget
8. Report the results

Next, we present these steps in some detail. In addition we will present an example of blood alcohol measurement by gas chromatography and illustrate how each of the steps can be applied. We will assume duplicate blood alcohol results of 0.081 and 0.082 g/dL for this example.

4.1 Clearly define the measurand

It is very important that the customer and the toxicologist have a clear understanding of exactly the property being measured. Interpretation will then be applied to a specific measurand in a specific context where FFP can be appropriately determined. For our example we will assume that the measurand is the venous whole blood alcohol concentration collected from a specific individual at a specific time and location.

4.2 Identify the measurement function

We will assume the following basic model for our measurement of blood alcohol concentration (BAC) by headspace gas chromatography:

$$C_{corr} = \frac{C_0 R}{\overline{X}} \cdot f_{dilutor} \qquad (14)$$

where: C_{corr} = the corrected BAC results
C_0 = the mean of the original measurement results
R = the traceable reference control value
$\overline{X} =$ the mean results from measuring the controls
$f_{dilutor}$ = the correction factor for the dilutor
Equation 14 is a basic multiplicative model that includes four components of uncertainty and corrects for analytical bias.

4.3 Identify the components of uncertainty

From equation 14 we see four components that contribute to the combined uncertainty in the corrected BAC. These include: (1) the original duplicate measurement results of the blood alcohol concentration, (2) the reference value (R) representing a traceable unbiased control standard purchased from a commercial laboratory having a certificate of analysis, (3) the mean of the replicate measurements $\left(\overline{X}\right)$ of the traceable control and (4) the correction factor $\left(f_{dilutor}\right)$ for the dilutor used in preparing both the controls and blood samples before analysis. We will assume $f_{dilutor} = 1$.

4.4 Quantify the standard uncertainties for the components

For our example we will assume the values for the four parameters are those shown in Table 1. The uncertainty for the reference value (R) is a Type B uncertainty which comes from the certificate of analysis provided by the vendor preparing the control standard. The uncertainty for the replicate measurements of the control standard is simply the standard deviation determined from n=8 measurements of the control standard. The uncertainty for the dilutor was determined from the certificate of analysis. Since the dilutor is designed to provide 10 ml volume we see a small bias exists. This is not corrected for since the same bias would influence both the control standard measurements as well as the blood samples. For this reason we assume $f_{dilutor} = 1$. The actual value of the $f_{dilutor}$ in table 1 (10.15ml), however, will be used to estimate its uncertainty. The uncertainty associated with the blood alcohol results reported in table 1 (0.00072 g/dL) requires some further explanation. The uncertainty associated with these BAC results represents total method uncertainty. This estimate will be determined from a large number of duplicate BAC results generated within the same laboratory over a long period of time (approximately one year). This would include variation from sample preparation, multiple instruments, multiple calibrations, multiple analysts, multiple uses of the dilutor and time. Figure 5 illustrates an uncertainty function generated from duplicate blood alcohol data analyzed in the forensic laboratory of New Zealand. (Stowell, e.tal., 2008) For illustration purposes, we will assume this model is relevant to our example. Each point in the plot represents the standard deviation associated with a single determination and is generated from the following equation for a pooled estimate:

$$u_B = \sqrt{\frac{\sum_{i=1}^{k} d_i^2}{2k}} \qquad (15)$$

where: u_B = the standard deviation for a single measurement of blood alcohol concentration
d_i = the difference between duplicate results for the ith sample
k = the total number of duplicate samples within the bin

Duplicate results are pooled into bins of 0.010 g/dL to generate the uncertainty estimates throughout the concentration range. The result is an estimate of the uncertainty as a function of concentration and reveals the general increase in variation with concentration. Some would advocate the use of a characteristic function rather than an uncertainty function. (Thompson and Coles, 2011) A characteristic function is generated from regressing the variance against the concentration squared. Before estimating our method uncertainty from these functions, we need to determine our corrected BAC result. This is done as follows:

$$C_{corr} = \frac{C_0 \, R}{\overline{X}} \cdot f_{dilutor} = \frac{(0.0815)(0.100)}{(0.0986)} \cdot 1 = 0.0827 \, g \, / \, dL \qquad (16)$$

We now use this corrected result to estimate our method uncertainty from the model in figure 5. Based on the linear uncertainty function in figure 5 we obtain a method uncertainty of 0.00076 g/dL. Developing the characteristic function for the same data set yields a method uncertainty estimate of 0.00072 g/dL. Therefore, we will use the value of 0.00072 g/dL for example, as we see in table 1.

Parameter	Values	Type	Standard Uncertainty	n
C_0	(0.082, 0.081 g/dL)0.0815 g/dL	A	0.00072g/dL	2
R	0.100 g/dL	B	0.0004 g/dL	1
$\overline{X}$	0.0986 g/dL	A	0.0008 g/dL	8
$f_{dilutor}$	10.15 ml	B	0.050 ml	10

Table 1. Estimates, standard uncertainties and the number of measurements for the four parameters assumed to contribute to the combined uncertainty of blood alcohol measurement
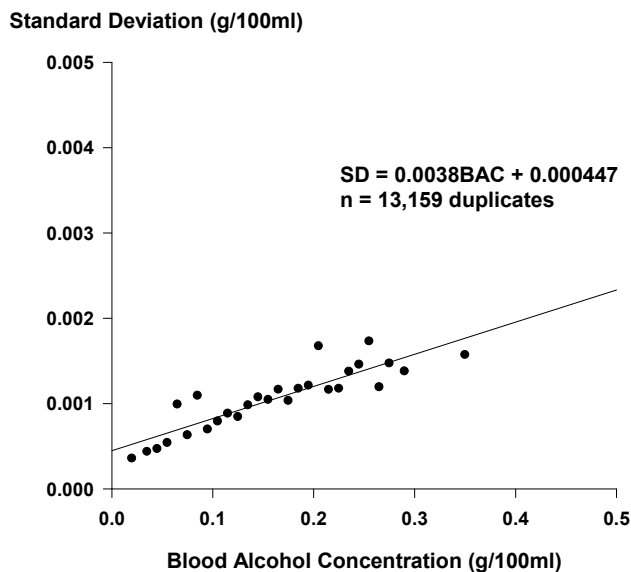

Fig. 5. An uncertainty function plotting pooled standard deviation estimates against their concentration determined from a large number of duplicate blood alcohol results

4.5 Combine the standard uncertainties and compute the combined uncertainty

We first determine our combined uncertainty using the general method of error propagation found in equation 8 assuming independence amongst the predictor variables. Putting our values determined from equation 16 into equation 8 we obtain equation 17. Since our measurement function is multiplicative we also estimate our combined uncertainty using equation 6 and assuming independence we obtain equation 18. Notice that we have included the actual estimate for $f_{dilutor}$ of 10.15 ml. This will ensure the appropriate value is determined for the uncertainty of the dilutor component. For purposes of bias correction in the measurement function of equation 14, however, we assume the value of $f_{dilutor} = 1.0$. From equations 17 and 18 we see that both the RSS method of equation 6 and the error propagation method of equation 8 yield nearly identical results.

$$u_{C_{Corr}} = \sqrt{\left[\frac{R}{\overline{X}}\cdot f_{dilutor}\right]^2 u_{C_0}^2 + \left[\frac{C_0}{\overline{X}}\cdot f_{dilutor}\right]^2 u_R^2 + \left[-\frac{C_0 R}{\overline{X}^2}\right]^2 u_{\overline{X}}^2 + \left[\frac{C_0 R}{\overline{X}}\right]^2 u_{f_{dilutor}}^2}$$

$$u_{C_{Corr}} = \sqrt{\left[\frac{0.100}{0.0986}\cdot 1\right]^2 \left[\frac{0.00072}{2}\right]^2 + \left[\frac{0.0815}{0.0986}\cdot 1\right]^2 \left[\frac{0.0004}{1}\right]^2 + \left[-\frac{(0.0815)(0.100)}{(0.0986)^2}\cdot 1\right]^2 \left[\frac{0.0008}{8}\right]^2 + \left[\frac{(0.0815)(0.100)}{0.0986}\right]^2 \left[\frac{0.050}{10}\right]^2} \quad (17)$$

$$u_{C_{Corr}} = 0.00065\, g\,/\,dL$$

$$\frac{u_{\overline{Y}}}{\overline{Y}} = \sqrt{\left[\frac{\frac{u_{C_0}}{\sqrt{n}}}{C_0}\right]^2 + \left[\frac{\frac{u_R}{\sqrt{n}}}{R}\right]^2 + \left[\frac{\frac{u_{\overline{X}}}{\sqrt{n}}}{\overline{X}}\right]^2 + \left[\frac{\frac{u_{f_{dilutor}}}{\sqrt{n}}}{f_{dilutor}}\right]^2}$$

$$\frac{u_{\overline{Y}}}{0.0827} = \sqrt{\left[\frac{\frac{0.00072}{\sqrt{2}}}{0.0815}\right]^2 + \left[\frac{\frac{0.0004}{\sqrt{1}}}{0.100}\right]^2 + \left[\frac{\frac{0.0008}{\sqrt{8}}}{0.0986}\right]^2 + \left[\frac{\frac{0.050}{\sqrt{10}}}{10.15}\right]^2} \quad (18)$$

$$u_{\overline{Y}} = 0.0827\,(0.0081) = 0.00067\, g\,/\,dL$$

4.6 Compute the expanded uncertainty and uncertainty interval

The expanded uncertainty is denoted by the value U and is determined from: $U = k u_C$ where k = a coverage factor and u_C = the combined uncertainty. The expanded uncertainty is then used to generate an uncertainty interval as

$$\overline{Y} \pm k u_c \quad \Rightarrow \quad \overline{Y} \pm U \quad (19)$$

where: $\overline{Y}$ = the unbiased mean measurement result, k = the coverage factor and U = the expanded uncertainty. Notice that u_C is actually the standard deviation of the mean. This results from the fact that we included the appropriate sample sizes, where available, for each term in equations 17 and 18. Sample size also determines degrees of freedom and whether the normal distribution can be assumed or if the t-distribution should be employed. Sample size should be determined as part of the measurement design to ensure sufficient quality control and statistical power. Coverage factors of k=2 or k=3 are common and represent approximately 95% and 99% uncertainty intervals respectively. Selecting k=2 or 3 assumes large degrees of freedom (sample size ≥ 30). Sample sizes less than 30 should employ the Students t distribution. From table 1 we see that none of the sample sizes exceed ten. However, we could argue that the method uncertainty associated with the duplicate blood alcohol results (0.00072 g/dL), determined from the data in figure 5, was generated from over 11,000 duplicate blood alcohol results. This should clearly justify the use of k=2 or 3 for approximate estimates of the 95% and 99% expanded uncertainty intervals. For our present example, however, we will assume we have the limited number of observations noted in table 1 and illustrate the calculation of what is called the "effective degrees of freedom", which may be necessary in some forensic contexts. For this purpose we employ the Welch-Satterthwaite equation which assumes the estimation of the effective degrees of

freedom for a probability distribution formed from several independent normal distributions as in equation 20. (Ballico, 2000, Kirkup and Frenkel, 2006)

$$v_{eff} = \frac{u_C^4}{\sum_{i=1}^{k} \frac{u_i^4}{v_i}} \tag{20}$$

where: v_{eff} = the effective degrees of freedom

u_C^4 = the combined uncertainty

u_i^4 = the uncertainty associated with the ith component

k = the number of components contributing to the combined uncertainty

The uncertainty terms $\left(u_i^4\right)$ can be determined either from the coefficients of variation (CV) or from partial derivatives determined from the measurement function in equation 14. If the CV estimates are used we do not incorporate the sample size n for each term. We will determine the CV estimates for our example. We first compute the combined uncertainty again as in equation 21.

$$\frac{u_C}{C_{Corr}} = \sqrt{\left[\frac{u_{C_0}}{C_0}\right]^2 + \left[\frac{u_R}{R}\right]^2 + \left[\frac{u_{\overline{X}}}{\overline{X}}\right]^2 + \left[\frac{u_{f_{dilutor}}}{f_{dilutor}}\right]^2} \tag{21}$$

$$\frac{u_C}{0.0827} = \sqrt{\left[\frac{0.00072}{0.0815}\right]^2 + \left[\frac{0.0004}{0.100}\right]^2 + \left[\frac{0.0008}{0.0986}\right]^2 + \left[\frac{0.050}{10.15}\right]^2} = 0.0011\, g\,/\,dL$$

Next, we incorporate these results into equation 20 as follows:

$$v_{eff} = \frac{u_C^4}{\sum_{i=1}^{k} \frac{u_i^4}{v_i}} = \frac{\left[\frac{0.0011}{0.0827}\right]^4}{\frac{\left[\frac{0.00072}{0.0815}\right]^4}{1} + \frac{\left[\frac{0.0004}{0.100}\right]^4}{\infty} + \frac{\left[\frac{0.0008}{0.0986}\right]^4}{7} + \frac{\left[\frac{0.050}{10.15}\right]^4}{9}} = 4.6 \approx 4$$

From this computation we see that the effective degrees of freedom can be some non-integer value, in which case the value is generally truncated. Notice also that the uncertainty associated with the reference value (R) has an infinite number of degrees of freedom. This is because it is a Type B uncertainty determined from a certificate of analysis where we assume the uncertainty in the uncertainty estimate (0.0004 g/dL) is zero with correspondingly large degrees of freedom. As a result this term disappears from the computation. Each of the other degrees of freedom is determined from n-1. From these results we would estimate our value from the t-distribution to be: $t_{0.975,4} = 2.776$ for estimating a 95% uncertainty interval. Using these results along with our combined uncertainty determined from equation 18 we would obtain a 95% uncertainty interval of: $\overline{Y} \pm k\,u_C \Rightarrow 0.0827 \pm 2.776(0.00067) \Rightarrow 0.0827 \pm 0.0019 \Rightarrow 0.0808\ to\ 0.0846\, g\,/\,dL$.

We now have an interval within which we would expect a large fraction (approximately 95%) of the expected values of the measurand to exist. If we were to assume k=2 to generate an approximate 95% uncertainty interval we would obtain: $0.0827 \pm 2(0.00067) \Rightarrow 0.0827 \pm 0.0013 \Rightarrow 0.0814$ to $0.0840 \, g/dL$. We see that this interval is slightly narrower than that employing the effective degrees of freedom estimate. Choosing the appropriate coverage factor will be a decision made within each forensic laboratory. A 99% interval (k=3) will provide a higher degree of confidence that may be important in forensic applications. This is particularly true where results are near prohibited legal limits. Whatever decision is made, the value for k should be clearly identified in the program policy or SOP manuals and strictly adhered to in practice. In this example we have assumed our expanded interval to be an "uncertainty interval" rather than a "confidence interval". The *GUM* document prefers the term "uncertainty interval" or "level of confidence". (ISO/GUM, 2008) Others, however, interpret U as representing a confidence interval which has a specific definition in the classical statistical sense.

4.7 Produce the uncertainty budget

Table 2 illustrates one form of an uncertainty budget for our example. The uncertainty budget lists the components contributing to the combined uncertainty along with the percent of their contribution to the total. The percent contributions were determined from the terms under the radical sign in equation 18. This is very useful for identifying which components are the major contributors and which may be reasonably ignored. The *GUM* document states that any contributions less than one-third of the largest contributor can be safely ignored. (ISO/GUM, 2008) Based on this we see that the analytical and dilutor components could be safely ignored in this example. However, from a forensic perspective it may be better to include all components considered, providing full disclosure. We see that the total method contributes the largest component at 59%. This is expected because of all of the contributing sub-components involved: analysts, calibrations, time, dilutions, etc. This analysis does not include, however, the venous blood sampling performed by the phlebotomist who typically performs only one venipuncture. Moreover, many laboratories do not even consider sampling as a component of their combined uncertainty. They simply consider their uncertainty estimates corresponding to the sample "as received in the laboratory". Jones, for example, has considered sampling as a source of uncertainty in some of his published work. (Jones, 1989)

Source	Type	Distribution	Standard Uncertainty	Percent[1]
Traceability	B	Normal	0.0004 g/dL	24%
Analyical	A	Normal	0.0008 g/dL	13%
Dilutor	B	Normal	0.050 ml	4%
Total Method	A	Normal	0.00072 g/dL	59%
Combined Uncertainty			0.00067 g/dL	
Expanded Uncertainty (k=2.776)			0.0019 g/dL	
95% confidence interval			0.0808 to 0. 0846 g/dL	

[1]Percent of contribution to total combined uncertainty

Table 2. Uncertainty budget for the illustrated example

4.8 Report the results

One of the most important, yet often overlooked, elements of determining measurement uncertainty is reporting the results. A great deal of thought should be given to this aspect of measurement. The end-user should be consulted to determine exactly what is needed for their application. There should be sufficient information so the results and their associated uncertainty are fully interpretable and unequivocal for a specific application without reference to additional documentation. This will necessitate some textual explanation in addition to the numerical results. One possibility for our blood alcohol example above is:

The duplicate whole blood alcohol results were 0.082 and 0.081 g/dL with a corrected mean result of 0.0827 g/dL. An expanded combined uncertainty of 0.0019 g/dL assuming a coverage factor of k=2.776 with an effective degrees-of-freedom of 4 and a normal distribution was generated from four principle components contributing to the uncertainty. An approximate 95% confidence interval for the true mean blood alcohol concentration is 0.0808 to 0.0846 g/dL.

In addition to the statement, a figure similar to that of figure 3 could be provided which might assist the court in placing the results in some geometric perspective. The format for reporting the results should be considered flexible. As time goes on there will no doubt be the need for revision to ensure clarity in communication and interpretation.

4.9 Assumptions of this approach

There were a number of assumptions employed in estimating the uncertainty illustrated above. The customer should appreciate these assumptions to allow for full and clear interpretation. Very generally, the assumptions are:

1. The blood alcohol measurement results are normally distributed
2. All standard uncertainties are valid estimates
3. The method uncertainty is probably over estimated due to some "double counting"
4. The method of confidence interval estimation will be robust
5. With a fixed mean (μ), 95% of the intervals will bracket μ
6. The confidence interval expresses the uncertainty due to sampling variability only
7. This entire approach to estimating the uncertainty is uncertain.
8. We have assumed that all uncertainty components are independent

We would not advocate that these assumptions be listed as part of the reported results. Rather, they should be available if requested by the end-user and toxicologists should be prepared to discuss them.

5. Breath alcohol measurement example

Our next example illustrates the uncertainty estimation for a breath alcohol measurement. We will assume the following measurement function which was presented earlier as equation 12:

$$\overline{Y}_{Corr} = \frac{\overline{Y}_0 \cdot GC_{Sol} \cdot R}{\overline{X} \cdot K \cdot GC_{Cont}} \tag{22}$$

where: $\overline{Y}_0 = $ the mean of the original n measurements

$GC_{Sol} = $ the mean of the simulator solution measurements by gas chromatography

$R = $ the traceable reference value

$\overline{X}$ = mean of the breath test instrument measuring the simulator solution heated to 34^0C

K = 1.23 the ratio of partition coefficients

GC_{Cont} = the mean results from measuring the traceable controls on the gas chromatograph

For this example we assume that simulator solutions are prepared and tested by gas chromatography within the toxicology laboratory. Commercially purchased standards (CRM) are used as calibrators and controls on the gas chromatograph. Certificates of analysis are used as Type B uncertainties to establish the traceability. For this example we will assume the following data are available for the six components of equation 22: Duplicate BrAC results: 0.081 and 0.085 g/210L, $\overline{Y}_0$ = 0.0830 g/210L, GC_{Sol} : mean = 0.0985 g/dL u = 0.0007 g/dL n=15, R = 0.100 g/dL u = 0.0003 g/dL, $\overline{X}$: mean = 0.0795 g/210L u = 0.0012 g/210L n=10, K = 1.23 u = 0.012 and GC_{Cont} : mean = 0.1015 g/dL u = 0.0006 g/dL n=28. We begin by computing the corrected mean BrAC results according to:

$$\overline{Y}_{Corr} = \frac{(0.0830\,g\,/\,210L)(0.0985\,g\,/\,dL)(0.100\,g\,/\,dL)}{(0.0795\,g\,/\,210L)(1.23)(0.1015\,g\,/\,dL)} = 0.0824\,g\,/\,210L$$

The estimate for the uncertainty in $\overline{Y}_0$ will come from an uncertainty function seen in figure 6 and developed from a large number of duplicate breath alcohol tests using equation 15. The total method uncertainty for our example determined from the linear model in figure 6 and using the corrected mean BrAC of 0.0824 g/210L is 0.0031 g/210L. Since our model in equation 22 is multiplicative we employ the RSS for the CV values and assume independence amongst all components. The combined uncertainty estimate is seen in equation 23. Next we estimate the 95% uncertainty interval and obtain:

$$\overline{Y} \pm k u_C \quad \Rightarrow \quad \overline{Y} \pm U \Rightarrow \quad 0.0824 \pm 2(0.00239) \quad \Rightarrow \quad 0.0824 \pm 0.0048$$
$$0.0776\,to\,0.0872\,g\,/\,210L$$

Since the n for estimating the uncertainty function in figure 6 was very large, we assume an infinite degrees of freedom and use k=2 for estimating an approximate 95% confidence interval. Table 3 shows the uncertainty budget for this analysis. From the uncertainty budget we see that the total method accounted for the majority of the combined uncertainty (84%). This is not surprising since the breath sampling component, contained within the total method uncertainty function of figure 6, has significant variation. The budget also shows that the reference traceability, the GC measurement of the controls and the GC measurement of the simulator solution all provide 1% or less to the combined uncertainty. They could reasonably be ignored in this example. We now report our results as follows:

The duplicate breath alcohol results were 0.081 and 0.085 g/210L with a corrected mean result of 0.0824 g/210L. An expanded combined uncertainty of 0.0048g/210L assuming a coverage factor of k=2 with an infinite number of degrees-of-freedom and a normal distribution was generated from six principle components contributing to the uncertainty. An approximate 95% confidence interval for the true mean breath alcohol concentration is 0.0776 to 0.0872 g/210L.

$$\frac{u_{\overline{Y}}}{\overline{Y}} = \sqrt{CV_{\overline{Y}_0}^2 + CV_{GC_{Sol}}^2 + CV_R^2 + CV_{\overline{X}}^2 + CV_K^2 + CV_{GC_{Cont}}^2}$$

$$\frac{u_{\overline{Y}}}{\overline{Y}_{Corr}} = \sqrt{\left[\frac{\frac{u_{\overline{Y}_0}}{\sqrt{n}}}{\overline{Y}_0}\right]^2 + \left[\frac{\frac{u_{GC_{Sol}}}{\sqrt{n}}}{GC_{Sol}}\right]^2 + \left[\frac{\frac{u_R}{\sqrt{n}}}{R}\right]^2 + \left[\frac{\frac{u_{\overline{X}}}{\sqrt{n}}}{\overline{X}}\right]^2 + \left[\frac{\frac{u_K}{\sqrt{n}}}{K}\right]^2 + \left[\frac{\frac{u_{GC_{Cont}}}{\sqrt{n}}}{GC_{Cont}}\right]^2}$$

(23)

$$\frac{u_{\overline{Y}}}{0.0824} = \sqrt{\left[\frac{\frac{0.0031}{\sqrt{2}}}{0.0824}\right]^2 + \left[\frac{\frac{0.0007}{\sqrt{15}}}{0.0985}\right]^2 + \left[\frac{\frac{0.0003}{\sqrt{1}}}{0.100}\right]^2 + \left[\frac{\frac{0.0012}{\sqrt{10}}}{0.0795}\right]^2 + \left[\frac{\frac{0.012}{\sqrt{1}}}{1.23}\right]^2 + \left[\frac{\frac{0.0006}{\sqrt{28}}}{0.1015}\right]^2}$$

$$u_{\overline{Y}} = 0.0824(0.0290) = 0.00239 \, g / 210L$$

The approximate 95% uncertainty interval estimated for this example shows that the lower limit falls below the critical legal driving level of 0.080 g/210L. We may be interested in knowing the probability that the true population mean BrAC is above 0.080 g/210L. This can be estimated by first considering our confidence interval in the following form:

$$P\left[\overline{Y} - Z_{(1-\alpha/2)} S_{\overline{Y}} \leq \mu \leq \overline{Y} + Z_{(1-\alpha/2)} S_{\overline{Y}}\right] = \pi \qquad (24)$$

Since we are interested in determining the probability that μ exceeds the lower limit we rewrite equation 24 as follows:

$$P\left[\overline{Y} - Z_{(1-\alpha/2)} S_{\overline{Y}} \leq \mu \leq \infty\right] = \pi \qquad (25)$$

We set the lower limit expressed in equation 25 equal to 0.080 g/210L and solve for $Z_{(1-\alpha/2)}$:

$$\overline{Y} - Z_{(1-\alpha/2)} S_{\overline{Y}} = 0.080 \Rightarrow 0.0824 - Z_{(1-\alpha/2)}(0.00239) = 0.080 \Rightarrow Z_{(1-\alpha/2)} = 1.0$$

Standard Deviation (g/210L)

Breath Alcohol
SD = 0.0260BrAC + 0.00095
n = 27,995 duplicates

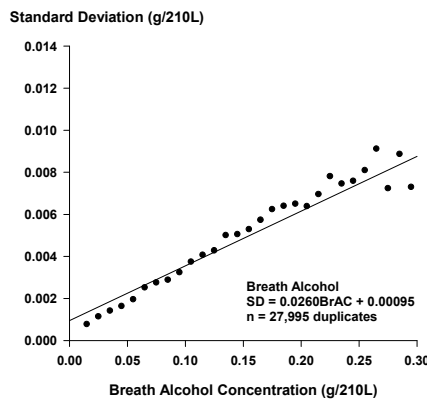**Breath Alcohol Concentration (g/210L)**

Fig. 6. An uncertainty function plotting pooled standard deviation estimates against their concentration determined from a large number of duplicate breath alcohol results

Source	Type	Distribution	Standard Uncertainty	Percent[1]
Total Method	A	Normal	0.0031 g/210L	84%
GC Solution	A	Normal	0.0007 g/dL	0.5%
Reference	B	Normal	0.0003 g/dL	1%
Breath Instrument	A	Normal	0.0012 g/210L	3%
Simulator Part. Coef	B	Normal	0.012	11%
GC Controls	A	Normal	0.0006 g/210L	0.5%
Combined Uncertainty			0.00239 g/210L	
Expanded Uncertainty (k=2)			0.0048 g/210L	
95% confidence interval			0.0776 to 0. 0872 g/210L	

[1] Percent of contribution to total combined uncertainty

Table 3. Uncertainty budget for the illustrated breath alcohol example

Next, we rearrange our probability statement, introduce the value for $Z_{(1-\alpha/2)}$, and refer to the standard normal tables:

$$P\left[\overline{Y} - Z_{1-\alpha/2}\, S_{\overline{Y}} \le \mu\right] = P\left[\frac{\overline{Y} - \mu}{S_{\overline{Y}}} \le Z_{1-\alpha/2}\right] = P\left[Z \le Z_{1-\alpha/2}\right] = P\left[Z \le 1.0\right] = 0.8413$$

There is a probability of 0.8413 that the individual's true mean BrAC exceeds 0.080 g/100ml. This may or may not rise to the level of proof beyond a reasonable doubt, depending on the opinion of the court. This example illustrated the use of simulator control standards produced within a local toxicology laboratory including their associated uncertainties. Some jurisdictions, however, choose to purchase simulator control standards rather than prepare their own. If that were the case in this example, we could have eliminated the GC solutions and GC controls from our uncertainty estimates. The simulator partition coefficient would have remained while the reference value would have been obtained from the certificate of analysis from the manufacturer and considered a Type B uncertainty. Therefore, rather than having to include the GC solution and GC control components separately in the combined uncertainty estimate, they should already be included within the manufacturer's estimate of combined uncertainty, depending, of course, on how the solution standards were prepared and tested.

6. Dealing with measurement bias

Our principle objective here will be to illustrate several ways for treating uncorrected bias. Bias or systematic error is common in all measurements. Some consider different types of bias such as: (1) method bias, (2) laboratory bias and (3) run bias. (O'Donnell and Hibbert, 2005) Not all, however, would agree with the need for classifications of bias. (Kadis, 2007, O'Donnell and Hibbert, 2007) Regardless of its classification or source, all forms of bias should ideally be determined and corrected for employing traceable control standards. As this is done, the uncertainty of that correction must be included as one of the components in the combined uncertainty. Occasionally, the analyst may determine that the bias is small

and insignificant and not correct for it. There are ways to handle uncorrected bias as well by adding an additional component to the combined uncertainty. We will consider some examples here. Estimations for bias can come from internal quality control, proficiency test data, collaborative studies or method validation data. (Kane, 1997)

6.1 Preparing an alcohol in water control solution

We will assume in this example that we desire to prepare an ethanol in water solution to be used as a control standard. We want to prepare this solution to have a concentration of approximately 0.10 g/dL. Our measurement function will be as follows: (Philipp et.al., 2010)

$$C = \frac{m_{Etoh} \, P D}{m_{Solution}} \tag{26}$$

where: C = the concentration of ethanol in water
m_{Etoh} = the mass measurement of ethanol
P = the purity of the ethanol
D = the density of the ethanol
$m_{Solution}$ = the mass measurement of the combined solution of ethanol and water

Preparing a control standard gravimetrically has advantages. (Gates, et.al., 2009) There is better traceability for the mass measurements and no concern regarding the uncertainty in volume measurements. We will assume the purity (P) to be 0.995 with a Type B standard uncertainty of 0.002 determined from the certificate of analysis. We further assume that the density (D) of the solution is 0.997 g/ml (OIML, King and Lawn, 1999) with a Type B standard uncertainty of 0.00054 g/ml (King and Lawn, 1999), determined from the certificate of analysis from the manufacturer of a density meter. For both the purity and the density we will assume the uniform distribution in order to estimate their standard uncertainties. The values for the density are obtained from published tables for ethanol/water solutions. The density of the solution will be a function of the mass fraction of ethanol. The higher the mass fraction of ethanol the closer the density will be to 0.789 g/ml - the density of pure ethanol. The lower the mass fraction of ethanol the closer the density will be to 1.00 g/ml - the density of water. Since the density of the solution depends on the mass fraction of ethanol and we have selected a density of 0.997 g/ml (corresponding to a mass fraction of approximately 0.101%) and we desire a total solution mass of 1800 g, we need to have the mass of ethanol equal to 1.82 g. We will need to weigh 1.82 g of ethanol and place it into solution with water and add water until we have a total mass of 1800g. We will assume that the total solution mass is weighed on a scale that has had replicate measurements (n=30) of a 2 Kg traceable check weight (Type B uncertainty of 0.016 Kg) with a mean result of 1,940 g and a standard uncertainty of 30 g. This will be used to estimate the standard uncertainty in the measurement of $m_{Solution}$. We now recognize that there is a bias in the weighing of the total solution. The measured mass of the solution is low by 3.0%. This will affect the mass of the ethanol necessary to maintain the density of 0.997 and mass fraction of 0.101%. As a result the mass of the ethanol will need to be 1.87 g. The mass of ethanol was weighed on a different scale that also has a set of replicate measurements (n=23) of a 2.0 g traceable check weight (Type B standard uncertainty of 0.014g) with a mean result of 2.08 g and a standard uncertainty of 0.02g. This scale has a bias of +4.0%. We now incorporate our assumed measurement information into equation 26:

$$C = \frac{m_{Etoh}\, PD}{m_{Solution} \cdot \dfrac{R_{2Kg}}{\overline{X}}} = \frac{(1.87\,g)(0.995)(0.997\,g\,/\,ml)}{(1800\,g)\left(\dfrac{2000\,g}{1940\,g}\right)} = 0.00100\,g\,/\,ml = 0.1000\,g\,/\,dL \quad (27)$$

where: $\dfrac{R_{2Kg}}{\overline{X}} = $ the correction factor for the bias in the scale used to weigh the total solution

Notice that we only correct for the bias in the scale used to weigh the total solution but not for the scale used to weigh the ethanol. The question now is how to deal with the +4.0% bias in the one scale. We begin by estimating the combined uncertainty ignoring the bias (assuming it is zero) and assuming independence of all variables. Since equation 27 is a multiplicative model we employ the RSS of the CV's squared as in equation 28. Notice that the standard uncertainty in the solution mass measurement comes from the repeatability measurements of the 2.0 Kg traceable check standards. There is no separate uncertainty estimate for the single measurement of the total solution of 1800 g. Employing the Welch-Sattherwaite equation to compute the effective degrees of freedom for our example we obtain:

$$\nu_{eff} = \frac{\left[\dfrac{0.00210}{0.1000}\right]^4}{\dfrac{\left[\dfrac{.02}{2.08}\right]^4}{22} + \dfrac{\left[\dfrac{.002}{0.995}\right]^4}{\infty} + \dfrac{\left[\dfrac{0.00054}{0.997}\right]^4}{\infty} + \dfrac{\left[\dfrac{30}{1800}\right]^4}{29} + \dfrac{\left[\dfrac{16}{2000}\right]^4}{\infty}} = 63.8 \approx 63$$

$$\frac{u_C}{C} = \sqrt{CV_{m_{Etoh}}^2 + CV_P^2 + CV_D^2 + CV_{m_{Sol}}^2 + CV_{R_{2Kg}}^2}$$

$$\frac{u_C}{C} = \sqrt{\left[\dfrac{\dfrac{u_{m_{Etoh}}}{\sqrt{n}}}{m_{Etoh}}\right]^2 + \left[\dfrac{\dfrac{u_P}{\sqrt{n}}}{P}\right]^2 + \left[\dfrac{\dfrac{u_D}{\sqrt{n}}}{D}\right]^2 + \left[\dfrac{\dfrac{u_{m_{Sol}}}{\sqrt{n}}}{m_{Sol}}\right]^2 + \left[\dfrac{\dfrac{u_{R_{2Kg}}}{\sqrt{n}}}{R_{2Kg}}\right]^2}$$

$$(28)$$

$$\frac{u_{\overline{Y}}}{0.1000} = \sqrt{\left[\dfrac{\dfrac{0.02}{\sqrt{23}}}{2.08}\right]^2 + \left[\dfrac{\dfrac{0.002}{\sqrt{3}}}{0.995}\right]^2 + \left[\dfrac{\dfrac{0.00054}{\sqrt{3}}}{0.997}\right]^2 + \left[\dfrac{\dfrac{30}{\sqrt{30}}}{1800}\right]^2 + \left[\dfrac{\dfrac{16}{\sqrt{1}}}{2000}\right]^2}$$

$$u_{\overline{Y}} = 0.1000(0.0089) = 0.00089\,g\,/\,dL$$

The 95% confidence interval for our estimated concentration would be:

$$\overline{Y} \pm t_{0.975,63}\, u_c \;\Rightarrow\; 0.1000 \pm 2.00(0.00089) \;\Rightarrow\; 0.1000 \pm 0.00178 \;\Rightarrow\; 0.0982\ to\ 0.1018$$

The next option for dealing with the bias in the mass measurement of the ethanol is to correct for it. This is always the recommended practice and consistent with the *GUM* document. Correcting the ethanol mass for the +4.0% bias yields a result of 1.80 g. Placing this corrected value into equation 27 yields a corrected concentration of 0.000962 g/ml or 0.0962 g/dL. Now we must account for the uncertainty in the 2.0g reference check weight by including its Type B uncertainty in equation 28 where we add the additional term:

$$\left[\frac{\frac{u_{R_{2g}}}{\sqrt{1}}}{R_{2g}}\right]^2 = \left[\frac{\frac{0.014}{\sqrt{1}}}{2.00}\right]^2$$ and, when including the corrected concentration, we obtain:

$u_C = 0.0986(0.0113) = 0.00111\,g\,/\,dL$. The uncertainty budget is shown in table 4 both when ignoring the bias and when including the bias correction. From table 4 we see that including the additional balance bias, the combined uncertainty increased by 25% and contributed 38% to the combined uncertainty. The bias, in this example, is clearly significant and as a result should be corrected for. Before illustrating our next approach to handling uncorrected bias, we will evaluate the bias in our example to determine its significance. To do so we employ the following t-test:

$$t = \frac{C-R}{\sqrt{u_C^2 + u_R^2}} = \frac{2.08 - 2.00}{\sqrt{\left[\frac{\frac{0.02}{\sqrt{23}}}{2.08}\right]^2 + \left[\frac{\frac{0.014}{\sqrt{1}}}{2.00}\right]^2}} = 10.9 \tag{29}$$

The critical value for a two-tailed test with $\alpha = 0.05$ and effective degrees of freedom of 51 from the t-distribution is $t_{0.975,51} = 2.01$. The results from equation 29 show the bias to be largely significant and should be corrected for. There are times when measurement bias is known to exist but is not corrected for. The analyst may believe the bias to be small and insignificant or it may be too complex to correct for. There are several methods that have been proposed for including the uncertainty due to uncorrected bias. (Maroto,et.al., 2002, Petersen, et.al., 2001) All of these effectively increase the expanded uncertainty by some amount to account for the uncorrected bias. Moreover, including an uncertainty component

Source	Type	Standard Uncertainty	Percent[1]	
			Ignoring Bias	Correcting Bias
Mass of Ethanol	A	0.02 g	5%	3%
Purity of Ethanol	B	0.002	1%	1%
Density of Solution	B	0.00054 g/ml	1%	1%
Mass of Solution	A	30 g	12%	7%
2.0 Kg Reference	B	16 g	81%	50%
2.0 g Reference	B	0.014 g		38%
Combined Uncertainty			0.00089 g/dL	0.00111 g/dL

[1] Percent of contribution to total combined uncertainty

Table 4. Uncertainty budget for the preparation of the control ethanol solution

resulting from a corrected bias is always less than the uncertainty component resulting from uncorrected bias. (Synek, 2005, Linsinger, 2008) One approach is to include the bias within the radical sign and estimate the expanded uncertainty (U) as follows:

$$U = kC\sqrt{CV^2_{m_{Etoh}} + CV^2_P + CV^2_D + CV^2_{m_{Sol}} + CV^2_{R_{2Kg}} + bias^2} \qquad (30)$$

Since all of the other terms within the radical sign are dimensionless relative variances, we must transform the bias into dimensionless relative units. Doing this with our example and assuming k=2 we obtain:

$$U = 2(0.1000)\sqrt{\left[\frac{\frac{0.02}{\sqrt{23}}}{2.08}\right]^2 + \left[\frac{\frac{0.002}{\sqrt{3}}}{0.995}\right]^2 + \left[\frac{\frac{0.00054}{\sqrt{3}}}{0.997}\right]^2 + \left[\frac{\frac{30}{\sqrt{30}}}{1800}\right]^2 + \left[\frac{\frac{16}{\sqrt{1}}}{2000}\right]^2 + \left[\frac{\frac{0.08}{\sqrt{23}}}{2.00}\right]^2}$$

$$U = 2(0.1000)(0.0122) = 0.0024\, g\,/\,dL$$

The combined uncertainty with this approach is 0.00122 g/dL compared to 0.00111 g/dL when correcting for the bias and 0.00089 g/dL when ignoring the bias. Another approach is to incorporate the coverage factor k into the radical sign but without effecting the bias term as follows:

$$U = C\sqrt{k^2\left[CV^2_{m_{Etoh}} + CV^2_P + CV^2_D + CV^2_{m_{Sol}} + CV^2_{R_{2Kg}}\right] + bias^2} \qquad (31)$$

With this approach the combined uncertainty remains the same but the expanded uncertainty becomes 0.00196 g/dL. As expected, this is slightly less than the expanded uncertainty determined from equation 30 which was 0.0024 g/dL. A third approach is basically the same as correcting for the bias and is expressed as:

$$Y \pm U + bias \quad \Rightarrow \quad \overline{y} - (U + bias) \le Y \le \overline{y} + (U - bias) \qquad (32)$$

For our example, the bias in the mass of the ethanol was +0.08g. The corrected mass of the ethanol should be 1.79 g rather than the 1.87 g value measured. Using the correct value of 1.79 g, the corrected concentration of the ethanol should be 0.0957 g/dL. This indicates that we have a bias in the estimated concentration of +0.0043 g/dL. Using this value for our bias and assuming an approximate 95% confidence interval, equation 32 becomes:

$$0.1000 - (2(0.00089) + 0.0043) \le Y \le 0.1000 + (2(0.00089) - 0.0043)$$
$$0.0939 \le Y \le 0.0975\, g\,/\,dL$$

Notice that this interval is not symmetric around our estimated, yet biased, concentration of 0.1000 g/dL. Instead, it has accounted for the +0.0043 g/dL bias and adjusted for this.

The next proposal for handling uncorrected bias is to simply add the absolute value of the bias to the expanded uncertainty as: $\overline{Y} \pm U + |bias|$. For our example this would result in:

$$\overline{Y}-\left(U+\left|bias\right|\right)\leq Y\leq\overline{Y}+\left(U+\left|bias\right|\right) \;\Rightarrow\; 0.1000-0.00608\leq Y\leq0.1000+0.00608$$
$$0.0939\leq Y\leq0.1061\,g\,/\,dL$$

This clearly would yield the largest uncertainty interval compared to the preceding methods and is probably larger than necessary. The final method we will consider yields an expanded uncertainty interval that is also asymmetric about the measurement result. (Phillips, et.al., 1997) This method computes the confidence interval based on the expanded uncertainty (U) estimated as follows:

$$\overline{Y}-U_{-}\leq Y\leq\overline{Y}+U_{+}$$

$$where: \qquad U_{+}=\begin{cases} ku_c-bias & if\ ku_c-bias>0 \\ 0 & if\ ku_c-bias\leq0 \end{cases} \qquad (33)$$

$$and \quad U_{-}=\begin{cases} ku_c+bias & if\ ku_c+bias>0 \\ 0 & if\ ku_c+bias\leq0 \end{cases}$$

Using this approach for our example would yield:

$$0.1000-\left[2\left(0.00089\right)+0.0043\right]\leq Y\leq0.1000+0 \;\Rightarrow\; 0.0939\leq Y\leq0.1000\,.$$

The asymmetry with this method has accounted for the positive bias and yields the same lower limit as the two preceding methods above. This results from the fact that our estimate is biased high by +0.0043 g/dL and was not corrected for. This last approach has more desirable statistical properties compared to the previous methods and has the advantage of avoiding negative expanded uncertainty limits (where the lower limit is below zero) which could occur at low concentrations. (Phillips, et.al., 1997)

6.2 Estimating bias by recovery

Another approach to estimating and handling bias is with recovery analysis. (Thompson, et.al., 1999) Recovery is the ratio, expressed as a percent, of the measurement result to the reference or true measurand value described by:

$$\%R=\left[\frac{C_0}{C_{Ref}}\right]100 \qquad (34)$$

where: %R = percent recovery
C_0 = the measured value
C_{Ref} = the true value of the measurand
Percent recovery is a metric more commonly applied in analytical contexts involving complex matrices with several steps of extraction, sample preparation and analysis of a specified sub-sample. The requirements of this complex procedure for extraction and analysis often results in a loss of the analyte prior to its actual quantitative determination. Hence, we have the concept of %Recovery. The accuracy of the analytical method is determined by its ability to quantify (recover) the full amount of the analyte in the original

matrix. Simply spiking alcohol in a blood sample and measuring it is not a typical application of percent recovery. The recovery is often determined during the method validation phase where a known blank matrix is spiked with a known mass of the relevant analyte. This is often referred to as a "reference recovery" or a "method recovery". (Barwick and Ellison, 1999) When recovery estimates are applied to correct subsequent samples, it is very important that the concentrations and matrix are appropriately similar and that the same full analytical protocol is followed. Measurements of recovery from several spiked samples may be performed with the mean and standard deviation of the percent estimates determined, providing uncertainty estimates for the percent recovery in future measurements. The fractional recovery can be employed as a correction factor in the measurement equation as follows:

$$C_{Corr} = \frac{C_0}{\overline{R}} \tag{35}$$

where: C_{Corr} = the corrected analytical result
C_0 = the original measurement
$\overline{R} =$ the mean fractional recovery

Assume that we are interested in determining the percent recovery of a specific drug for a particular analytical method. Assume that we have two vials of a subject's blood, each containing 1.0 ml and each containing some unknown concentration of the drug of interest. To one tube we add 0.1ml of a known analyte standard having a concentration of 20mg/dL.

We have now added a concentration of: $\frac{20\,mg}{dL}\left[\frac{0.1\,ml}{0.1\,ml+1.0\,ml}\right] = 1.82\frac{mg}{dL}$. To the other tube

we simply add 0.1 ml of water. We now measure the concentration of the analyte in each tube in replicate (at least twice) and determine the means to be: Tube with added analyte: 10.8 mg/dL Tube with added water: 9.3 mg/dL. We now compute the percent recovery according to:

$$\% \operatorname{Re}\text{covery} = \left[\frac{Measured\ Difference}{Concentration\ Added}\right]\cdot 100 = \left[\frac{10.8mg\,/\,dL - 9.3mg\,/\,dL}{1.82mg\,/\,dL}\right]\cdot 100 = 82.4\% \tag{36}$$

Assume that we have done this recovery experiment during method validation using blood specimens spiked with the analyte and obtained a mean % recovery of $\overline{R} = 84\%$ with a standard uncertainty of 6% determined from 45 spiked samples. Assume further that we now have a suspect's blood sample and we wish to provide an unbiased estimate of the analyte's concentration using this recovery data. We determine the suspect's sample results to be C_0 = 15.4mg/dL with a standard uncertainty of 0.92mg/dL determined from n=56 measurements of past quality control data. We further assume there are no other significant sources of bias, other than that estimated by the %Recovery. First we could determine whether the mean recovery of 84% was significantly different from 1.0 or not with the following t-test:

$$t = \frac{\left|\overline{R} - 1\right|}{u_{\overline{R}}} = \frac{\left|0.84 - 1\right|}{0.06\,/\,\sqrt{45}} = -17.9 \tag{37}$$

The p-value for t = 17.9 with df=44 is <0.00001. We conclude that the mean recovery is very significantly different from 1.0. The recovery estimate should be used to correct the analytical results. Using our mean recovery to correct our analytical results yields: $C_{Corr} = \dfrac{C_0}{\overline{R}} = \dfrac{15.4}{0.84} = 18.3\,mg/dL$. The combined uncertainty in our corrected estimate can now be determined from the RSS method using the CV's squared since we have a multiplicative model and we assume independence according to:

$$\frac{u_C}{C_{Corr}} = \sqrt{CV_{C_0}^2 + CV_R^2} \;\Rightarrow\; \frac{u_C}{18.3} = \sqrt{\left[\frac{\frac{0.92}{\sqrt{56}}}{18.3}\right]^2 + \left[\frac{\frac{0.06}{\sqrt{45}}}{0.84}\right]^2} \;\Rightarrow\; u_C = (18.3)(0.0126) = 0.231\,mg/dL$$

This results in a relative combined uncertainty of approximately 1.3%. Moreover, the analytical component contributed 45% while the recovery component contributed 65% to the combined uncertainty. The same analysis can be done when spiking blank specimens with a known concentration of the analyte. If we added the same 0.1ml of 20mg/dL concentration to 1.0ml of blank specimen, and quantified the specimen with our analytical method and obtained 1.65 mg/dL, this would become the numerator in equation 36 and we would obtain a recovery estimate of:

$$\%\,Recovery = \left[\frac{Measured\ Concentration}{Concentration\ Added}\right]\cdot 100 = \left[\frac{1.65mg/dL}{1.82mg/dL}\right]\cdot 100 = 90.7\% \quad .$$

Both methods of spiking blank samples or spiking samples already containing the analyte are used in recovery studies. Moreover, it is important to remember with recovery studies the assumption that no other bias exists. We have briefly considered several ways that have been proposed to handle uncorrected bias. Ideally, bias should always be corrected for - even when statistically insignificant. When the bias is not corrected for, the combined uncertainty statement should include some additional component, thus increasing its magnitude, accounting for the uncorrected bias. Moreover, the customer should be made aware, either in the uncertainty statement or otherwise, when uncorrected bias exists and how it has been accounted for.

7. Uncertainty in post-mortem drug analysis

This example summarizes work recently published where methadone was measured in post-mortem cases. (Linnet, et.al., 2008) One sample of blood was taken from each femoral vein in 27 post-mortem autopsies. LC-MS/MS was the analytical method used to quantify both methadone and its main metabolite, 2-ethyl-1,5-dimethyl-3,3-diphenylpyrrolinium (EDDP). For our present example we will focus only on the quantitative measurement of methadone. While the study did not explicitly present a measurement function, the following would be a reasonable approximation:

$$C_{Corr} = \frac{C_0 \cdot C_{Cal}}{\overline{C}_A} = \frac{C_0 \cdot \dfrac{m_{Meth}\,P}{V}}{\overline{C}_A} \tag{38}$$

where: C_{Corr} = the corrected measurement of methadone

C_0 = the original quantitative measurement result of the methadone by LC-MS/MS

C_{Cal} = the reference calibration and/or control value

$\overline{C}_A$ = the mean quantitative measurement of the reference value

m_{Meth} = mass of the reference methadone added to the calibration/control solution

P = the purity of the methadone

V = the volume of the calibration/control methadone solution

The study also presented the following uncertainty estimates, expressed as %CV's, for each of the components in equation 38: $u_{\overline{C}_A} = 3.65\%$ $u_P = 0.29\%$ $u_{m_{Meth}} = 0.53\%$ $u_V = 0.05\%$.

The uncertainty in the purity was determined from employing the uniform distribution and the manufacturer's certificate of analysis stating the purity was 99.99% ± 0.5%. The uncertainty in the original measurements (C_0) was determined from the duplicate sampling, one from each femoral vein. The standard uncertainty for a single determination was determined from each of these results according to:

$$u_M = \sqrt{\frac{\sum_{i=1}^{N}(rd)_i^2}{2N}} = \sqrt{\frac{\sum_{i=1}^{N}d_i^2}{2N}} \tag{39}$$

Equation 39, expressing the computation in two equivalent forms, was designed to estimate the total method (u_M) component of uncertainty. A major part of this was due to the sampling technique from each of the femoral veins. This component was termed pre-analytical (PA). Once the computations were determined from equation 39, the pre-analytical component was determined according to:

$$CV_M^2 = CV_{PA}^2 + CV_A^2 \tag{40}$$

Finally, the combined uncertainty was determined according to:

$$CV_T^2 = CV_{PA}^2 + CV_A^2 + CV_{Cal}^2 = CV_{PA}^2 + CV_A^2 + CV_{m_{Meth}}^2 + CV_P^2 + CV_V^2 \tag{41}$$

Incorporating the uncertainty estimates outlined in Table 1 of the study we obtain:

$CV_T = \sqrt{18.95\%^2 + 3.65\%^2 + 0.53\%^2 + 0.29\%^2 + 0.05\%^2} = 19.3\%$. With this estimate we, and the authors of the study, have assumed independence of the components and a multiplicative measurement model. The uncertainty budget for this example is shown in Table 5, from which we see that the pre-analytical or sampling component contributes by far the most to the combined uncertainty. This is not unexpected since it represents the sampling component. Sampling, when included as a component in the combined uncertainty estimate, is typically the largest contributor. The study reported that amongst the 27 cases, the concentration of methadone ranged from 0.005 to 2.29 mg/kg with a median value of 0.472 mg/kg. The median was appropriately reported, rather than the mean, because the distribution of results was positively skewed. Therefore, we would be interested in this case in computing a 95% confidence interval for the median. The most common approaches to estimating confidence intervals for a median do not involve uncertainty estimates. This results from the fact that the median is a quantile, specifically, the 50th percentile. One method for estimating the approximate 95% confidence interval for the median presented in

this study is to compute estimates of r and s as in equation 42. (Altman, et.al., 2000) For our sample size of n=27 and rounding the estimates to the nearest integer we obtain the results seen in equation 43. This would indicate that the 8th and 20th ordered observations would provide an approximate 95% confidence interval for the population median. The exact level of confidence for this example based on the binomial distribution would be 98.1%. (Altman, et.al., 2000)

Source	Type	%CV	Percent[1]
Pre-Analytical	A	18.95%	96%
Analytical	A	3.65%	3.9%
Mass of Methadone	A	0.53%	0.08%
Purity	B	0.29%	0.02%
Volume	B	0.05%	0%
Combined Uncertainty		19.3%	100%

[1]Percent of contribution to total combined uncertainty

Table 5. Uncertainty budget for the post-mortem measurement of methadone in femoral blood

$$r = \frac{n}{2} - \left[Z_{1-\alpha/2} \cdot \frac{\sqrt{n}}{2} \right] \qquad s = 1 + \frac{n}{2} + \left[Z_{1-\alpha/2} \cdot \frac{\sqrt{n}}{2} \right] \qquad (42)$$

$$r = \frac{27}{2} - \left[1.96 \cdot \frac{\sqrt{27}}{2} \right] = 8.4 \approx 8 \qquad s = 1 + \frac{27}{2} + \left[1.96 \cdot \frac{\sqrt{27}}{2} \right] = 19.6 \approx 20 \qquad (43)$$

8. Uncertainty in a blood alcohol analysis

The unique aspect of this example will be the addition of the uncertainty due to calibration. We will assume that duplicate blood alcohol results of 0.104 and 0.107 g/dL were obtained from the same headspace gas chromatograph. The following is our assumed measurement function:

$$C_{corr} = \frac{C_0 R}{X_{Cont}} \cdot f_{dilutor} \cdot f_{Calib} \qquad (44)$$

where: C_{corr} = the corrected BAC results
C_0 = the mean of the original measurement results
R = the traceable reference control value
$\overline{X}_{Cont}$ = the mean results from measuring the controls
$f_{dilutor}$ = the correction factor for the dilutor
f_{Calib} = the correction factor for the calibration
We have added an additional correction factor $\left(f_{Calib} \right)$ in equation 44 which we also set equal to one and also include its uncertainty component. We will assume that the instrument was calibrated with a linear five point calibration curve generated by the use of

five traceable control standards. The calibration curve was generated by linear least squares yielding the following function:

$$Y = a + bX \tag{45}$$

where: Y = instrument response, X = known control concentration values and a and b are model parameters. The objective in developing a calibration curve is to estimate the true value of a future unknown concentration (X) given some instrument response (Y). Therefore, we find the inverse of equation 45:

$$X = \frac{Y - a}{b} . \tag{46}$$

For our purposes, we are interested in determining the uncertainty in X found in equation 46. The parameters a and b, however, are correlated. We can eliminate the parameter a by solving for a according to $a = \overline{Y} - b\,\overline{X}$ and then substituting this into equation 46 according to:

$$X_0 = \frac{Y_0 - \left(\overline{Y} - b\,\overline{X}\right)}{b} \quad \Rightarrow \quad X_0 = \frac{Y_0 - \overline{Y}}{b} + \overline{X} \tag{47}$$

where: X_0 = a future single estimate of concentration
Y_0 = a future single instrument response
$\overline{Y}$ = the mean of the instrument responses during calibration
$\overline{X}$ = the mean of the control samples used during calibration
From equation 47 we see that X_0 is a function of only three random variables: Y_0, $\overline{Y}$, and b. Solving for the uncertainty in X_0 by the method of error propagation we obtain:

$$u_{X_0} = \frac{S_{Y|X}}{b} \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{\left(Y_0 - \overline{Y}\right)^2}{b^2 \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}} \tag{48}$$

where: $S_{Y|X}$ = standard error from regression of Y on X in developing the calibration curve
b = the slope of the calibration curve
m = the number of measurements used to estimate X_0
n = the number of measurements used to generate the calibration curve
We will assume specific values for the terms in equation 48 and solve for the uncertainty according to:

$$u_{X_0} = \frac{(0.005)}{(1.02)} \sqrt{\frac{1}{2} + \frac{1}{5} + \frac{(0.1055 - 0.1516)^2}{(1.02)^2 (0.046)}} = 0.0042$$

Now, for our example we will assume the variables for equation 44 found in Table 6. For purposes of determining the uncertainties in each of the correction factors we assume $f_{Dilutor}$ to be 10.65 and f_{Calib} to be 0.1058 g/dL. However, for estimating the corrected blood alcohol concentration in equation 44 we assume each to be 1.0. Next, we can estimate our corrected blood alcohol concentration according to:

$$C_{corr} = \frac{C_0\,R}{\overline{X}} \cdot f_{dilutr} \cdot f_{Calib} \;=\; \frac{(0.1055)(0.100)}{(0.1025)} \cdot 1 \cdot 1 \;=\; 0.1029\,g\,/\,dL$$

We now combine the standard uncertainty components to determine the combined uncertainty according to equation 49. Estimating an approximate 95% uncertainty interval would yield:

$$0.1029 \pm 2(0.0020) \;\Rightarrow\; 0.1029 \pm 0.0040 \;\Rightarrow\; 0.0989\;to\;0.1069\,g\,/\,dL\,.$$

The percent contribution from each component to the combined uncertainty in this example is: C_0 10%, R 2%, $\overline{X}_{Cont}$ 1% , $f_{Dilutor}$ 1% and f_{Calib} 86%. From this we see that the calibration uncertainty contributed by far the most to the combined uncertainty. This may have resulted from the values assumed for this example and may not reflect most forensic programs. Each laboratory would need to determine this for their particular context. It should also be noted that equation 48 includes the uncertainty only of the least squares estimates and not that of the reference standards used as calibrants. These could be added as separate components. There are other methods to account for the uncertainty in calibration as well. For example, the maximum vertical deviation between the line of identify and the least squares regression line can be divided by the square root of three, assuming the uniform distribution, and

Variable	Estimate	Uncertainty	n
C_0	0.1055	0.0009	2
R	0.100	0.0003	1
$\overline{X}_{Cont}$	0.1025	0.0008	16
$f_{Dilutor}$	10.65	0.05	10
f_{Calib}	0.1058	0.0042	5

Table 6. The values of specific variables assumed for our blood alcohol measurement model

$$\frac{u_{\overline{C}_{Corr}}}{\overline{C}_{Corr}} = \sqrt{ \left[\frac{\frac{u_{C_0}}{\sqrt{n}}}{C_0}\right]^2 + \left[\frac{\frac{u_R}{\sqrt{n}}}{R}\right]^2 + \left[\frac{\frac{u_{\overline{X}}}{\sqrt{n}}}{\overline{X}}\right]^2 + \left[\frac{\frac{u_{f_{dilutor}}}{\sqrt{n}}}{f_{dilutor}}\right]^2 + \left[\frac{\frac{u_{f_{Calib}}}{\sqrt{n}}}{f_{Calib}}\right]^2 }$$

(49)

$$\frac{u_{\overline{C}}}{0.1029} = \sqrt{ \left[\frac{\frac{0.0009}{\sqrt{2}}}{0.1029}\right]^2 + \left[\frac{\frac{0.0003}{\sqrt{1}}}{0.100}\right]^2 + \left[\frac{\frac{0.0008}{\sqrt{16}}}{0.1025}\right]^2 + \left[\frac{\frac{0.050}{\sqrt{10}}}{10.65}\right]^2 + \left[\frac{\frac{0.0042}{\sqrt{5}}}{0.1058}\right]^2 }$$

$$u_{\overline{Y}} = 0.1029(0.0192) = 0.0020\,g\,/\,dL$$

divided by the concentration value of X at that point. This is often termed a "lack of linearity" component.

The preceding examples presented here have been illustrative only. There was no intention that the uncertainty estimates assumed were the only ones to be considered or even represented any specific laboratory program. They were presented simply to illustrate the computations involved. Indeed, there are surely other components to be considered. (Sklerov and Couper, 2011) These must be identified by the forensic toxicologist considering their particular laboratory, protocol, instruments, customers and the required fitness-for-purpose.

9. Different methods for estimating uncertainty

We have illustrated above several examples for estimating the combined uncertainty in contexts relevant to forensic toxicology. These examples have presented the standard bottom-up approach recommended largely by the *GUM* document. There are, however, several other approaches to dealing with uncertainty that have been proposed in the forensic toxicology and metrological literature. Wallace, for example, has proposed a number of different methods for estimating measurement uncertainty. (Wallace, 2010)

9.1 Use of proficiency test data

One method advocated by Wallace is the use of proficiency test data. (Wallace, 2010) Proficiency testing basically consists of an organizing laboratory which, employing well established and traceable methods, prepares and tests the concentrations of several samples. These samples are then sent blindly to participating laboratories with instructions on how the measurements are to be performed, recorded and then returned to the organizing laboratory. The samples are to be treated by the participating laboratories as routine case samples and tested according to their routine protocols. The organizing laboratory summarizes the data reporting means, standard deviations and various plots, including, for example, Z-scores. The standard deviations at various mean concentrations can be used to generate uncertainty functions. Clearly, these estimates will exhibit rather large variation due to the different laboratories, instruments, protocols, analysts, time, etc. These estimates, conditioned on the appropriate concentration, can be used as the total method component in the combined uncertainty estimate. Consider an example where we have duplicate blood alcohol results obtained in the toxicology laboratory of 0.118 and 0.116 g/dL. The laboratory participated in a proficiency study which yielded the uncertainty function observed in figure 7. This figure was actually generated from data available from Collaborative Testing Services [CTS]. For this example we will assume the following measurement function:

$$C_{Corr} = \frac{C_0 R}{\overline{X}} \tag{50}$$

where: C_{Corr} = the corrected measurement result
C_0 = the mean of the original duplicate measurements
R = the reference value for the controls
$\overline{X}$ = the mean result for measuring the reference controls

The mean of our assumed duplicate results is 0.1170 g/dL. The reference value is R=0.100 g/dL with a Type B standard uncertainty of 0.0003 g/dL. The mean measurement of the

controls were $\overline{X} = 0.1024\,g\,/\,dL$ with n=34 measurements and a standard uncertainty of 0.0009 g/dL. Computing our corrected estimate from equation 56 we obtain 0.1143 g/dL. Using this value to estimate our method uncertainty from the equation found in figure 8 we obtain: $u_M = 0.0369(0.1143) + 0.00129 = 0.0055\,g\,/\,dL$. Assuming independence and the multiplicative model of equation 50, we now estimate our combined uncertainty as seen in equation 51. The approximate 95% confidence interval (k=2) for the true mean blood alcohol concentration in this example would be:

$$\overline{Y} \pm 2u_{\overline{Y}} \Rightarrow 0.1143 \pm 2(0.0039) \Rightarrow 0.1065 \; to \; 0.1221\,g\,/\,dL\,.$$

The risk in using proficiency data in this manner is that the actual uncertainty associated with a particular laboratory may be overestimated. Another limitation to keep in mind is that the proficiency data may not have been generated with the same analytical protocol employed within a particular laboratory. Proficiency data, however, does have a large source of variation, which may be acceptable within the forensic context. The uncertainty budget for these results is found in Table 7. The method uncertainty determined from the proficiency test data in this example, contributed by far the most to the combined uncertainty while the reference and analytical components could effectively be ignored.
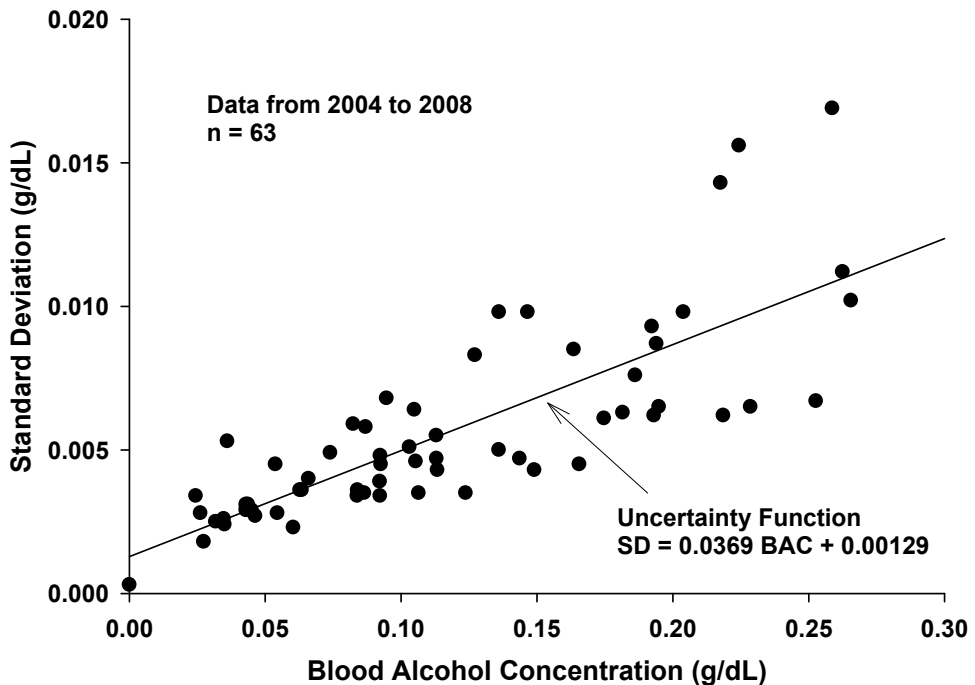

Fig. 7. Plot of the standard deviation against concentration and determination of an uncertainty function from CTS proficiency test blood alcohol data

$$\frac{u_C}{C} = \sqrt{CV_{C_0}^2 + CV_R^2 + CV_{\overline{X}}^2} \quad = \quad \frac{u_C}{C} = \sqrt{\left[\frac{\frac{u_{C_0}}{\sqrt{n}}}{C_0}\right]^2 + \left[\frac{\frac{u_R}{\sqrt{n}}}{R}\right]^2 + \left[\frac{\frac{u_{\overline{X}}}{\sqrt{n}}}{\overline{X}}\right]^2}$$

(51)

$$\frac{u_{\overline{Y}}}{0.1143} = \sqrt{\left[\frac{\frac{0.0055}{\sqrt{2}}}{0.1143}\right]^2 + \left[\frac{\frac{0.0003}{\sqrt{1}}}{0.100}\right]^2 + \left[\frac{\frac{0.0009}{\sqrt{34}}}{0.1024}\right]^2} \quad \Rightarrow \quad u_{\overline{Y}} = 0.1143(0.0342) = 0.0039\, g\,/\,dL$$

Source	Type	%CV	Percent[1]
Method (Proficiency)	A	5%	99%
Reference	B	0.3%	0.8%
Analytical	A	0.9%	0.2%
Total			100%

[1] Percent of contribution to total combined uncertainty

Table 7. Uncertainty budget resulting from the use of proficiency test data as the estimate for method uncertainty

9.2 Using the guard band approach

Employing a guard band is another approach to accounting for measurement uncertainty. (EURACHEM/CITAC, 2000) Use of the guard band is a tool for determining compliance within specified limits. It establishes a decision rule, particularly relevant where there are critical or prohibited analytical limits which may define, for example, binary outcomes such as pass/fail, guilty/not guilty, etc. These can be important in drunk-driving prosecution where alcohol results (either blood or breath) are introduced to establish whether the subject exceeded the legal limit. Consider the example where an individual provided duplicate breath alcohol results of 0.092 and 0.098 g/210L. A traceable commercially purchased simulator control standard having a reference value of 0.0824 g/210L and a Type B combined uncertainty of 0.0008 g/210L was measured by the breath test instrument. The mean of n=46 measurements with this control was 0.0856 g/210L with a standard uncertainty of 0.0010 g/d10L. We wish to determine an upper limit to the guard band, above which we will be 99% confident that the individual's true mean breath alcohol concentration exceeds 0.080 g/210L. This can be visualized in figure 8 where we see that the upper limit of the guard band is the value 0.080 + ku_C. We must first find the combined uncertainty (u_C) and then the appropriate value of k. The value of k will actually be from the t-distribution in this example and will need to correspond to a 98% confidence interval. The degrees of freedom will be determined from the Welch-Satterthwaite equation. We begin by identifying our measurement function as follows:

$$C_{Corr} = \frac{C_0\,R}{X}$$

(52)

where: C_{Corr} = the corrected breath alcohol concentration

R = the traceable control reference value

$\overline{X}$ = the mean of replicate (n=18) measurements of the reference control standard
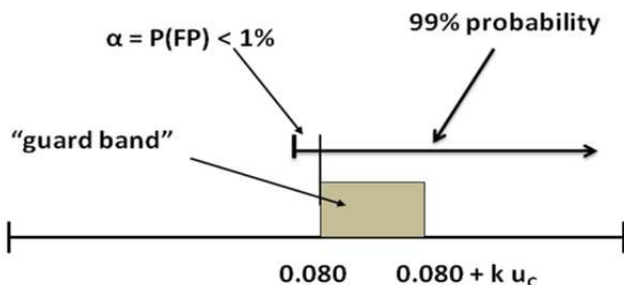

Fig. 8. Illustrating the construction and use of a "guard band" to determine compliance with a specified limit

The corrected mean breath alcohol results in our example is found from equation 52 to be:

$C_{Corr} = \dfrac{(0.0950)(0.0824)}{(0.0856)} = 0.0914\,g\,/\,210L$. We will assume that our method uncertainty is determined from the uncertainty function found in figure 6 which results in: $u = 0.0260(0.0914) + 0.00095 = 0.0033\,g\,/\,210L$. Given that our measurement function in equation 52 is multiplicative, we now find our combined uncertainty, assuming independence, as in equation 51:

$$\frac{u_{\overline{Y}}}{0.0914} = \sqrt{\left[\frac{\frac{0.0033}{\sqrt{2}}}{0.0914}\right]^2 + \left[\frac{\frac{0.0008}{\sqrt{1}}}{0.0824}\right]^2 + \left[\frac{\frac{0.0010}{\sqrt{46}}}{0.0856}\right]^2} \quad \Rightarrow \quad u_{\overline{Y}} = 0.0914(0.0274) = 0.0025\,g\,/\,210L .$$

We now find the relative combined uncertainty by removing the values of n according to:

$$\frac{u_{\overline{Y}}}{0.0914} = \sqrt{\left[\frac{0.0033}{0.0914}\right]^2 + \left[\frac{0.0008}{0.0824}\right]^2 + \left[\frac{0.0010}{0.0856}\right]^2} \quad \Rightarrow \quad u_{\overline{Y}} = 0.0914(0.0391) = 0.0036\,g\,/\,210L .$$

Now we determine our effective degrees of freedom using the Welch-Satterthwaite equation as follows:

$$v_{eff} = \frac{\left[\dfrac{0.0036}{0.0914}\right]^4}{\dfrac{\left[\dfrac{.0033}{0.0914}\right]^4}{\infty} + \dfrac{\left[\dfrac{.0008}{0.0824}\right]^4}{\infty} + \dfrac{\left[\dfrac{0.0010}{0.0856}\right]^4}{45}} = 5814.8 \approx \infty$$

Notice that in the Welch-Satterthwaite equation we have changed our degrees of freedom for the total method component to infinity. This is because the standard uncertainty estimate

(0.0033 g/210L) from figure 6 is based on much more than one degree of freedom ($n>27,000$). The degrees of freedom for the reference standard in the Welch-Satterthwaite equation is set to infinity because it is a Type B uncertainty without information on the degrees of freedom provided. Since we have essentially an infinite number of effective degrees of freedom we select our k (or t) value of 1.96. We can now compute the upper limit for our guard band: 0.080 + 2.33(0.0025) = 0.0858 g/210L. Since the subject's corrected mean breath alcohol concentration exceeds the upper guard band limit of 0.0858 g/210L we conclude there is 99% confidence that the individual's true mean breath alcohol concentration exceeds 0.080 g/210L. Values exceeding $0.080 + ku_c$ could be considered within the "rejection zone". For a measurement in this region, the probability of a "false rejection" is less than α, the probability of the false-positive error. (Desimoni and Brunetti, 2007) One must also keep in mind that for guard band estimates at different concentrations, the combined uncertainty estimates need to incorporate the method uncertainty appropriate to that concentration. The guard band approach could also be generated based on a large set of historical data and then employed for a period of time. The estimates could be updated annually, for example, to ensure the system remains in statistical control. The assumptions with this approach is that the individuals continue to be tested on the same instrumentation and protocols used to generate the guard band limits and that the system remains in statistical control. The United Kingdom is one jurisdiction that employs a guard band approach. (Walls and Brownlee, 1985) A value of 6mg/dL is subtracted from the mean of duplicate blood alcohol results below 100 mg/dL and 6% is deducted from results over 100 mg/dL. The results of this deduction must exceed their legal limit of 80 mg/dL for prosecution. Denmark employs a similar approach where they deduct 0.1 g/Kg to compute their level for prosecution. (Kristiansen and Petersen, 2004) Similarly, Sweden employs the guard band approach to uncertainty estimation by requiring that the lower 99.9% confidence interval limit for mean results must exceed their legal limit. (Jones and Schuberth, 1989) Guard band calculations could also be incorporated into computerized breath test instruments for immediate determination of critical limits for purposes of prosecution.

9.3 Uncertainty estimation from total allowable error

There is considerable debate regarding the best method for estimating measurement uncertainty and whether it is even necessary. Many argue that measurement uncertainty is unnecessary because it may be misunderstood by the customer or confuse the interpretation. Since bias is only determined with regard to a reference standard, many analytes do not have standards available while others have several. As a result, it is argued that bias may not be validly determined in the first place. Some that argue against the use of measurement uncertainty would advocate the use of total allowable error (TE_a). (Westgard, 2010) Total allowable error is determined from the following linear model:

$$TE_a = |bias| + k\,u_C \quad . \tag{53}$$

The total allowable error combines both bias and random components and estimates the upper limit. In some cases this may over estimate the actual capability of the analytical method or laboratory performance. Moreover, the method of total allowable error does not correct for bias - it simply includes the maximum level allowable. If we were to allow a maximum bias of 4% and the relative combined uncertainty for the method was 2% and we

selected a coverage factor of k=2, we would obtain: $TE_a = |4\%| + 2(2\%) = 8\%$. This would provide an upper limit estimate for the customer who could be assured, with a high degree of probability, that the total error would not exceed this limit. One might report the final results in this context as: *The whole blood alcohol results were 0.094 and 0.096 g/dL having a mean of 0.0950 g/dL which did not have an associated total allowable error of more than 8% with approximately 95% probability.* One context appropriate for the application of the total error method is where a single control is measured as part of an analytical run. If the control exceeded the total allowable error, one would not know whether it was due to bias or random sources. However, the result would be caught and the system corrected before resuming routine measurements. One of the criticisms of the method of total allowable error method is that it allows bias to exist without correcting for it. (Dybkaer, 1999, Dybkaer, 1999) Admittedly, the total error method provides a very conservative estimate, a maximum actually, for interpreting measurement uncertainty.

9.4 Monte Carlo methods

Monte Carlo methods are simulation techniques that are more computationally intensive. With faster computers available, these methods are becoming more popular. Monte Carlo methods require assumptions regarding the measurement function along with the distributional form and parameters for each of the input components, being themselves random variables. Random data are then simulated from each of the component distributions, placed into the measurement function, followed by the computation of the measurand. This is done a large number of times, generating a distribution of response values. From these results, the distribution, the expected value and the standard uncertainty of the response variable can be determined. As a result we do not need to assume some distributional form for the response variable and we have a direct, empirically determined estimate of uncertainty. Monte Carlo methods also avoid two limitations of the GUM method – the required linear relationship between the response variable and the components and the justified application of the central limit theorem. (Fernandez, et.al., 2009) Consider the following example of a breath alcohol measurement function where we have six input variables: $\overline{Y}_{Corr} = \dfrac{\overline{Y}_0 \cdot GC_{Sol} \cdot R}{\overline{X} \cdot K \cdot GC_{Cont}}$ and where we assume the following distributions for each of the six input variables: $\overline{Y}_0 \sim N(0.1250, 0.0047^2)$ the mean of the original n measurements, $GC_{Sol} \sim N(0.1025, 0.0008^2)$ the mean of the simulator solution measurements by gas chromatography, $R \sim N(0.100, 0.0003^2)$ the traceable reference value, $\overline{X} \sim N(0.0825, 0.0012^2)$ the mean of the breath test instrument measuring the simulator solution, $K \sim Unif(1.21, 1.25)$ the ratio of partition coefficients in the simulator heated to 34^0C and $GC_{Cont} \sim N(0.0980, 0.0008^2)$ the results from measuring the traceable controls on the gas chromatograph. We employ a routine written in R that simulates random results from each of these distributions and computes the response variable $(\overline{Y}_{Corr})$. This is done 10,000 times. The resulting distribution for the response variable is seen in figure 9. The

expected value for the response variable is 0.1287 g/210L with an empirical 95% confidence interval of 0.1215 to 0.1360g/210L, determined from the distribution of results in figure 9. The sampling/method component was also correctly identified as having the largest contribution to total uncertainty of 85%.
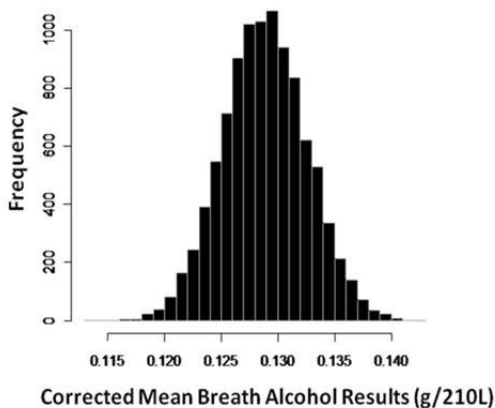

Fig. 9. A distribution of 10,000 Monte Carlo simulated measurement results

10. Uncertainty in qualitative analysis

Several measurements performed in forensic toxicology are qualitative in nature. These measurements typically take the form of a binary response (i.e., pass/fail, yes/no, over/under, present/absent, etc.). They are classification in nature where materials are assigned to discrete groups based on measurement results. Diagnostic tests are one important example of qualitative analyses. Their qualitative results are important indicators of whether some specified threshold has been exceeded or not and are important for the determination of further confirmatory analyses. In some cases the measurement system will respond simply with binary results (green light/red light). At other times the measurement system is quantitative on a continuous scale which can be dichotomized. For example, a pre-arrest breath test instrument employing a fuel cell might measure the breath alcohol on a continuous concentration scale but is interpreted as being greater than or equal to 0.080 g/210L or less than 0.080 g/210L. In either case, the response is considered binary and thus qualitative. The uncertainty associated with qualitative analyses has received much less attention than that of quantitative analysis. The uncertainty in qualitative analyses is basically probabilistic in nature - that is, we are interested in the probability of being correct in our decision. We are concerned primarily with the probability of false positive and false negative results. While there are a number of statistical methods for estimating the uncertainty associated with qualitative or diagnostic test results, there is no consensus as to which is to be preferred. (EURACHEM/CITAC, 2003, Pulido, et.al., 2003, Ellison, et.al., 1998) Some methods involve the simple determination of false-positive (FP) and false-negative (FN) fractions which in turn assess the probability of making a wrong decision. (Pepe, 2003) Other qualitative and quantitative methods employ Baye's Theorem which is argued by many as a superior approach to estimating and interpreting measurement

uncertainty. (Gleser, 1998, Weise and Woger, 1992, Kacker and Jones, 2003, Phillips, et.al., 1998). Space does not permit further discussion of these important and useful methods.

11. Discussion

Several examples have been presented here for estimating measurement uncertainty in the context of forensic toxicology. By no means do these examples imply that all possible uncertainty components have been considered. These examples were intended primarily to illustrate the general approach and computations involved. Moreover, while an example may have assumed a blood alcohol context, it could just as well have been applied in the context of breath or drug analysis. While the general approach will be relevant to most methods in forensic toxicology, each laboratory will need to identify and quantify its uncertainty components unique to its protocols and instrumentation. The examples and discussion presented here have also assumed independence among the input or predictor variables. This is certainly not always a valid assumption. In some measurement contexts there will be significant correlation between input variables which must be accounted for. (GUM, EURACHEM/CITAC, Ellison, 2005) While these concepts may be new to some practicing toxicologists, the concept of measurement uncertainty should not raise concerns for the forensic sciences. The emphasis should be on their ability to quantify confidence of measurement results. They should be presented in a manner that emphasizes and demonstrates their fitness-for-purpose. Modern technology should enhance and simplify these computations as well. Spreadsheet programs can be developed which require only the entry of specific values followed by the generation of all uncertainty results. Moreover, such computations can even be incorporated into the software of aalytical instruments. Such technology, when validated, should greatly simplify the process.

Several factors are responsible for the emphasis today on reporting measurement results along with their uncertainty. These include legal, economic, liability, accrediting and technological considerations. As professional toxicologists concerned with providing measurement results of the highest possible quality, we must be prepared to make this extra effort of providing the relevant uncertainty. Since there is no consensus regarding the best approach for computing uncertainty at this time, toxicologists should be familiar with the several approaches suggested here and then select and validate the one which best suits their analytical, procedural and legal context. The literature is rich with material regarding measurement uncertainty and should be carefully reviewed by toxicologists. (Drosg, 2007, Williams, 2008, Fernandez, 2011 Ekberg,et.al., 2011) This effort wll enhance the quality and interpretability of our measurement results and help establish a foundation of "evidence based forensics". The unavoidable fact of measurement uncertainty results in the risk of making incorrect decisions. While ignoring the uncertainty increases this risk, providing the uncertainty reduces and quantifies the risk for the decision maker. This fact alone should motivate the legal community to request and forensic toxicologists to rigorously estimate and provide such estimates.

12. References

Altman, D.G., Machin, D., Bryant, T.N. and Gardner, M.J. (2000). *Statistics with confidence*, (2nd edition), British Medical Journal, London, ISBN 0-7279-1375-1

ASCLD/LAB. (2011) ASCLD/LAB Policy on Measurement Uncertainty, August 2011, Available from: http://www.ascld-lab.org/documents/AL-PD-3051.pdf,

Ballico, M. (2000). Limitations of the Welch-Satterthwaite Approximation for Measurement Uncertainty Calculations. *Metrologia*, Vol.37, pp. 61-64.

Barwick, V.J. & Ellison, S.L.R. (1999). Measurement Uncertainty: Approaches to the Evaluation of Uncertainties Associated with Recovery. *Analyst*, Vol.124, pp. 981-990.

Bich, W., Cox, M. and Harris, P. (2006). Evolution of the 'Guide to the Expression of Uncertainty in Measurement'. *Metrologia*, Vol.43, pp. S161-S166.

Daubert vs. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Deldossi, L. and Zappa, D. (2009). ISO 5725 and GUM: Comparison and Comments. *Accred Qual Assur*, Vol.14, pp. 159-166.

Desimoni, E. & Brunetti, B. (2010). About Acceptance and Rejection Zones as Defined in the EURACHEM/CITAC Guide (2007) 'Use of Uncertainty Information in Compliance Assessment'. *Accred Qual Assur*, Vol.15, pp. 45-47.

Drosg, M. (2007). *Dealing with Uncertainties: A Guide to Error Analysis*, Springer, New York.

Dybkaer, R. (1999). Setting Quality Specifications for the Future with Newer Approaches to Defining Uncertainty in Laboratory Medicine. *Scand J Clin Lab Invest*, Vol.59, pp. 579-584.

Dybkaer, R. (1999). From Total Allowable Error via Metrological Traceability to Uncertainty of Measurement of the Unbiased Result, *Accred Qual Assur*, Vol.4, pp. 401-405.

Ekberg, C., Englund, S. & Liljenzin, J.O. (2011). Uncertainty Analysis of Correlated Stability Constants. *Accred Qual Assur*, Vol.16, pp. 185-189.

Ellison, S.L.R., Gregory, S. & Hardcastle, W.A. (1998). Quantifying uncertainty in qualitative analysis. *Analyst*, Vol.123, pp. 1155-1161.

Ellison, L. (2005). Including Correlation Effects in an Improved Spreadsheet Calculation of Combined Standard Uncertainties. *Accred Qual Assur*, Vol.10, pp. 338-343.

EURACHEM/CITAC Guide. (2000). *Quantifying Uncertainty in Analytical Measurement*, 2nd Ed., May 2011, Available from:
http://www.measurementuncertainty.org/mu/QUAM2000-1.pdf

EURACHEM/CITAC Guide. (2003). *The Expression of Uncertainty in Qualitative Testing*, May 2011, Available from:
http://www.nmschembio.org.uk/dm_documents/LGCVAM2003048_cOjjl.pdf

EURACHEM/CITAC Guide. (2007). *Use of uncertainty in compliance assessment*. Eds. Williams, A. and Ellison, S.L.R., April 2011, Available from:
http://www.citac.cc/EURACHEM-CITAC%20Guide%20
%20Use%20of%20uncertainty%20information%20%20in%20compliance%20assess
ment%20-%202007.pdf

Fernandez, M. (2011). On the use of the 'Uncertainty Budget' to Detect Dominant Terms in the Evaluation of Measurement Uncertainty. *Accred Qual Assur*, Vol.16, pp. 83-88.

Fernandez, M., Calderon, J. & Diez, P. (2009). Implementation in MATLAB of the Adaptive Monte Carlo Method for the Evaluation of Measurment Uncertainties. *Accred Qual Assur*, Vol.14, pp. 95-106.

Gates, K., Chang, N., Dilek, I., Jian, H., Pogue, S. & Sreenivasan, U. (2009). The Uncertainty of Reference Standards - A Guide to Understanding Factors Impacting Uncertainty, Uncertainty Calculations, and Vendor Certifications. *Journal of Analytical Toxicology*, Vol.33, pp. 532-539.

Gleser, L.J. (1998). Assessing Uncertainty in Measurement. *Statistical Science*, Vol.13 No.3, pp. 277-290.

International Organization for Standardization (2008). *International Vocabulary of Basic and General Terms in Metrology (VIM)*, 3rd ed., ISO, Geneva. May 2011, Available from:

http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008_pdf

International Organization for Standardization (ISO). (2008). *Guide to the Expression of Uncertainty in Measurement*, ISO, Geneva, 2008. June 2011, Available at:

http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

International Organization of Legal Metrology (OIML), *International Alcoholometric Tables*, Paris, available online at: http://www.oiml.org/publications/R/R022-e75.pdf

ISO/IEC 17025. (2000). General requirements for the competence of testing and calibration laboratories. London: British Standards Institution, pp. 1-26.

Jones, A.W. (1989). Differences Between Capillary and Venous Blood-alcohol Concentrations as a Function of Time After Drinking, with Emphasis on Sampling Variations in Left vs Right Arm. *Clin. Chem.*, Vol.35 No.3, pp. 400-404.

Jones, A.W. & Schuberth, J. (1989). Computer-aided Headspace Gas Chromatography Applied to Blood-alcohol Analysis: Importance of Online Process Control. *Journal of Forensic Sciences*, Vol.34 No.5, pp. 1116-1127.

Kacker, R. & Jones, A. (2003). "On Use of Bayesian Statistics to Make the *Guide to the Expression of Uncertainty in Measurement* Consistent. *Metrologia,* Vol.40, pp. 235-248.

Kacker, R., Sommer, K. & Kessel, R. (2007). Evolution of Modern Approaches to Express Uncertainty in Measurement. *Metrologia*, Vol.44, pp. 513-529.

Kacker, R., Kessel, R. & Sommer K. (2010). Assessing Differences Between Results Determined According to the Guide to the Expression of Uncertainty in Measurement. *Journal of Research of the National Institute of Standards and Technology*, Vol.115 No.6, pp. 453-459.

Kadis, R. (2007). Do We Really Need to Account for Run Bias when Producing Analytical Results with Stated Uncertainty? Comment on 'Treatment of Bias in Estimating Measurement Uncertainty' by G.E. O'Donnell and D.B. Hibbert, *Analyst*, Vol.132, pp. 1272-1274.

Kane, J.S. (1997). Analytical Bias: The Neglected Component of Measurement Uncertainty. *Analyst*, Vol.122, pp. 1283-1288.

King, B. & Lawn, R. (1999). International Interlaboratory Study of Forensic Ethanol Standards. *Analyst*, Vol. 124, pp. 1123-1130.

Kirkup, L. & Frenkel, B. (2006). *An introduction to uncertainty in measurement*. Cambridge University Press, Cambridge, UK.

Kristiansen, J. (2001). Description of a Generally Applicable Model for the Evaluation of Uncertainty of Measurement in Clinical Chemistry. *Clin Chem Lab Med*, Vol.39 No.10, pp. 920-931.

Kristiansen, J. (2003). The Guide to Expression of Uncertainty in Measurement Approach to Estimating Uncertainty: An Appraisal. *Clinical Chemistry*, Vol.49 No.11, pp. 1822-1829.

Kristiansen, J. & Petersen, H.W. (2004). An Uncertainty Budget for the Measurement of Ethanol in Blood by Headspace Gas Chromatography. *J of Analytical Toxicology*, Vol.28, pp. 456-463.

Krouwer, J. (2003). Critique of the Guide to the Expression of Uncertainty in Measurement Method of Estimating and Reporting Uncertainty in Diagnostic Assays. *Clinical Chemistry*, Vol.49 No.11, pp. 1818-1821.

Ku, H.H. (1966). Notes on the Use of Propagation of Error Formulas. *J. Res. Natl. Bureau Standards -C, Eng. Instrum.* 70C (4), pp. 916-921.

Linnet, K., Johansen, S.S., Buchard, A., Munkholm, J. & Morling, N. (2008). Dominance of Pre-analytical Over Analytical Variation for Measurement of Methadone and its Main Metabolite in Postmortem Femoral Blood. *Forensic Science International*, Vol. 179, pp. 78-82.

Marato, A., Boque, R., Riu, J. & Rius, F.X. (2002). Should Non-significant Bias be Included in the Uncertainty Budget? *Accred Qual Assur*, Vol.7, pp. 90-94.

Mnookin, J.L., Cole, S.A., Dror, I.E., Fisher, B.A.J., Houck, M.M., Inman, K., Kaye, D.H., Koehler, J.J., Langenburg, G., Risinger, D.M., Rudin, N., Siegel, J. & Stoney, D.A. (2011). The Need for a Research Culture in the Forensic Sciences. *UCLA Law Review*, Vol.58, pp. 725-779.

National Academy of Science. (2009). *Strengthening Forensic Science in the United States: A Path Forward*, National Research Council, The National Academies Press, Washington, D.C.

O'Donnell, G.E. & Hibbert, D.B. (2005). Treatment of Bias in Estimating Measurement Uncertainty. *Analyst*, Vol.130, pp. 721-729.

O'Donnell, G.E. & Hibbert, D.B. (2007). Reply to 'Do We Really Need to Account for Run Bias When Producing Analytical Results with Stated Uncertainty? Comment on 'Treatment of Bias in Estimating Measurement Uncertainty'. *Analyst*, Vol.132, pp. 1275-1277.

Pepe, M.S. (2003). T*he statistical evaluation of medical tests for classification and prediction.* Oxford University Press, Oxford.

Petersen, P.H., Stockl, D., Westgard, J.O., Sandberg, S., Linnet, K. & Thienpont, L. (2001). Models for Combining Random and Systematic Errors. Assumptions and Consequences for Different Models. *Clin Chem Lab Med*, Vol.39 No.7, pp. 589-595.

Philipp, R., Hanebeck, O., Hein, S., Bremser, W., Win, T. & Nehls, I. (2010). Ethanol/Water Solutions as Certified Reference Materials for Breath Alcohol Analyzer Calibration. *Accred Qual Assur*, Vol.15, pp. 141-146.

Phillips, S.D., Eberhardt, K.R. and Parry, B. (1997). Guidelines for Expressing the Uncertainty of Measurement Results Containing Uncorrected Bias. *J of Research of the National Institute of Standards and Technology*, Vol.102 No.5, pp. 577-585.

Phillips, S.D., Estler, W.T., Levenson, M.S. and Eberhardt, K.R., (1998), "Calculation of Measurement Uncertainty Using Prior Information", *J. of Research of the National Institute of Standards and Technology*, Vol.103 No.6, pp. 625-632.

Pulido, A., Ruisanchez, I., Boque, R. & Rius, F.X. (2003). Uncertainty of Results in Routine Qualitative Analysis. *Trends in Analytical Chemistry*, Vol.22 No1.10, pp. 647-654.

Rocke, D.M. & Lorenzato, S. (1995). A Two-Component Model for Measurement Error in Analytical Chemistry. *Technometrics*, Vol.37 No.2, pp. 176-184.

Sklerov, J.H. & Couper, F.J. (2011). Calculation and Verification of blood Ethanol Measurement Uncertainty for Headspace Gas Chromatography. *J Anal Toxicol*, Vol.35, pp. 402-410.

Stowell, A.R., Gainsford, A.R. & Gullberg, R.G. (2008). New Zealand's Breath and Blood Alcohol Testing Programs: Further Data Analysis and Forensic Implications. *Forensic Science International*, Vol.178, pp. 83-92.

Synek, V. (2005). Attempts to Include Uncorrected Bias in the Measurement Uncertainty. *Talanta*, Vol.65, pp.829-837

Taylor, B.N. & Kuyatt, C.E. (1994). NIST Technical Note 1297, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, Physics Laboratory, Gaithersburg, MD: National Institute of Standards and Technology. March 2011, Available from: http://physics.nist.gov/Document/tn1297.pdf

Thompson, M. & Fearn, T. (1996). What Exactly is Fitness For Purpose in Analytical Measurement. *Analyst*, Vol.121, pp. 275-278.

Thompson, M. (1997). Comparability and Traceability in Analytical Measurements and Reference Materials. *Analyst*, Vol.122, pp. 1201-1205.

Thompson, M. (1998). Do we really need detection limits? *Analyst*, Vol.123, pp. 405-407.

Thompson, M., Ellison, S.L.R., Fajgelj, A., Willetts, P. & Wood, R. (1999). Harmonized Guidelines for the Use of Recovery Information in Analytical Measurement. *Pure and Applied Chemistry*, Vol.71 No.2, pp. 337-348.

Thompson, M. & Coles, B.J. (2011). Use of the 'Characteristic Function' for Modeling Repeatability Precision. *Accred Qual Assur*, Vol.16, pp. 13-19.

Wallace, J. (2010). Proficiency Testing as a Basis for Estimating Uncertainty of Measurement: Application to Forensic Alcohol and Toxicology Quantitations. *J Forensic Sci*, Vol. 55, pp. 767-773.

Wallace, J. (2010). Ten Methods for Calculating the Uncertainty of Measurement. *Science and Justice*, Vol.50 No.4, pp. 182-186.

Walls, H.J. & A.R. Brownlie. (1985). *Drink, Drugs and Driving*. 2nd ed., Sweet and Maxwell, London, p. 108.

Weise, K. & Woger, W. (1992). A Bayesian Theory of Measurement Uncertainty, *Meas Sci Technol*, Vol.3, pp. 1-11.

Westgard, J.O. (2010). Managing Quality vs. Measuring Uncertainty in the Medical Laboratory. *Clin Chem Lab Med*, Vol.48 No.1, pp. 31-40.

Williams, A. (2008). Principles of the EURACHEM/CITAC Guide 'Use of Uncertainty Information in Compliance Assessment'. *Accred Qual Assur*, Vol.13, pp. 633-638.
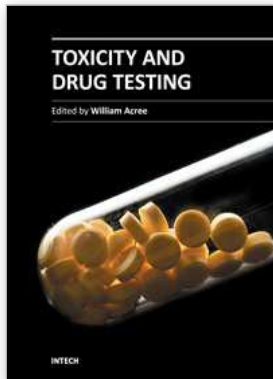
Toxicity and Drug Testing

Edited by Prof. Bill Acree

Modern drug design and testing involves experimental in vivo and in vitro measurement of the drug candidate's ADMET (adsorption, distribution, metabolism, elimination and toxicity) properties in the early stages of drug discovery. Only a small percentage of the proposed drug candidates receive government approval and reach the market place. Unfavorable pharmacokinetic properties, poor bioavailability and efficacy, low solubility, adverse side effects and toxicity concerns account for many of the drug failures encountered in the pharmaceutical industry. Authors from several countries have contributed chapters detailing regulatory policies, pharmaceutical concerns and clinical practices in their respective countries with the expectation that the open exchange of scientific results and ideas presented in this book will lead to improved pharmaceutical products.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Rod G. Gullberg (2012). Measurement Uncertainty in Forensic Toxicology: Its Estimation, Reporting and Interpretation, Toxicity and Drug Testing, Prof. Bill Acree (Ed.), ISBN: 978-953-51-0004-1, InTech, Available from: http://www.intechopen.com/books/toxicity-and-drug-testing/measurement-uncertainty-in-forensic-toxicology-its-estimation-reporting-and-interpretation

INTECH
open science | open minds