Quality Assurance and Safety of Crops & Foods

ICC MoniQA

Quality Assurance and Safety of Crops & Foods 2011, **3**, 123–134

ORIGINAL ARTICLE

What is the best way to ensure that valid analytical methods are used for food control?

Martin Rose[1], Roland Poms[2,3], Roy Macarthur[1], Bert Pöpping[4] & Franz Ulberth[5]

1 The Food and Environment Research Agency, Sand Hutton, York, YO41, UK
2 ICC - International Association for Cereal Science and Technology, Marxergasse 2, 1030 Wien, Vienna, Austria
3 ESR – Institute for Environmental Science and Research, 27 Creyke Road, Christchurch, New Zealand
4 *Eurofins UK*, Pocklington, UK
5 *European Commission, Joint Research Centre* Institute for Reference Materials and Measurements, Retieseweg 111, B-2440 Geel, Belgium

Keywords
analytical criteria; fitness for purpose; food control; standard methods; uncertainty.

Correspondence:
Martin Rose, The Food and Environment Research Agency, Sand Hutton, York YO41 1LZ, UK.
Email: martin.rose@fera.gsi.gov.uk

Received 3 November 2010; revised 10 January 2011; accepted 15 January 2011.

doi:10.1111/j.1757-837X.2011.00101.x

Abstract

Monitoring and Quality Assurance in the Food Supply Chain is a Network of Excellence funded by the European Union. This Network of Excellence aims to make food safer by harmonizing the quality of methods used for food control. Part of this process involves the development and validation of new methods, including rapid methods and emerging technologies; the production of practical harmonized guidance on method validation and criteria for analytical methods; and the production of reference and testing materials to be used for food control.

New technologies and analytical research enable us to measure new and emerging food contaminants and other chemicals that may be a threat to the health of the consumer. Once risk assessment confirms the threat, legislation may be enacted to limit the amount of these chemicals present in food that is sold. In order to enforce this legislation it is necessary to be able to detect the presence of chemical and measure its concentration in food. Hence, the consumer protection provided by enforcement depends on how well the measurement method performs. We need to be confident that measurement methods are performing sufficiently well to protect the consumer, without leading to the rejection of large quantities of food that comply with legislation. A number of approaches that can be used to provide confidence include: the use of standard methods, the use of analytical criteria that describe the performance of a method, and consideration of fitness for purpose based on measurement uncertainty. This paper examines the utility and ease of application of the different approaches. In addition a simple method for assessing fitness for purpose, the uncertainty profile, is discussed.

Rose M, Poms R, Macarthur R, Pöpping B & Ulberth F (2011). What is the best way to ensure that valid analytical methods are used for food control?. *Quality Assurance and Safety of Crops & Foods*, 3, 123–134

Introduction

Food control laboratories are often generalist facilities that are required to offer a wide range of capability, often not only in the food sector, and while under variable and sometimes considerable financial restraint. Similarly the level of technology and analytical expertise may vary between laboratories. Hence, many such laboratories rely on the availability of standard methods from a variety of sources that are easy to implement, sufficiently robust and that have been shown to meet the requirements of legislation.

Identification of chemicals in food that might cause concern for human health is not straightforward because, typically, toxicologists and analytical chemists each wait for

the other to make the first move. Toxicologists are reluctant to study chemicals unless there is evidence of exposure (to measure exposure requires an analytical method). In turn, analytical chemists do not prioritize compounds for method development unless there is evidence that they may cause harm, which requires output from a toxicological evaluation.

Evidence of new exposure to chemicals from the diet is usually first uncovered as a result of research, often using new analytical methods, sometimes based on new technology, that extends the scope of analytical chemistry into new territory: lower concentrations, new analytes, or difficult sample types. The most famous example of this is the detection of organo-chlorine pesticides at previously undetectable concentrations in a wide range of samples following the development of the electron capture detector by James Lovelock (1958).

If chemicals newly identified in food raise concern among regulators, it is important to be able to gather reliable data on exposure to form a basis for risk assessment. The gathering of reliable data requires analytical methods with known performance characteristics to ensure that risk assessments are based on accurate quantitative measurements, and that regulations can be reliably and defensibly enforced. This means that the new analytical method based on new technology, applied by 'rocket scientists', must be converted to, or replaced by, a method that can be used more widely, and that the performance of this method can be characterized. One way of doing this is to produce a 'standard method'.

Standard methods

So what are the drivers for converting a research method into a standard method and what is the process for achieving the conversion? The top level drivers for the production of standard methods of analysis are those that apply to *standardization per se* (CEN, n.d.). A standard can provide a definition of consensus among interested parties and stakeholders (e.g. vendors, buyers, enforcement agencies, academia, etc.) and possibly best practice in the sector to which it is applied, support free trade within its domain, and reduce costs associated with instability, fragmentation and overlap in practice and responsibilities. The economic benefits of standardization may be large. For example standardization (as a whole, not just in analytical methods) has been estimated to account for 13% of growth in labour productivity between 1948 and 2005 and made an annual contribution £ 2.5 billion to the UK economy. Studies in Germany, France and Denmark suggest that standardization benefits Gross Domestic Product by approximately 1% (CEN, n.d.).

For analyses undertaken in support of food consumer protection in the legal context of 'free trade' [e.g. between World Trade Organisation members or within the European Union (EU)] there is a requirement for *agreed* methods of analysis or methods with *agreed* performance. If imported goods are to be rejected on the basis of non-compliance with regulations, there needs to be an *agreement* on how these controls are enforced in different countries and *agreement* that this is done in a uniform manner. Within the EU, there are agreements for free trade, but there are 27 different competent authorities with an even larger number of National Reference Laboratories and many more official control laboratories. Hence there is a need for standardization of some kind to remove 'instability, fragmentation and overlap' (CEN, n.d.). Therefore EU Directive 85/591/EEC: 'Introduction of Community methods of sampling and analysis' and Regulation 882/2004: 'Official control of foodstuffs' have been introduced to enforce feed and food law, animal health and welfare rules and monitor and verify that the relevant requirements therein are fulfilled by business operators at all stages of production, processing distribution and processing within the EU.

Directive 85/591/EEC concerning the introduction of community methods of sampling and analysis for the monitoring of foodstuffs intended for human consumption says in the preamble: 'Whereas the *methods of* sampling and *analysis* used for this purpose can have direct repercussions on the establishment and functioning of the common market; whereas they *should,* therefore, *be harmonized*. . .'. Article 2 goes on to state that 'the introduction of the measures provided for in Article 1 (1) *shall not preclude* Member States *from using other tested and scientifically valid methods* provided that this does not hinder the free movement of products recognized as complying with the rules by virtue of community methods'. However, in the event of differences in the interpretation of results, those obtained by the use of community methods shall be determinant.

As designing and describing methods of analysis or other technical product specifications is not the primary objective of EU policy making, the current approach, which is fully in line with the 'New Approach' (Council Resolution of 7 May 1985 on a new approach to technical harmonization and standards) and the 'Better Regulation' initiative, is to leave it to the European standardization system to develop standards in support of EU policies and legislation. This general approach is also reflected in Regulation 882/2004 where a 'hierarchy' of methods of analysis to be used for

official control purposes is described. Article 11 lays down that sampling and analysis methods used in the context of official controls shall comply with relevant community rules or,

(a) if no such rules exist, with internationally recognized rules or protocols, for example those that the European Committee for Standardization (CEN) has accepted or those agreed in national legislation, or,
(b) in the absence of the above, with other methods fit for the intended purpose or developed in accordance with scientific protocols.

The practical advantages associated with the use of standard methods are:

- They are generally methods that are based on widely accepted principles with sufficient validation data and proven transferability to other laboratories.
- They give a clear description with all details including calibration and calculation.
- They have been agreed by the interested parties and stakeholders.
- Standard methods are usually designed to use equipment and techniques that can be accessed by as wide a range of laboratories as possible.
- Accreditation bodies would only need to review a standard method once in detail.
- Many standards are available in more than one language (CEN produces standards in English, German and French).
- They are particularly useful if it is necessary to demonstrate to, and gain agreement from, all stakeholders that actions based on the results of analytical tests are a necessary protection for consumers rather than potential barrier to free trade.
- They are also a starting point for new laboratories, for laboratories involved with a wide range of functions where a variety of analyses are undertaken.

However, there are some disadvantages associated with standardization and standard methods. For example, the process of converting a good analytical method into a *standard* method can be laborious. The basis for any method used to enforce food safety regulatory requirements is providing evidence that a method delivers valid results. A newly developed and single-laboratory validated method will then normally be subject to formal validation by collaborative trial, usually organized by the method provider or sometimes by a standards body such as AOAC, CEN or the like using the agreed international protocols (ISO 5725-2, 1994; Horwitz, 1995). The performance data from such an exercise can be used to give a firmer indication of fitness for purpose across a number of laboratories. Valid sets of results from at least eight laboratories are usually required for such a ring trial to give sufficient data to calculate repeatability and reproducibility. The method may then go through a process of being considered, approved and eventually issued as a standard. The process of converting a method that is considered to have demonstrated sufficiently good performance into as a *standard method* will usually take at least 2 years (Figure 1).

Prescribing (usually by legislation) or agreeing (trade bodies) on the use of a standardized method has been the traditional approach to harmonization in food and feed analysis which was applied by the Codex Alimentarius Commission and EU. While this approach may be simpler for all parties when deciding which of a range of possible analytical methods to use, prescribing a specific method of analysis means: the analyst is denied freedom of choice and thus may be required to use a standard method in some situations where there may be other methods which could do a better job; the use of automation and up to date methods is inhibited; it is administratively difficult to change a method found to be unsatisfactory or inferior to another (often new) method.

Alternative approaches to select analytical methods for official control purposes

A current alternative to the use of agreed or prescribed analytical methods is the use of analytical methods with agreed or prescribed performance. There are a number of ways in which the performance of methods may be described which may be particularly useful for different stakeholders. Broadly, method performance might be described using: analytical 'criteria' such as those traditionally used by analysts (the criteria approach), measurement uncertainty as applied in analytical chemistry since around the turn of the century (the standard uncertainty approach), or by evaluating the consequences of measurement uncertainty for stakeholders (the uncertainty profile approach). The strengths and features of these approaches are discussed here.

The criteria approach

Under this approach a range of acceptable values is defined for a number of parameters that describe the performance of the method. Typically these are parameters used by analysts working in the laboratory and may include technique-specific technical parameters such as chromatographic resolution (Ettre, 1993) or DNA quality (The European Network of Genetically Modified Organism Laboratories, 2008)

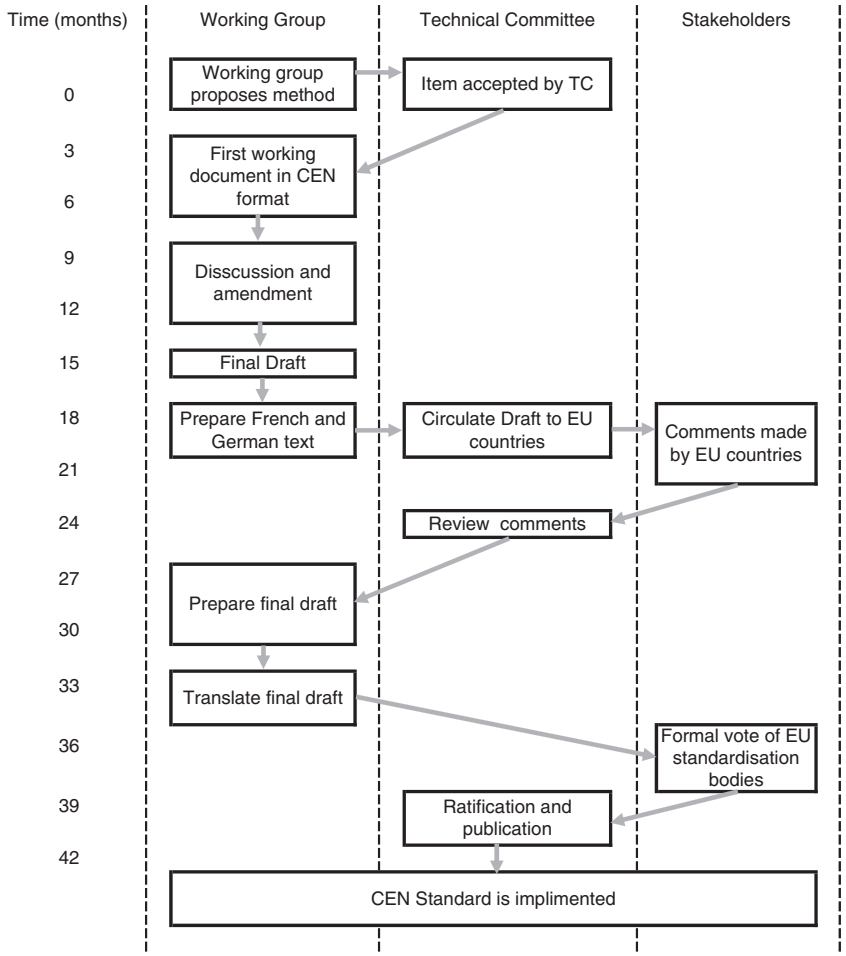

Figure 1 Process of making a European Standard (EN) in CEN National Members (European Union countries plus Croatia, Norway, Iceland and Switzerland).

or higher level general parameters such as repeatability, and reproducibility standard deviations (Currie & Svehla, 1994), and recovery (Thompson *et al.*, 1999).

The earliest example of the criteria approach in European legislation is found in Directive 98/53/EC (EU, 1998). This Directive, contains 'recommended' criteria for a number of analytes. For the analysis of Aflatoxin M1 in liquid milk they are: that 'blank' samples should give a 'negligible' response; that recovery should be between 50% and 120% for samples that contain between 0.01 and 0.5 $\mu g\,L^{-1}$ analyte and between 70% and 110% for samples that contain more than 0.5 $\mu g\,L^{-1}$ analyte; and it is recommended that the reproducibility relative standard deviation should be no greater than that given by the unmodified Horwitz equation (Horwitz *et al.*, 1980) with a 'maximum permitted value' not greater than twice the value given by the Horwitz equation. These criteria align well with those used to assess the performance of a method by collaborative trial, and probably represent the minimum set of criteria that can be used to control method performance effectively (one criterion to do with bias and one to do with variation). They lead to the position where *any* method that has performed reasonably well in a collaborative trial can be used for official control.

Sometimes a more extensive set of criteria are applied. For example Regulation 1883/2006 (EU, 2006a) (methods for sampling and analysis for dioxins in some food stuffs) uses the criteria approach and, for 'confirmatory methods' sets out the basic (one for precision, one for bias) criteria that reproducibility standard deviation shall be < 15 % and trueness shall lie within ± 20 % for dioxins toxic equivalents at 0.5 $\times$, 1 $\times$ and 2 $\times$ 'the level of interest'.

However, many other criteria are also set out in the Regulation:

- An upper limit for 'detectable quantities'.
- A requirement for 'high selectivity'.

• Recovery must be between 60% and 120 % (unless the congener makes a small enough contribution to the total toxic equivalents).
• A requirement that gas chromatographic resolution of two particular isomers be sufficiently high.
• A maximum limit for the difference between the 'upper-bound' and 'lowerbound' estimates [effectively another target for limit of detection (LOD)].

In addition to the criteria for method performance a number of methodological and quality control procedures are prescribed, which may be thought of as fixed, or standardized parts of the analytical method.

Another important example of the criteria approach is Decision 2002/657/EC (EU, 2002) which gives both criteria for methods and designs for the experiments necessary to show that criteria are met.

The criteria approach as used in EU legislation gives greater flexibility than the standard method approach, mainly by removing the bureaucratic barriers to the use of new or modified methods. This avoids the situation of having many good methods of analysis available, which meet requirements as regards method performance characteristics, but which are not considered by Codex, EU or other bodies simply because of time and organizational constraints. The amount of laboratory work necessary to demonstrate that a method meets criteria is not much reduced compared with that generally used in the standard method approach if reproducibility standard deviation is explicitly included as a criterion because some kind of collaborative trial is still required. The time-consuming standardization process is not necessary; however, sometimes to the detriment of making only a limited method documentation available.

There remain some limitations to, and potential problems for, the criteria approach:

• The criteria approach cannot be directly applied to empirical methods because results must be comparable to apply it.
• It can be difficult for all stakeholders to tell which of two methods, each with different values for six or seven factors used as criteria, is best for them.
• As the number of criteria increases there is a choice between more adequate methods 'failing'[1] or having to widen the acceptable range of values for each factor which may allow inadequate methods to 'pass'.

If these issues are a problem, they can be dealt with using measurement uncertainty to assess the performance of analytical methods.

The standard uncertainty approach

Under this more recent approach to method validation, the performance of a method is described by the *expected standard uncertainty*[2] associated with measurement results. The standard uncertainty associated with a method is a single parameter that gives an estimate of the combined effect of the individual factors that describe the method on how far we can expect a measurement result to lie from a true concentration. The tipping point from traditional multi-criteria method validation towards modern approaches to validation occurred with the publication of the second edition of the Eurachem guide on analytical measurement uncertainty (Eurachem, 2000) and the IUPAC/ISO/AOAC harmonized guidelines for Single Laboratory Validation of Analytical methods (Thompson *et al.*, 2002). The harmonized guidelines retained traditional validation parameters but introduced measurement uncertainty as a central part of method validation and included the very valuable observation that '... method validation is tantamount to the estimation of measurement uncertainty'.

Standard uncertainty can be estimated using two broad classes of method: first, those based on the bottom-up approach:

1. Describe the method.
2. Identify the individual sources of uncertainty associated with each component of the method.
3. Carry out experiments to get estimates of the size the uncertainty associated with each component of the method.

If there is a 95% probability that a sufficiently good method will produce results that meet one criterion, then the probability that the method will produce results that meets all seven such criteria may be as low as 70%. ($0.95^7 = 0.70$).

Strictly *standard uncertainty* is a number attached to an individual measurement result that describes the size of the uncertainty associated with a result. Methods do not have a standard uncertainty. However, we can use the concept of measurement uncertainty to describe method performance by estimating an *expected standard uncertainty*: the size of uncertainty we can expect to be associated with results when we use the method, provided that internal quality control measures are in place and the method is shown to be under statistical control. When the term measurement uncertainty is applied to a method, as it often is, it should be understood to mean expected measurement uncertainty.

4. Combine the uncertainties mathematically to get an estimate of the standard uncertainty.

And second, those based on the top-down approach:

1. Describe the method.
2. Identify the individual sources of uncertainty associated with each component of the method.
3. Use the method to undertake measurements under conditions that allow all of the method components to vary over their natural range.
4. Use the observed variation in results as an estimate of measurement uncertainty.

A critical review (EU, 2004) of the bottom-up approach describes it as 'absurd and budget busting', and helpfully provides a fifth Step, 'You can come back later and add in those factors that you initially overlooked or which are pointed out to you by your colleagues or by your friendly assessor months after the report has been delivered and forgotten'.

In practice, the expected standard uncertainty may be most reliably estimated with the minimum of mathematical fuss using a mostly top-down approach whereby the size of the measurement variation is estimated using a collaborative trial (ISO 5725-2, 1994) or a single-laboratory (Horwitz, 1995) study, and the size of the uncertainty associated with bias is estimated using a certified reference material, if available, or spiking experiments. Another method that may give a very simple and reliable estimate of uncertainty for a single laboratory, but may take some time to achieve (because a minimum of about eight sets of results are needed), is to use results from proficiency testing to estimate standard uncertainty (Castle *et al.*, 2004).

Decision 2002/657/EC (EU, 2002) does not include approaches to validation based on consideration of expected standard uncertainty. However, subsequent 'Guidelines for the Implementation of Decision 2002/657/EC' (EU, 2008) state that 'when determined correctly by systematically taking into account all relevant influencing factors possibly affecting the measurement results, the within-laboratory reproducibility can be regarded as a good estimator for the combined measurement uncertainty of the individual methods. Further prerequisites are the use of recovery-corrected data and the fact that the uncertainty of the recovery was taken into account the one or the other way'. The guidelines also say that factors used to calculate expanded uncertainty should be the same as those used to calculate the decision limit CCα and the detection capability CCβ in 2002/657/EC (EU, 2002).

Although we have said here that expected standard uncertainty is a single parameter that describes method performance it is important to remember that, in general, the absolute size of measurement uncertainty varies with the concentration of the analyte. In analytical chemistry the form of the relation is usually given using a two-component model (Rocke & Lorenzato, 1995; Eurachem, 2000):

$$u(x) = \sqrt{u_0^2 + RSU^2.x^2} \quad (1)$$

where $u(x)$ is the expected standard uncertainty associated with results of measurements at concentration x; u_0 is the fixed component of uncertainty, the value to which standard uncertainty tends as concentration approaches zero; and *RSU* is the proportional component of uncertainty, the value that relative standard uncertainty approaches as analyte concentration increases. Hence, a criteria can be set such that at a concentration x (e.g. 'the concentration of interest' such as a legislative limit), the expected measurement uncertainty must be no greater than that calculated using Equation (1) with u_0 *and RSU* set to maximum acceptable values.

An example of the standard uncertainty approach applied to methods for the official control of lead, cadmium, mercury, inorganic tin, 3-monochloropropane-1,2-diol (3-MCPD) and benzo(a)pyrene in foodstuffs is found in Commission Regulation 333/2007 (EU, 2007). The approach may be used where a limited number of fully validated methods of analysis exist. Here, an equation of the form of Equation (1) [Equation (2) below) is used to define a maximum standard uncertainty:

$$uf = \sqrt{(LOD/2)^2 + (aC)^2} \quad (2)$$

. The equation is implicitly based on the assumption that LOD has been calculated as $2 \times u_0$ [see Equation (1)]. A range of different values for α are given (Table 1) which are a little less than those produced by the modified Horwitz equation (Thompson, 2000). For 3-MCPD an upper limit for LOD of $5\,\mu g\,kg^{-1}$ is given in the analytical criteria

Table 1 Factors used to set maximum standard uncertainty in European legislation

Concentration (C, $\mu g\,kg^{-1}$)	α
≤ 50	0.20
51–500	0.18
501–1000	0.15
1001–10 000	0.12
$> 10\,000$	0.10

section. Hence, the maximum allowed standard uncertainty for the measurement of 3-MCPD is given by Figure 2. The discontinuity in the curve is caused by 'α' changing from 0.20 to 0.18 when moving from 50 to 51 $\mu g\,kg^{-1}$ (Table 1).

An equation with the form of Equation (2) and the values for α shown in Table 1 are used to define the maximum standard uncertainty for official control by chemical analysis for all purposes where the 'standard uncertainty' approach is included in legislation. However, this can sometimes lead to problems. For example (EU, 2006b) includes the standard uncertainty approach for methods for the official control of mycotoxins in foodstuffs with the usual equation for maximum standard uncertainty and values for α. However, no upper limit for LOD is specified anywhere in the legislation because 'The detection limits of the methods used are not stated as the precision values are given at the concentrations of interest'. This means that, strictly, there is no upper limit to the expected standard uncertainty for an analytical method that could, under certain circumstances, be used in accordance with the Regulation (EU, 2006b) (usually the LOD of the method is used instead).

The main advantage of the use of the expected standard uncertainty as the validation parameter is that it expresses the performance of a method in a single parameter on a scale that matters to most users of analytical results. In general the users of results do not care about the linearity of a calibration curve, the repeatability of results or the recovery associated with them, but they do want to know, or at least may understand the meaning of, how far away a measurement result might be from a true concentration.

A second advantage is that the standard uncertainty (or relative standard uncertainty) associated with a measurement result is a natural scale for expressing analytical performance that does not contain any hidden assumptions about the use to which measurement results may be put. This is in contrast to other quantities that can be used to express measurement performance such as LOD or limit of quantification.

In general we should expect that no more laboratory work should be necessary to estimate standard uncertainty than is necessary to evaluate the factors used in the criteria approach.

The main disadvantage of the standard uncertainty approach is the perceived complexity, for analysts, of *calculating* estimates of standard uncertainty using methods described in the most cited guides (ISO/IEC Guide 98, 1995; EU, 2002) which tend to focus on the more demanding bottom-up approaches. However, a range of options, including some very simple approaches for calculating standard uncertainty can be found within these and other guides:

- Using results generated during single laboratory validation (Horwitz, 1995; EU, 2004).
- Using proficiency test results (Castle *et al.*, 2004).
- Using collaborative trial results (Eurachem, 2000; EU, 2004).

There is perhaps a remaining issue: that it is not always obvious how stakeholders should use estimates of standard uncertainty when interpreting results though clear and relevant guidance exists (EU, 2004), and assessing whether a method is likely to give a fit for purpose result. The uncertainty profile is an attempt to provide a simple method to do this.

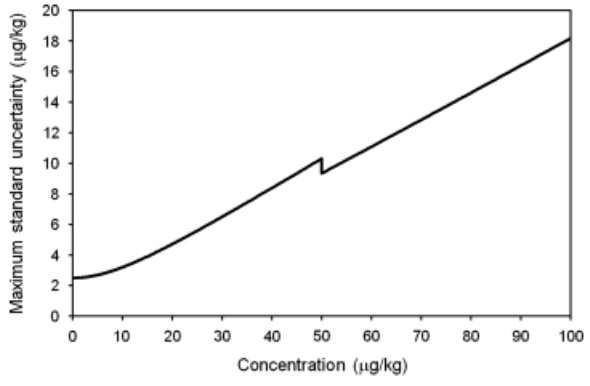

Figure 2 Maximum standard uncertainty for methods used for the official control of 3-monochloropropane-1,2-diol in food.

The uncertainty profile approach

An uncertainty profile (Macarthur *et al.*, 2010) is a graphical representation of the size of the expected measurement uncertainty associated with a method. An uncertainty profile is produced by plotting the expected mean measurement result, and a confidence interval within which a high proportion (usually 90% or 95%) of results can be expected to lie, across a range of estimated true concentrations. The profile can then be used to give an estimate of the range of concentrations for which the method can be expected to give fit for purpose results (sufficiently small measurement uncertainty), and other quantities such as the critical level (lowest measurement result that reliably demonstrates that the analyte is present above a threshold concentration), the limit of control (lowest true concentration that will reliably

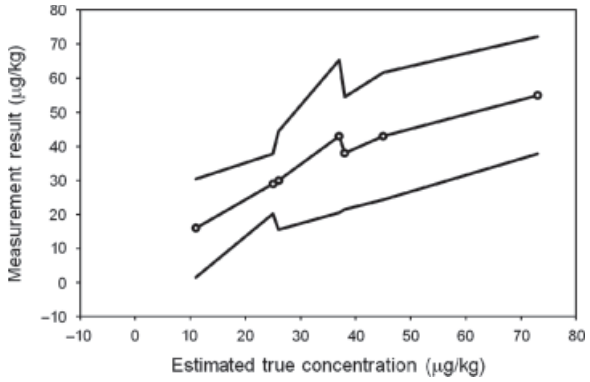

Figure 3 Uncertainty profile for a method for the measurement of 3-monochloropropane-1,2-diol in food. —O—, average of measurement results; ——, 95% confidence interval for measurement results.

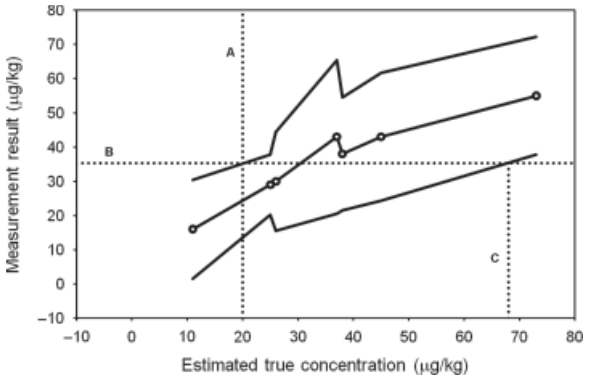

Figure 4 Use of the uncertainty profile to estimate the capability of a method for the measurement of 3-monochloropropane-1,2-diol (3-MCPD) in food to control the presence of 3-MCPD against a legislative limit. —O—, average of measurement results; ——, 95% confidence interval for measurement results; A ....., maximum limit for 3-MCPD; B ....., critical measurement result for demonstrating non-compliance with the limit; C ....., highest concentration of 3-MCPD that might not ($P = 0.025$) produce a result above the critical measurement result.

give a measurement result above the critical level) and limit of assurance (highest true concentration that will reliably give a measurement result below the critical level). An uncertainty profile can also be used to compare different measures of method validity such as comparison against criteria for precision and bias or to a target standard uncertainty. A method for constructing and using an uncertainty profile is given in detail in Macarthur *et al.* (2010).

Figure 3 gives an example of an uncertainty profile calculated using the results of a collaborative trial for a method for the measurement of 3-MCPD in food (Brereton *et al.*, 2001). The lines describing the profile are disjointed. This is often the case for uncertainty profiles based on collaborative trials, which rely on estimates of reproducibility standard deviation, from a relatively small number of laboratories, that vary between different concentrations.

The maximum limit for 3-MCPD in soy sauce and hydrolysed vegetable protein is 20 µg kg^{-1} (EU, 2006c). The uncertainty profile can be used to find the expected lowest measurement result that demonstrates that a sample does not comply with legislation (critical measurement result for demonstrating non-compliance, line B in Figure 4, 35 µg kg^{-1}), and an expected upper limit for the concentration of 3-MCPD that may be in a sample that has produced a result at the critical level (line C in Figure 4, 68 µg kg^{-1}).

Similarly, the profile can be used to test whether a method gives results that are equivalent to a method that meets criteria for performance based on precision and bias. For example Regulation 333/2007 sets the following criteria for methods for the measurement of 3-MCPD in foods.

Recovery between 75% and 110% at all concentrations and upper limits for 'precision'[3] of:

4 µg kg^{-1} at a concentration of 20 µg kg^{-1}.
6 µg kg^{-1} at a concentration of 30 µg kg^{-1}.
7 µg kg^{-1} at a concentration of 40 µg kg^{-1}.
8 µg kg^{-1} at a concentration of 50 µg kg^{-1}.
15 µg kg^{-1} at a concentration of 100 µg kg^{-1}.

The lower limit L of the range within which analytical results can be expected, for a method that meets the criteria, is given by

$$L = x.R_L - z_\gamma.\sigma$$

and the upper limit U is given by

$$U = x.R_U - z_\gamma.\sigma \quad (3)$$

where x is the estimated true concentration, R_L and R_U are lower and upper limits to recovery, z_γ is a coverage factor taken from the normal distribution associated with the coverage probability γ, and σ is the upper limit for precision at concentration x. For a 95% confidence interval γ is equal to 2.

A problem with these criteria is that the condition (reproducibility, intermediate, repeatability, sub-repeatability) under which precision is to be estimated is not specified. Reproducibility is assumed here.

Hence, a profile for within which we can expect to see 95% of results is given by a linear interpolation of the points (x, L, U) shown in Table 2.

Figure 5 shows a comparison between the uncertainty profile derived from collaborative trial results and the target profile derived from the analytical criteria (Table 2). The uncertainty profile for the method is outside of the target profile for concentrations $< 40\,\mu g\,kg^{-1}$. Hence, based on the collaborative trial results, the method cannot be expected to produce results that are consistent with the analytical criteria across the full concentration range, but may produce results that are consistent with the criteria for samples that contain more than $40\,\mu g\,kg^{-1}$ 3-MCPD.

A comparison between the uncertainty profile and criteria for standard uncertainty can also be made. Given a maximum limit for standard uncertainty given by *uf* [Equation (2)] based on an LOD ($5\,\mu g\,kg^{-1}$ in this case) and 'α' (Table 1), then the lower limit L of the range within which analytical results can be expected, for a method that meets the criteria for standard uncertainty, is given by:

$$L = x - z_\gamma.uf$$

and the upper limit is given by:

$$L = x + z_\gamma.uf \quad (4)$$

where x is the estimated true concentration, z_γ is a coverage factor taken from the normal distribution associated with the coverage probability γ.

Table 2 Uncertainty profile based on analytical criteria for methods for the measurement of 3-monochloropropane-1,2-diol in food (Regulation 333/2007)

Estimated true concentration (x, $\mu g\,kg^{-1}$)	Lower limit (L, $\mu g\,kg^{-1}$)	Upper limit (U, $\mu g\,kg^{-1}$)
20	7	30
30	10.5	45
40	16	58
50	21.5	71
100	45	140

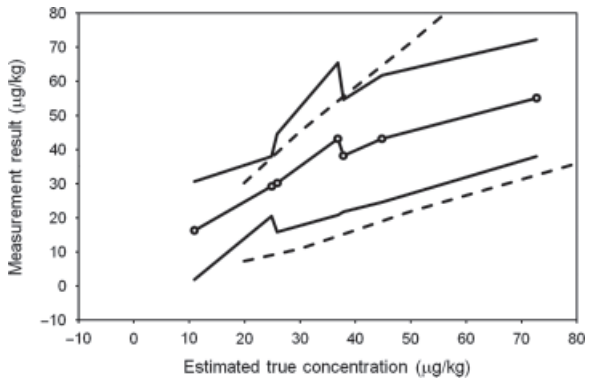

Figure 5 Comparison of uncertainty profile for measurement of 3-monochloropropane-1,2-diol in food to analytical criteria in Regulation 333/2007. —O—, average of measurement results; ——, 95% confidence interval for measurement results; – – –, 95% confidence interval for a method that just meets criteria for recovery and precision.

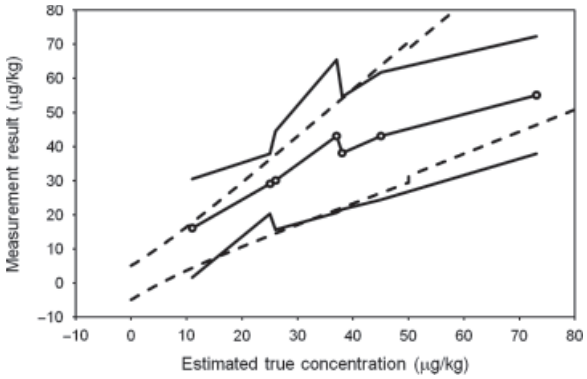

Figure 6 Comparison of uncertainty profile for measurement of 3-monochloropropane-1,2-diol in food to criteria for standard uncertainty in Regulation 333/2007. , average of measurement results; —O—, —— 95% confidence interval for measurement results; – – –, 95% confidence interval for a method that just meets legislative requirement for measurement uncertainty.

Figure 6 shows a comparison between the uncertainty profile derived from collaborative trial results and the target profile derived [Equation (4)] from the target standard uncertainty (Table 1) with a coverage factor (z_γ) of 2. Hence, the results produced during the collaborative trial are not consistent with the required standard uncertainty at any concentration: results tended to be too high at low concentrations and too low at high concentrations.

The uncertainty associated with estimates of the concentration of 3-MCPD using the method *in a particular laboratory* was estimated using 17 Food Analysis Performance and Assessment Scheme[4] proficiency test results produced by that laboratory. Proficiency test results are particularly useful because the assigned value in Food Analysis Performance and Assessment Scheme rounds is usually calculated as the robust mean of submitted results. The submitted results typically come from a large number of laboratories using a range of analytical methods. The assigned value is a good estimate of the true concentration

http://www.fapas.com

of analyte in the test materials. Also, because each round uses a new material, the uncertainty associated with the assigned value is included in the observed variation of the difference between a laboratory's results and the assigned values.

For each proficiency test result y the ratio $R = y/a$ was calculated where a was the assigned value for the concentration of 3-MCPD in the test material. The standard deviation (s) and average ($\bar{R}$)of the ratios was also calculated. A lower limit (L) for the range within which results could be expected to lie was given by

$$L = x.(\bar{R} - s.t_{\gamma.n-1})$$

and the upper limit was given by:

$$U = x.(\bar{R} + s.t_{\gamma.n-1}) \quad (5)$$

where $t_{\gamma,\mathrm{n}-1}$ is the value of the inverse t-distribution with $\mathrm{n}-1$ degrees of freedom at the $1-\gamma$th percentile, n is the number of results, and x is estimated true concentration.

A comparison of proficiency test results produced by the laboratory using the method validated by the collaborative trial, and an estimate of uncertainty derived from them, to a target profile derived from the target standard uncertainty (Figure 7) shows that the laboratory does produce results that meet the target for standard uncertainty. Hence, *the laboratory*'s results generated using the method are fit for the purpose of testing samples against the legislative limit because the criterion for standard uncertainty is met.

The uncertainty profile can also be used to directly compare targets for method performance that are expressed on different scales. For example, a plot of the target profile based on analytical criteria for the measurement of 3-MCPD [Equation (3), Table 2] with the target profile based on the target for standard uncertainty [Equation (2), Table 1] shows that the performance target based on criteria is a little more generous then that based on standard uncertainty (Figure 8). This is a general pattern that can be observed across the other analytes mentioned in Regulation 333/2007, and in other legislation giving targets for method performance (e.g. for mycotoxins, EU, 2006d) and for fusarium toxins (EU, 2007).

The main point of examining the use of the uncertainty profile for looking at the performance of the method for the measurement of 3-MCPD in such detail is because it demonstrates how uncertainty profiles can be used to easily express and compare different targets for (analytical criteria, target standard uncertainty), and observations of (performance summary from collaborative trial, proficiency test results), analytical performance (Figures 5–8). Also, the use of the profile to set a critical level for measurement results

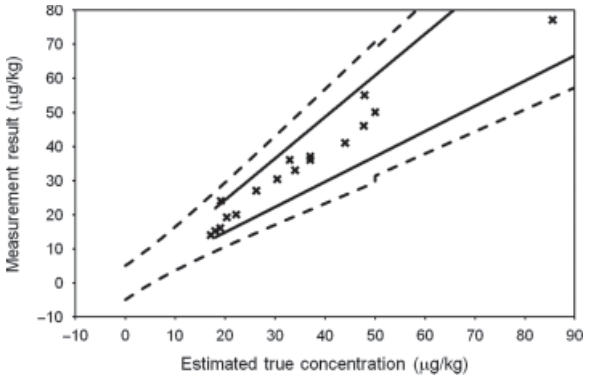

Figure 7 Comparison between proficiency test results for 3-monochloropropane-1,2-diol in food and criteria for measurement uncertainty in Regulation 333/2007. , proficiency test results; ——, 95% confidence interval for measurement results; – – –, 95% confidence interval for a method that just meets legislative requirement for measurement uncertainty.

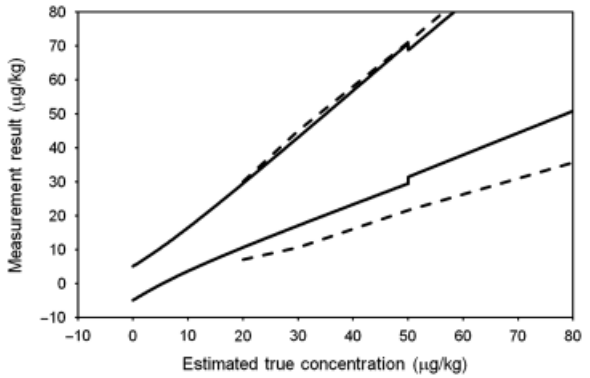

Figure 8 Comparison between analytical criteria and criteria for measurement uncertainty in Regulation 333/2007.

when testing for compliance with legislation (equivalent to CCα in Decision 2002/657/EC, EU, 2002) and assessing the capability of a method to detect non-compliance by finding the highest concentration that might not be detected as non-compliant (equivalent to CCβ in 2002/657/EC) was demonstrated (Figure 4).

Conclusion: which approach to selecting valid methods is best?

The answer suggested by our examination of the relations between approaches to judging the validity of methods is that all approaches: the use of standard methods, the use of methods that meet criteria for traditional validation parameters and approaches based on measurement uncertainty have value.

The wide availability of standard methods which are designed, and demonstrated, to be reliably and economically implementable by a large number of laboratories is a valuable starting point for laboratories who need to produce valid measurement results in support of food safety and legislation. The use of standard methods will not be superseded as long as there remains a demand from non-specialist laboratories that need ready access to methodology that will be reliable and easy to implement 'off the shelf'.

However the use of a standard method is not the end of the story for various reasons. The first reason is that while we can be confident that standard methods are reliable and robust within their scope of application, the users of the method need to provide an objective measure that they are able to apply the method correctly to have confidence in the results produced by standard methods. The second reason is that technology and to an even greater extent 'events' move faster than Standards. Standard methods will not always be there when we need them, or when they are available there may be other methods, which could do a better job. We need to be able to tell when methods will produce results that are good enough to deal with a particular event, and when methods will produce results that are at least as good as a standard method.

The paragraph above mentions several reasons why we might need to know about the performance of an analytical method. And in general 'users' of methods (and results produced by the methods) will want to know about method performance: analysts, risk managers, producers, consumers and legislators, all of whom might at some point need to understand what valid analytical methods can do. So we need to have descriptions of the performance of valid methods of analysis that can be easily understood or translated.

We have shown how describing method performance using measurement uncertainty provides a versatile approach that encompasses, and makes comparable (Figure 9), approaches based on formal validation by collaborative trial, the traditional parameters used by analysts to describe method performance and performance data generated during use of a method. Consideration of measurement uncertainty using approaches such as the uncertainty profile also makes it easy to assess and communicate the practical impact of uncertainty.

The main challenge for the analyst is that application of this approach relies more heavily on calculation and data analysis than on traditional multi-criteria laboratory validation. Largely this is a problem associated with the application of the bottom-up approach to uncertainty (estimating the size of each individual source of uncertainty and then combining mathematically), which may need many experiments undertaken especially to estimate uncertainty followed by many calculations and some algebra. However, in the last 10 years, simpler approaches to estimating uncertainty have become more popular, including some of the approaches described in this paper.

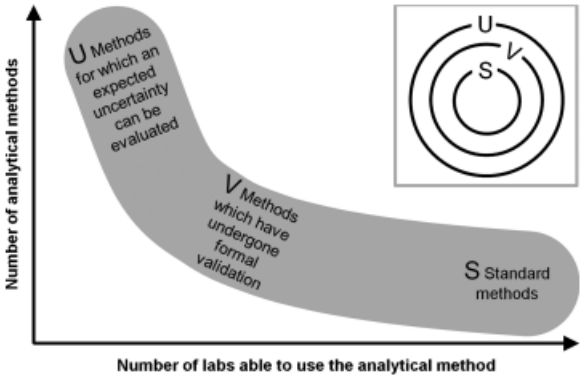

Figure 9 Relations analytical methods with different validation statuses.

Acknowledgement

This work was supported by the Monitoring and Quality Assurance in the Food Supply Chain Network of Excellence (contract no. FOOD-CT-2006-36337) within the 6th framework Topic T5.4.5.1; Quality and safety control strategies for food. The article reflects the authors' personal views and do not necessarily reflect those of either their employers or of the European Commission.

References

Brereton P., Kelly J., Crews C., Honour S., Wood R. (2001) Determination of 3-chloro-1,2-propanediol in foods and food ingredients by gas chromatography with mass spectrometric detection: collaborative study. *Journal of the Association of Official Analytical Chemists International*, **84**, 455–465.

Castle L., Macarthur R., Mead E.M., Read W.A. (2004) Measurement uncertainty associated with overall migration testing. *Food Additives and Contaminants A*, **21**, 256–264.

CEN. (N.d.) CENELC Position Paper. Available at http://www.din.de/sixcms_upload/media/2896/CEN CENELEC_official position paper_revision of the European Standardisation System.pdf [Last accessed 1 November 2010].

Currie L.A., Svehla G. (1994) Nomenclature for the presentation of results of chemical analysis. *Pure and Applied Chemistry*, **66**, 595–608.

Ettre L.S. 1993 Nomenclature for chromatography. *Pure and Applied Chemistry*, **65**, 819–872.

European Union (EU). (1998) Directive 98/53/EC, Laying down the sampling methods and the methods of analysis for the official control of the levels for certain contaminants in foodstuffs. *Official Journal of the European Union*, **L201**, 93–101.

European Union (EU). (2002) Commission Decision 2002/657/EC, implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. *Official Journal of the European Union*, **L221**, 8–36.

European Union (EU). (2004) Report on the relationship between analytical results, measurement uncertainty, recovery factors and the provisions of EU food and feed legislation, with particular reference to community legislation concerning contaminants in food and undesirable substances in feed, Annex 2 Available at http://ec.europa.eu/food/food/chemicalsafety/contaminants/report-sampling_analysis_2004_en.pdf [Last accessed 1 November 2010].

European Union (EU). (2006a) Commission Regulation 1883/2006, laying down methods of sampling and analysis for the official control of levels of dioxins and dioxin-like PCBs in certain foodstuffs. *Official Journal of the European Union*, **L364**, 32–43.

European Union (EU). (2006b) Commission Regulation 401/2006, laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. *Official Journal of the European Union*, **L70**, 12–34.

European Union (EU). (2006c) Commission Regulation 1881/2006, setting maximum levels for certain contaminants in foodstuffs. *Official Journal of the European Union*, **L364**, 558–577.

European Union (EU). (2006d) Commission Regulation 401/2006, laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. *Official Journal of the European Union*, **L70**, 12–34.

European Union (EU). (2007) Commission Regulation 333/2007, laying down the methods of sampling and analysis for the official control of the levels of lead, cadmium, mercury, inorganic tin, 3-MCPD and benzo(a)pyrene in foodstuffs. *Official Journal of the European Union*, **L88**, 29–38.

European Union (EU). (2008) European Commission, SANCO/2004/2726-rev 4 December. Guidelines for the implementation of decision 2002/657/EC. Available at http://ec.europa.eu/food/food/chemicalsafety/residues/cons_2004-2726rev4_en.pdf [Last accessed 1 November 2010].

Eurachem. (2000) Quantifying uncertainty in analytical measurement. 2nd edition. Available at http://www.measurementuncertainty.org/ [Last accessed 1 November 2010].

Horwitz W. (1995) Protocol for the design, conduct and interpretation of method performance studies. *Pure and Applied Chemistry*, **67**, 331–343.

Horwitz W., Kamps L.R., Boyer K.W., Boyer (1980) Quality assurance in the analysis of foods and trace constituents. *Journal of the Association of Official Analytical Chemists*, **63**, 1344–1354.

ISO/IEC Guide 98. (1995) Uncertainty of measurement–Part 3: Guide to the expression of uncertainty in measurement (GUM: 1995). Availability at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50461.

ISO 5725-2. (1994) Official Journal of the European Union L88. Accuracy (trueness and precision) of measurement methods and results–Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. Available at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=11834.

Lovelock J.E. (1958) A sensitive detector for gas chromatography. *Journal of Chromatography A*, **1**, 35–46.

Macarthur R., Feinberg M., Bertheau Y. (2010) Construction of measurement uncertainty profiles for quantitative analysis of genetically modified organisms based on interlaboratory validation data. *Journal of the Association of Official Analytical Chemists*, **93**, 1046–1056.

Rocke D.M., Lorenzato S. (1995) A two-component model for measurement error in analytical chemistry. *Technometrics*, **37**, 176–184.

The European Network of Genetically Modified Organism Laboratories. (2008) Definition of minimum performance requirements for analytical methods of GMO testing. Available at http://gmo-crl.jrc.ec.europa.eu/doc/Min_Perf_Requir_Analyt_methods_131008.pdf [Last accessed 1 November 2010].

Thompson M. (2000) Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing. *Analyst*, **125**, 385–386.

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*, **71**, 337–348.

Thompson M., Ellison SLR., Wood R. (2002) Harmonized guidelines for single-laboratory validation of methods of analysis. *Pure and Applied Chemistry*, **74**, 835–855.